Airborne hyperspectral imaging for multisensor data fusion

Flybåren hyperspektral avbildning for multisensorisk datafusjon

Agnieszka Kinga Kuras

Airborne hyperspectral imaging for multisensor data fusion

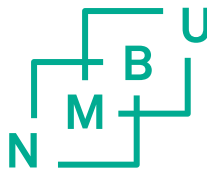Flybåren hyperspektral avbildning for multisensorisk datafusjon

Philosophiae Doctor (PhD) Thesis

Agnieszka Kinga Kuras

Norwegian University of Life Sciences

Faculty of Science and Technology

Ås (2022)

Supervision Team

Main supervisor
Professor Ingunn Burud
>Faculty of Science and Technology (REALTEK),
>Norwegian University of Life Sciences (NMBU),
>Ås, Norway

Co-supervisor
Professor Thomas Kringlebotn Thiis
>Faculty of Science and Technology (REALTEK),
>Norwegian University of Life Sciences (NMBU),
>Ås, Norway

Co-supervisor
Dr. Maximilian Brell
>Remote Sensing and Geoinformatics,
>German Research Centre for Geosciences (GFZ),
>Potsdam, Germany

Co-supervisor
Dr. Christian Rogass
>Helmholtz Centre for Environmental Research,
>Leipzig, Germany

Evaluation Committee

First opponent
Professor Nektarios Chrysoulakis
>Director of research – Head of RSLab,
>Foundation for research and technology (FORTH),
>Crete, Greece

Second opponent
Dr. Misganu Debella-Gilo
>Chief Engineer, Division of survey and statistics, geomatics,
>The Norwegian Institute of Bioeconomy Research (NIBIO),
>Ås, Norway

Committee coordinator
Associate Professor Oliver Tomic
>Faculty of Science and Technology (REALTEK),
>Norwegian University of Life Sciences (NMBU),
>Ås, Norway

Summary

Multisensor data fusion demand in Earth observations is constantly increasing thanks to technological advances and the willingness to explore the Earth in a multidisciplinary way. Recently hyperspectral imaging has become a promising tool for Earth monitoring purposes but has also emerged as suitable for fusion with other remote sensors for various applications.

This dissertation examines different types of multisensor data fusion, such as feature-level and application-level fusion, where each application is based on hyperspectral imaging at the airborne scale. In feature-level data fusion, hyperspectral imaging is combined with LiDAR (Light Detection and Ranging) to analyze urban environments, mainly focusing on urban land cover classification and implementing deep learning algorithms. In contrast, application-level data fusion presents the integration of hyperspectral imaging with magnetic data for material characterization of geologic complexes in remote and harsh environments, such as Greenland.

This PhD thesis focused on enhancing analysis outcomes by combining hyperspectral imaging with other sensors and precisely selecting applications in which one sensor is insufficient to obtain the required parameters.

The analysis of feature-level data fusion for hyperspectral and LiDAR data began with a detailed review of sensor key characteristics most representative of urban land cover analysis. These features were intended to segment land cover classes by considering 2D and 3D convolutional operations, where 2D convolutions involve spatial information and 3D convolutions add a spectral dimension allowing the inclusion of information about the interrelation of hyperspectral bands. The study on feature-level data fusion was completed with a multitemporal analysis, where a general framework was proposed towards automatical updating a local urban database.

The other part of the dissertation was based on the fusion of sensors operating in different feature vectors with a common factor: identifying iron and its magnetic properties. Iron in hyperspectral imaging also has distinct absorption features recognizable at the relatively low spatial resolution. Moreover, it is the only chemical

element capable of maintaining magnetic properties, which is the main aim of magnetic surveys.

This dissertation has contributed new approaches to various feature-level and application-level multisensor data fusion exploitations confirming the great potential and versatility and showing future directions of multidisciplinary research using remote sensing methods for Earth observation.

Sammendrag

Etterspørselen etter multisensorisk datafusjon i jordobservasjoner øker stadig, som følge av teknologiske fremskritt og viljen til å utforske jorden på en tverrfaglig måte. Nylig har hyperspektral avbildning blitt et lovende verktøy for jordovervåkingsformål, of viser seg også å være egnet for fusjon med andre eksterne sensorer for ulike applikasjoner.

Denne avhandlingen undersøker ulike typer multisensorisk datafusjon, slik som på egenskapsnivå- og applikasjonsnivå, hvor hver anvendelse er basert på flybåren hyperspektral avbildning. I datafusjon på egenskapsnivå kombineres hyperspektral avbildning med LiDAR (Light Detection and Ranging) for å analysere urbane miljøer, hovedsakelig med fokus på klassifisering av urbant land-dekke og implementering av dyplæringsalgoritmer. Datafusjon på applikasjonsnivå derimot, pressenterer integrasjonen av hyperspektral avbildning med magnetiske data for materialkarakterisering av geologiske komplekser i tøffe miljøer, som Grønland.

I denne doktorgradsavhandlingen har jeg fokusert på å forbedre analyseresultatene ved å kombinere hyperspektral avbildning med andre sensorer og nøyaktig velge applikasjoner der én sensor ikke er tilstrekkelig til å oppnå de nødvendige parameterne.

Analysen av datafusjon på egenskapsnivå for hyperspektrale og LiDAR-data begynte med en detaljert gjennomgang av de sensornøkkel-karakteristikker som er mest representative for analyse av urbant land-dekke. Disse egenskapene var ment å segmentere land-dekkeklasser ved å vurdere 2D- og 3D-konvolusjonsoperasjoner, der 2D-konvolusjoner involverer romlig informasjon og 3D-konvolusjoner legger til en spektral dimensjon som tillater inkludering av informasjon om sammenhengen mellom hyperspektrale bånd. Studien om datafusjon på egenskapsnivå ble fullført med en multitemporal analyse, der et generelt rammeverk ble foreslått for automatisk oppdatering av en lokal urban database.

Den andre delen av avhandlingen er basert på fusjon av sensorer som opererer i forskjellige egenskapsvektorer med en felles faktor: å identifisere jern og dets magnetiske egenskaper. Jern i hyperspektral avbildning har distinkte absorpsjonsegenskaper som også kan gjenkjennes ved den relativt lave romlige

oppløsningen, men er også det eneste kjemiske elementet som er i stand til å opprettholde magnetiske egenskaper, som er hovedmålet for magnetiske undersøkelser.

Denne avhandlingen har bidratt med nye tilnærminger til ulike multisensor datafusjonsutnyttelse på egenskapsnivå og applikasjonsnivå, som bekrefter det store potensialet og allsidigheten og viser fremtidige retninger for tverrfaglig forskning ved bruk av fjernmålingsmetoder for jordobservasjon.

Nøkkelord: multisensorisk datafusjon, hyperspektral avbildning, LiDAR, urban miljø, tøft miljø, dataintegrasjon, dyplæring, maskinlæring

Dedicated to my family

Acknowledgments

Contents

List of Figures

List of Tables

List of Abbreviations

CNN	Convolutional Neural Network
CRNN	Convolutional Recurrent Neural Network
DEM	Digital Elevation Model
DL	Deep Learning
HL-Fusion	Hyperspectral and LiDAR data fusion
HS	Hyperspectral
HSM Integration	Hyperspectral and magnetic data integration
LiDAR	Light Detection and Ranging
MCC	Matthew's Correlation Coefficient
MIFD	Modified Iron Feature Depth Index
ML	Machine Learning
nDSM	Normalized Digital Surface
ResU-Net	Residual U-Net model architecture
RF	Random Forest
RNN	Recurrent Neural Network
SAR	Synthetic Aperture Radar
SVM	Support Vector Machines
SWIR	Shortwave-Infrared
VIS	Visible light
VNIR	Visible and Near-Infrared

CHAPTER 1

Introduction

1.1. Motivation

In the field of remote sensing, hyperspectral (HS) imaging has gained attention in spatiotemporal Earth observations, ranging from monitoring of urban and suburban land cover (Jürgens 2001), land cover classifications with deep learning methods (Kuras et al., 2022c), agriculture monitoring (Wei et al., 2021), and geological characterization (Kuras et al., 2022b), among others. HS sensor is a passive camera acquiring spectral information in the reflective spectral range from visible (VIS) to short-wave infrared (SWIR) from 400 to 2500 nm (Signoroni et al., 2019; Ben-Dor 2001). The information in this spectral range of the electromagnetic spectrum enables the retrieval of chemical and physical properties of a material based on its absorption and feature shapes (Ortenberg 2011).

The application diversity and versatility of HS sensors do not exclude other analysis methods and technologies; instead, HS data can support the analysis of the target of interest by implementing multisensor data fusion. Information on materials in terms of chemical properties is mainly analyzed by HS imaging; the structure, texture, and form are determined by LiDAR scanning, and specific material information, such as magnetic properties, can be retrieved using magnetic surveys. Hence, complex surface properties require the application of multiple sensors. However, each sensor operates unique physics parameters, making the fusion process complex, and often challenging to work in the same feature vector.

Recently, there has been a growing demand for multidisciplinary research on multisensor data fusion technology in the remote sensing research community. This is due to rapid technological advances and the possibility of diverse applications in various scientific fields, yielding additional information and providing more reliable results than by only one sensor. However, multisensor data fusion can be approached in a variety of manners, focusing either on raw data-level, feature-level, or application-level fusion.

In this thesis, I am facing the challenge of multisensor data fusion for airborne-based HS data and different remote sensing methods in different applications in different complex environments. I approached this subject by studying HS and LiDAR (Light Detection and Ranging) fusion (HL-Fusion), relying on feature-level fusion in urban environments using deep learning algorithms for semantic segmentation and multitemporal analysis. I also employed application-level fusion by contributing to a novel integration of airborne HS data with helicopter-borne magnetic data (HSM Integration) for geological characterization in remote areas.

1.2. Research Questions and Objectives

Since interest in multisensor data fusion has been steadily growing in remote sensing over the last years, the opportunities to improve our understanding of Earth and beyond have also increased. The primary purpose of this thesis is to answer the research questions: "*What are the most efficient feasibilities in the use of airborne HS multisensor data fusion at feature and application levels for various applications?*" and "*What are the technical and analytical benefits for multisensor data fusion, taking into consideration additional time and costs, increased complexity of the analyses, but achieving significantly improved outcomes?*"

I addressed the primary objectives by dividing them into various problematics investigating:

Feature-level fusion for HL-Fusion:
Our primary goal in the feature-level fusion was to develop a general workflow for multitemporal airborne HS and LiDAR data for land cover classification in the urban environment. The advantage of using multitemporal data from one study of interest in the land cover analysis is the possibility of stable classification through an iterative process and the generation of by-products, such as a local spectral library, as an automatic procedure. The workflow for such HL-Fusion includes feature and model selection for segmentation (Paper A), segmentation process and comparison of segmentation model parameters (Paper B), and segmentation optimization for multitemporal data with the creation of a local spectral library and local map update (Paper C).

2

Application-level fusion for HSM integration:

HS data can also be implemented as an auxiliary tool for geophysics to, for example, improve mineral exploration in remote areas. However, since HS and geophysical methods (here: magnetics) operate in different feature vectors, I focused on application-level data fusion proposing a novel approach to such HSM Integration using physical proxies (Paper D). The integration of geophysics and remote sensing is still in its initial development, and our analysis is intended to serve as a starting point for the forthcoming studies.

1.3. Contributions and Thesis Outline

The contributions of this PhD thesis have been published in peer-reviewed papers in scientific journals and presented at many conferences. The journal papers (Paper A-D) and those not included in this thesis (Paper E) are attached in Appendix.

Chapter 2 shows the datasets used during the PhD study. I also present the main ideas behind multisensor data fusion, explaining different concepts and challenges in multisensor data fusion. Furthermore, I briefly describe the theoretical background of airborne HS imaging and address the multisensor data fusion of HS images with other remote sensing methods. Firstly, I describe airborne HS data fusion with LiDAR data for land cover classification in urban environments. Secondly, I depict the integration of airborne HS data with magnetics for geological characterization in harsh environments.

Chapter 3 summarizes, one by one, all the publications, highlighting the most important contributions and achievements. In the beginning, I present an extensive literature review of urban land cover classification based on HS and LiDAR data applying machine learning (ML) and deep learning (DL) algorithms (Kuras et al., 2021). Based on this review, I compared standard 2D segmentation methods and proposed 3D segmentation algorithms on fused HS and LiDAR data for urban land cover classes (Kuras et al., 2022c). The best segmentation approach from this article (Kuras et al., 2022c) served as the basis for multitemporal analysis on fused HS and LiDAR data. I developed a process chain for a self-improving segmentation algorithm based on temporal fused HS and LiDAR data (Kuras et al., 2022a). Through this iterative method, I can identify high and low frequent changes and create a local urban spectral library that helps update maps

automatically and effectively. In the final article, I present the potential of multisensor data fusion applications using airborne HS data with helicopter-borne magnetic data in remote areas in West Greenland, where I focused on iron properties for geological characterization, developing a new Modified Iron Feature Depth index (MIFD) (Kuras et al., 2022b).

These examples of multisensor data fusion applications answer the stated research questions in (Chapter 1.2.) and the conclusion is presented in Chapter 4, pointing out the limitations and providing solutions in future perspectives based on this doctoral thesis.

1.4. Publications

The publications from this doctoral thesis answer the scientific questions stated above and contributed to developments in multisensor data fusion in remote sensing. The scientific articles successively describe:

- Extensive literature review based on land cover classification approach in the urban environment in the last decade using ML and DL on HS and LiDAR data:
  Paper A (Section 3.1.): Kuras A., Brell M., Rizzi J., Burud I., Hyperspectral and Lidar Data Applied to the Urban Land Cover Machine Learning and Neural-Network-Based Classification: A Review*, Remote Sensing, 2021, 13(17), 3393.*

- Comparison of performance evaluations for semantic segmentation in two and three dimensions of fused HS and LiDAR data in the urban environment:
  Paper B (Section 3.2.): Kuras A., Jenul A., Brell M., Burud I., Comparison of 2D and 3D semantic segmentation in urban areas using fused hyperspectral and lidar data, *Journal of Spectral Imaging*, *2022, 11(a11), pp. 1-17.*

- An approach based on fused multitemporal HS and LiDAR data in an urban environment:
  Paper C (Section 3.3.): Kuras A., Brell M., Liland K.H., Burud I., Multitemporal Feature-Level Fusion on Hyperspectral and LiDAR Data in the Urban Environment, *Submitted on December 2022 to Remote Sensing.*

- The possibilities of application-based HS data integration with magnetics in geological characterization in remote areas

  Paper D (Section 3.4.): Kuras A., Heincke B.H., Salehi S., Mielke C., Köllner N., Rogass C., Altenberger U., Burud I., Integration of Hyperspectral and Magnetic Data for Geological Characterization of the Niaqornarssuit Ultramafic Complex in West-Greenland, *Remote Sensing, 2022, 14(19), 4877.*

1.5. Conferences

- **IGARSS**, September 26 – October 2, 2020, Virtual event
  - Oral presentation: Geological characterization of Niaqornarssuit Complex based on hyperspectral and magnetic data fusion, oral presentation.

- **WHISPERS**, March 24-26, 2021, Virtual event
  - Oral presentation: Machine learning methods for road edge detection on fused airborne hyperspectral and LiDAR data
  - Conference paper (Appendix: Related papers): Senchuri R., Kuras A., Burud I., Machine learning methods for road edge detection on fused airborne hyperspectral and LiDAR data, *2021 11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2021, pp. 1-5, doi: 10.1109/WHISPERS52202.2021.9484007.

- **EARSeL**, March 30 – April 1, 2021, Virtual event
  - Oral presentation: Blue-green microstructures – detection of geometrical and permeability features of microstructures

- **EARSeL**, June 22-24, 2022, Potsdam, Germany
  - Oral presentation: change detection in urban areas from airborne-based hyperspectral and lidar data

- **IASIM**, 2022, July 3-6, 2022, Esbjerg, Denmark
  - Oral presentation: Hyperspectral-LiDAR fusion applied for multitemporal analysis of urban areas

- **WHISPERS**, September 13-16, 2022, Rome, Italy
  - Oral presentation: Land cover classification based on hyperspectral-LiDAR fusion in urban areas: a case study

  - **JURSE**, May 17-19, 2023, Crete, Greece
  - Feature-Level Based Hyperspectral and Lidar Data Fusion for Urban Analysis

CHAPTER 2

Background

Remotely sensed data can be acquired using passive or active sensors. Passive sensors, such as HS imaging spectrometers, measure the naturally available energy reflected from the surface (Meer et al., 2001). Active sensors use an artificial energy source to illuminate the ground receiving the backscattered radiation, e.g., LiDAR scanner (Gupta 2003). However, besides optimal sensing, geophysical investigations, such as magnetics, use physical signals sent to the ground to measure Earth´s magnetic field and its anomalies (Essa et al., 2022).

2.1. Datasets

In this thesis, I used airborne HS data, combining them with LiDAR in an urban environment and magnetics in harsh environments. Our datasets can be divided according to the applications used, such as urban analysis (Paper B, C) and geological characterization in remote areas (Paper D).

2.1.1. Høvik, Norway: Urban environment

In Paper B and Paper C, airborne HS and LiDAR datasets have been fused and analyzed. Both HS and LiDAR data were acquired under cloud-free conditions simultaneously by the Terratec AS Company on four flight campaigns in August 2019, April and September 2020, and June 2021 over Bærum municipality near Oslo, Norway (Figure 1).
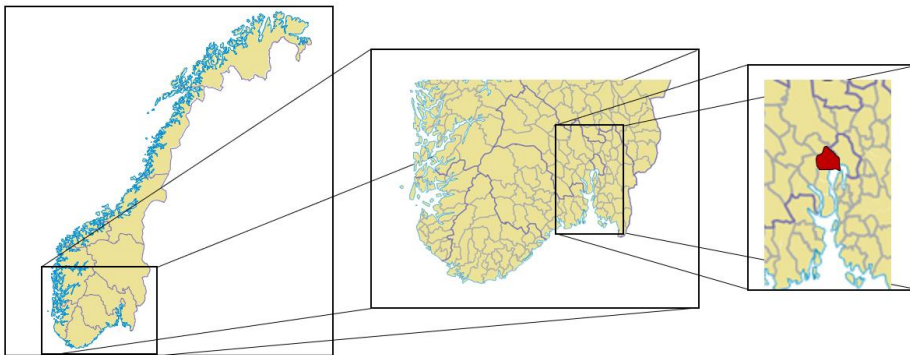
Figure 1. Map of Norway showing the location of study of interest in Bærum municipality.

In particular, the HS data included images from two HySpex sensors: VNIR−1800 (400 − 1000 nm) and SWIR−384 (1000 – 2500 nm), with 0.3 m and 0.7 m spatial resolution, respectively. The LiDAR data were acquired using Riegl VQ−1560i, with five pulses per m^2 and intensity at 1064 nm.

The particular area of interest is located in Høvik with a coordinate extent of 588060, 6641500; 588878, 6641735 WGS 84 / UTM zone 32N (Figure 2).

Figure 2. The study area in Høvik.

HS and LiDAR data have been preprocessed before the analysis.

The HS data were georeferenced and orthorectified by the Terratec AS (Field Group) using the PARGE software (Parametric Geocoding and Orthorectification for Airborne Optical Scanner Data). PARGE reconstructed the geometry for each HS image pixel using a digital elevation model (DEM) of the region, GPS position, and attitude (ReSe Applications). The geocoded radiance data were topographically corrected to reflectance, applying atmospheric correction using ATCOR-4 (Atmospheric and Topographic Correction for airborne imagery). In the analyses, I excluded absorption features associated with H_2O and OH close to bands at 1.4 μm and 1.9 μm, and noisy

bands and outliers were set to bad bands. The images from all flight lines with residual 176 channels for VNIR and 262 channels for SWIR were mosaicked separately using ENVI software (Environment for Visualizing Images). The resulting mosaic layers of VNIR and SWIR were stacked by applying the Savitzky-Golay filter (Ruffin et al., 1999) to all spectra and multiplicative scaling factor from VNIR to SWIR for leveling adjustment between ending VNIR and beginning SWIR spectral range around 1000 nm.

First, from the LiDAR data (3D point clouds), the noise and outliers were removed, after which the point cloud was converted into 2.5D rasters of 0.3 m pixel size using Quick Terrain Modeler software. In Paper B and Paper C, the rasters contained information about slope, multiple returns, the intensity from the first return, normalized Digital Surface Model (nDSM), and point density (Figure 3).
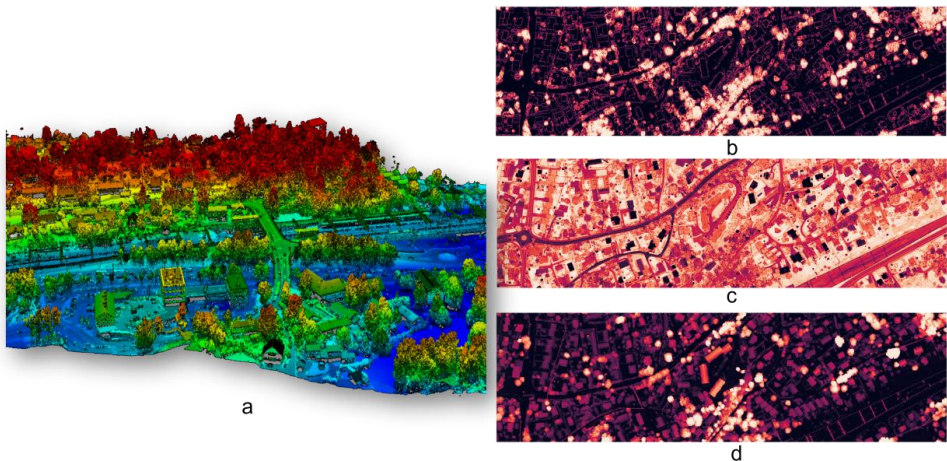
Figure 3. a) Cutout from a 3D point cloud from the area of interest. b-d) Raster examples representing multiple returns (b), the intensity from the first return (c), and nDSM (d).

2.1.2. Niaqornarssuit complex: West Greenland

The data used for geological characterization in West Greenland contain airborne HS (HyMAP (Kruse et al., 2000)) (Figure 4a) images and helicopter-borne magnetic data (Paper D (Kuras et al., 2022b)) (Figure 4d). The study area is a geological complex (Niaqornarssuit Complex) (Figure 4b, c) located in West Greenland (66.83°N −52.02°E long/lat WS 84). The Niaqornarssuit Complex is an elliptically formed ultramafic complex of 1.8 km x 0.9 km extent (Gothenborg et al., 1977). Well-exposed iron-rich (ultramafic) rocks characterize the complex at the surface, satisfying conditions to

acquire HS images. The area was entirely covered with a regional HS survey (HyMAP (Tukiainen et al., 2005)) in 2002. The HS camera was mounted on a Piper Navajo Chieftain aircraft flying at an altitude of $\sim$ 2500 m with a scanner swath width of 3000 m (Cocks et al., 1998; Kruse et al., 2000).
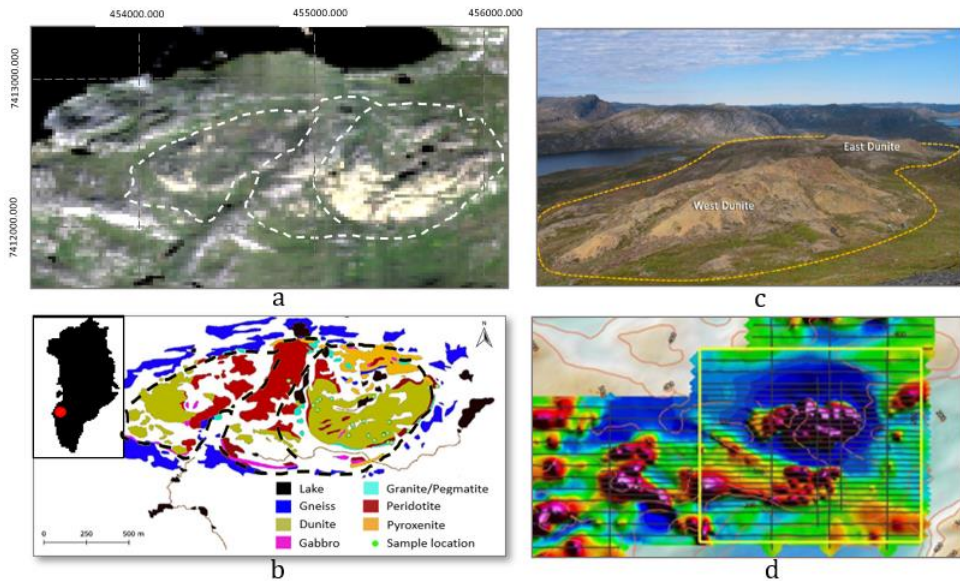
Figure 4. The Niaqornarssuit Complex a) an HS HyMap image in RGB color representation, b) a simplified geological map of the Niaqornarssuit Complex from (Østergaard 2011), c) the geological complex from the southwest direction, representing a gently undulating rolling terrain with moderate relief, d) magnetic map of the study of interest. The yellow rectangle outlines the area around the geological complex.

The geocoded radiance data were converted to reflectance using ATCOR4 with rugged terrain optimization (Salehi et al., 2017). The absorption features associated with H_2O and OH close to bands at 1400 nm and 1900 nm were excluded from the analysis. The HS images, with the remaining 106 channels, were mosaicked using ENVI. Clouds, shadows, water, snow, ice, and rugged terrain with low illumination were masked from the HS scene.

The magnetic data were acquired in 2012 in a magnetic helicopter-borne survey by Geotech Ltd. (Geotech 2012). A GPS navigation system and a radar altimeter were built-in, generating a local Digital Elevation Model (DEM) of the terrain. The helicopter flew at an average speed of 80 km/h and an altitude of 87 m. The magnetic sensors were positioned on average 63 m above the ground. The total magnetic intensity map (Figure

4d) was obtained through preprocessing, including several corrections and micro-leveling (Geotech 2012).

2.2. Airborne Hyperspectral imaging

Imaging spectroscopy (hyperspectral imaging) provides a three-dimensional cube that includes two-dimensional spatial information – *x*, *y* and adds spectral information – *λ* in the reflective spectral range from 400 – 2500 nm at each position (pixel) as a spectrum in the third dimension (Clark 1999) (Figure 5).
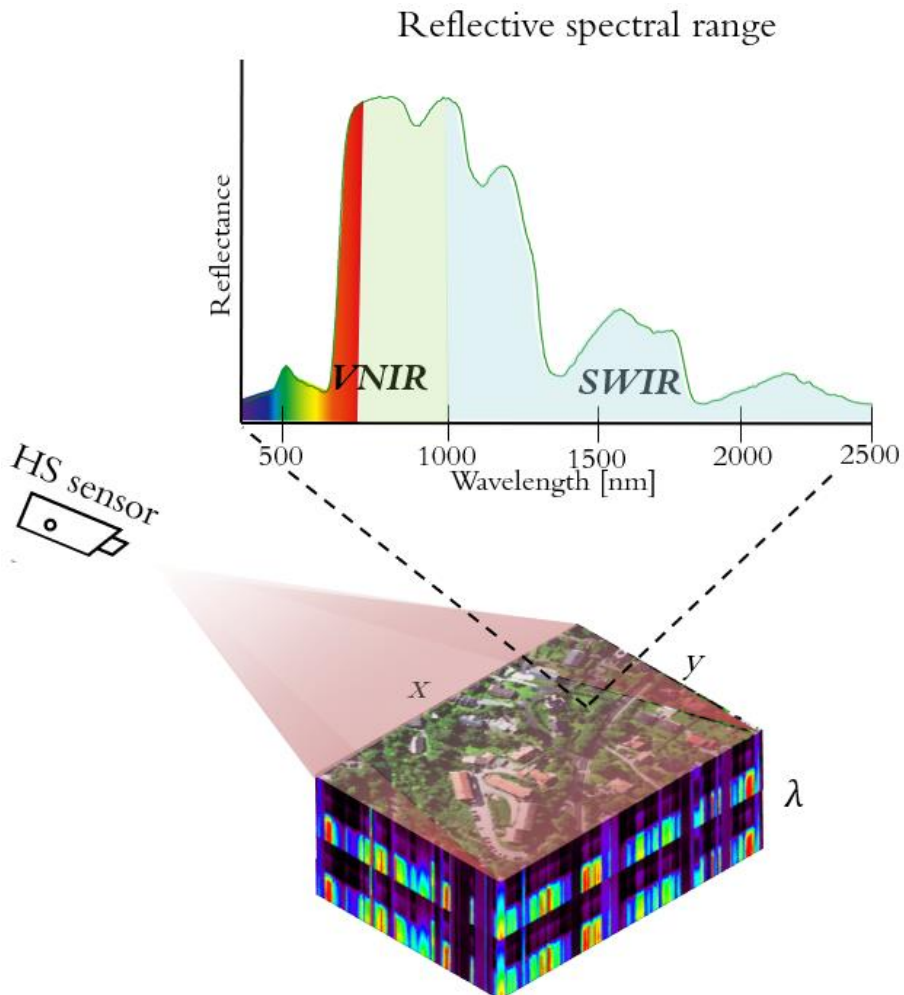
Figure 5. The HS imaging principle represents an HS datacube and a random vegetation spectrum extracted from one highlighted pixel with black dashed lines.

The benefit of HS imagery is identifying an object based on its unique absorption features – spectral fingerprints. Therefore, the optical remotely sensed data are used to classify and map land covers based on physical and chemical material properties (Gu et al., 2012). However, HS images result in high dimensional data leading to computationally expensive analyses. Other limitations are the insufficiency of reference and training spectra and spatial variation in the spectral information (Santara et al., 2016). Additionally, imaging spectroscopy-based classification is limited to several classes that contain the same material or similar spectral responses, such as roads and cycleways made from asphalt (Li et al., 2018).

The accuracy and interpretability of object classification based only on a single sensor in complex, dense urban areas are often insufficient (Ghamisi et al., 2016). The HS data are limited in object differentiation made from the same material (Debes et al., 2014). The airborne images from passive sensors contain shadows that can lead to object misinterpretation or incomplete classification results (Hui et al., 2008). By combining the two types of data, the environmental-insensitive LiDAR sensor overcomes the shadow issue and complements the analysis with detailed information about the elevation (Dalponte et al., 2008). Multisensor data fusion can provide robust and reliable land cover classification, including spectral, spatial, and topographical features (Li et al., 2018).

The advanced literature on methods for information extraction from HS data includes denoising, spectral unmixing, dimensionality reduction, and data classification (Paoletti et al., 2019). Depending on the classification aim, different advanced mapping methods are applied to achieve the goal. HS imaging data are usually vectors expanding to higher dimensions (Manolakis et al., 2016). Generally, the first step of the classification process of the HS data is a spectral dimensionality reduction to the relevant components by applying linear spectral transformations (Manolakis et al., 2016). Standard techniques for dimensionality reduction are principal component analysis (PCA), linear discriminant analysis (LDA), and other unsupervised classification methods.

HS imagery analysis is often challenging due to non-linear relations between spectral characteristics and similar materials (Li et al., 2019). Furthermore, the spectrum of any pixel in the HS image is a mixture of diverse materials. Therefore, ML and DL provide

helpful tools for confronting non-linear issues (Li et al., 2019). Other challenging problems in analyzing HS data are related to the data quality, spatial dependency of the spectra, and little reference spectral library for supervised classification methods (Camps-Valls et al., 2005).

2.3. Multisensor Data Fusion

In general, multisensor data fusion integrates observations from different sensors to improve and provide a stable and robust analysis of a target of interest (Durrant-Whyte et al., 2008). In remote sensing, multisensor data fusion has many advantages through improvement in classifications, detections, reliability, and reduction in data ambiguity, with the expansion of spatial, temporal, feature, and/or parameter dimensions (Khaleghi et al., 2013; Hall et al., 1997). However, to conduct data fusion, it is required to account for several challenges.

Data errors

Data produced by multiple sensors always contain technical errors, certain inaccuracies, and measurement uncertainties. Fusion-based data analysis is expected to be effectively insensitive to such imperfections and take advantage of data redundancy to mitigate their implications (Khaleghi et al., 2013). Before the fusion application, it is required to preprocess data from all sensors separately to eliminate errors and prevent error propagation in multisensor data fusion.

Data alignment

Fusion from multiple sensors involves preprocessing that must be done before the fusion operation. In the situation that the functionality of the integrated sensors differs, it is necessary to perform geometric coregistration to a common frame that involves appropriate calibration and sensor coalignment (Brell et al., 2016).

Acquisition time

Sensor fusion refers to the integration of data acquired simultaneously, e.g., a flight campaign with an HS camera and LiDAR scanner (Kuras et al., 2022c), and the integration of data collected at different times of day and year (Kuras et al., 2022b). In the latter case, a fusion-based analysis must deal with time variations without any loss of information.

Data dimensionality

Each additional sensor added to the analysis means introducing a new parameter and dimension, which increases the complexity of the analysis as well as its processing time.

Multisensor data fusion can be performed based on raw data, product, and application level (Figure 6), applying either physical or empirical approaches (Brell et al., 2016; Torabzadeh et al., 2014). Physical approaches consider sensor parameters, principles of the measurements, illumination sources, sensor position, and other quantities in the data preprocessing step (Asner et al., 2012; Brell et al., 2016). In HL-Fusion, the intensity values of the overlapping wavelengths in the two sensors describe a physical link that enables the creation of a hyperspectral point cloud (Brell et al., 2019). Empirical approaches neglect sensor characteristics; however easily correct geometric errors and adjust the spatial resolution of one sensor to another empirically (Kuras et al., 2021).
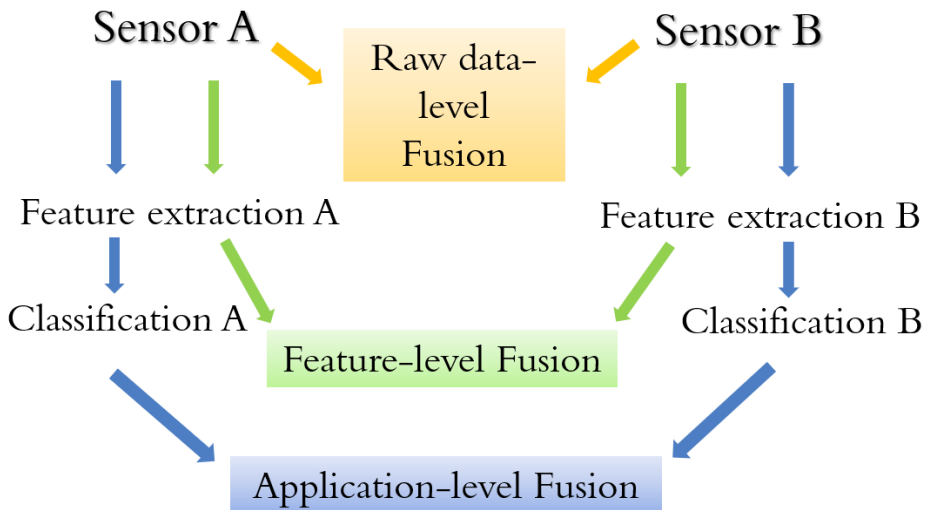
Figure 6. Different multisensor data fusion levels, including raw data-level fusion (orange signs), feature-level fusion (green signs), and application-level fusion (blue signs).

Raw data-level fusion

Sensor fusion based on raw data (without feature extraction) refers to combining data at the lowest processing level to merge physical parameters and thus expand the dimensionality of the final fused vector. The fused vector maintaining original information should improve classification results and be characterized by reliability and

transferability, which are neglected when using only one sensor. Such fusion is known in the remote sensing field, e.g., HS data combined with high spatial resolution panchromatic imagery (Kaufman et al., 2015), different HS sensors covering different spectral ranges such as VNIR and SWIR are merged (Senchuri et al., 2021), HS fused with Synthetic Aperture Radar (SAR) (Dabbiru et al., 2015). However, dimension expansion of already high-dimensional data can lead to a curse of dimensionality issues, mainly when the acquired data are limited (Hughes 1968).

Feature-level data fusion

Feature-based fusion is one form of fusion where the data from each sensor are subjected to preprocessing and feature extraction separately, building a new normalized feature space with compressed dimensions for each data. Subsequently, extracted features are merged before the main classification task, such as in the fusion of SAR and HS images (Hsu et al., 2003) and HS and LiDAR data in the urban environment (Senchuri et al., 2021; Man et al., 2015; Kuras et al., 2022c; Liao et al., 2017; Luo et al., 2017; Hasani et al., 2017; Khodadadzadeh et al., 2015) (Paper C). In the classification problem, such analysis empowered an upgrade from pixel-based to object-based classification by extracting spectral features from HS data and raster-based LiDAR features, such as height and its derivatives, intensity values (from the first return), point density, among others (Kuras et al., 2021).

Application-level data fusion

In the case of application-based (decision-based) fusion, the data from each sensor are analyzed separately, after which the final results are merged into a common decision or compared, e.g., in the form of a map or an image (Man et al., 2015; Luo et al., 2017), such as in HS data fused with geophysics (Kuras et al., 2022b; Bedini et al., 2018; Jackisch et al., 2020). This fusion is employed when combining data from individual sensors into a common feature vector is difficult (Sharma et al., 1998).

2.4. Hyperspectral-LiDAR Data Fusion for Urban Land Cover Segmentation

HL-Fusion is based on integrating the spectral-spatial information provided by the HS data and the spectral-spatial-geometric-structural information of the LiDAR data on the feature level. Although differences in physics characterize HS sensors and LiDAR scanners, their properties can be combined.

The LiDAR scanner´s monochromatic laser intersects into one of the following wavelength regions: 532 nm (VIS), 1064, or 1550 nm (SWIR), as the HS data cover the reflective spectral range (e.g., from 400 to 2500 nm). The combination of HS sensors and LiDAR significantly influences the advancement of remote sensing applications, paving the way for full 3D analysis (Hakala et al., 2012). However, HL-Fusion does not just stand for classification in full 3D geometry. Most applications are based on a geometric simplification of high-dimensional data, reducing both HS and LiDAR data to 2.5-dimensional rasters. In HL-Fusion, the geometrical errors are usually mitigated empirically by adjusting the spatial resolution of one sensor to the other, neglecting sensor characteristics, such as scan or incident angles (Kuras et al., 2021). Despite the dimensional compression, HL-Fusion can perform better in land cover classification than one single sensor when integrating spectral, spatial, and structural features into the analysis. Attempts at such fusion have been made in the past decade, e.g., merging intensity, height and its derivatives, and other feature information extracted from LiDAR-based point clouds to spectral features derived from HS data. Individual studies on HL-Fusion rely on matching spectral data to the first return from LiDAR data. This helps preserve the 3D close-to-reality geometry and structures and strengthens the efficiency and robustness of the analysis of fused data (Brell et al., 2019).

2.4.1. Airborne LiDAR Scanning

The LiDAR scanner as an active sensor is a whiskbroom-type instrument emitting radiation from one bandwidth or more for multispectral LiDAR systems. A LiDAR system calculates the distance d based on the time the emitted light travels to the surface and back to the receiver based on the equation:

$$d = \frac{ct}{2} \qquad\qquad\qquad (1)$$

where *c* is the speed of light.

A laser in a LiDAR system is monochromatic, often available in three wavelengths, such as visible – 532 nm for bathymetric analyses, near-infrared – 1064 nm, and 1550 nm for vegetation and urban analyses. The laser hitting the object's surface measures the time passed between the pulse emission and reflection, enabling the extraction of information in a 3D point cloud (*x,y,z*) (Figure 7). From such a point cloud, we can derive elevation, single and multiple returns, the intensity from backscattered signals, texture information, and more (Yan et al., 2015; Wehr et al., 1999). Since the last decade, airborne LiDAR has become a standard method for diverse applications in urban and rural environments for analysis of terrain and coastal areas, 3D objects modeling, forestry, high-resolution topographic mapping, archeology, city planning, disaster management, and urban flooding simulation (Hui et al., 2008; Tóvári et al., 2012; Teo et al., 2017; Zhou et al., 2018). Also, the benefit o fusing airborne LiDAR is that it is resistant to complex variations in terrain and illumination conditions, preserving the full 3D geometry of data (Yan et al., 2015).
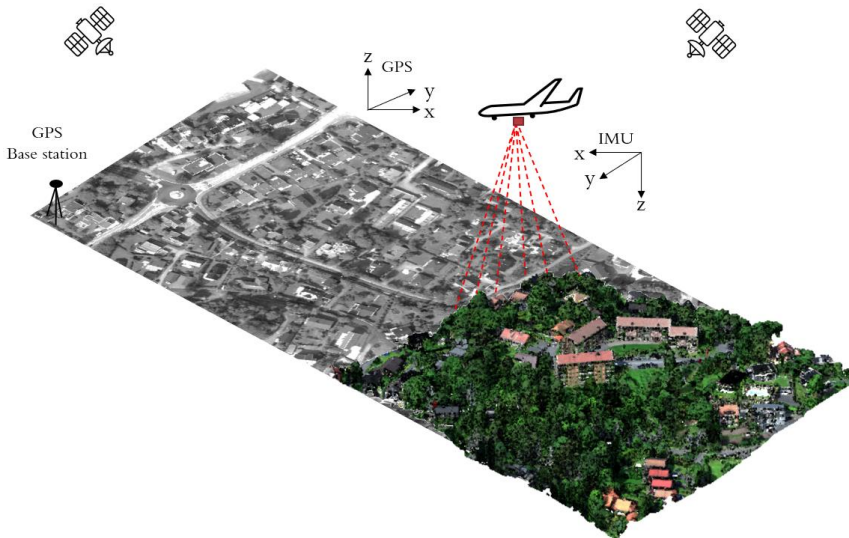
Figure 7. Airborne LiDAR scanning principles.

2.5. Hyperspectral-Magnetic Integration for Geological Characterization

In harsh environments, such as Greenland, standard mineral exploration techniques are problematic due to limited access and remoteness. Airborne-based remote sensing methods such as HS imaging and magnetics provide a time- and cost-efficient tool for geological characterization in an early exploration phase, especially for iron-rich geological complexes. Furthermore, both HS data and magnetics cover large areas by measuring physical parameters describing the spatial distribution of a structure of a geological complex and its lithology.

Iron is the most significant component that allows the integration of HS and magnetic data in an iron-rich complex. In HS imaging, iron has distinct absorption features at 650 nm (ferric iron Fe^{3+}) and 1000 nm (ferrous iron Fe^{2+}) in alteration minerals. Only iron and a few iron minerals are responsible for the magnetic properties of rock, which are caused by built certain iron ratio of Fe^{3+}/Fe^{2+}. Therefore, magnetic measurements are suited to retrieve magnetic information in iron-rich geological complexes. However, rock magnetization is related only to magnetite, iron-titanium oxide minerals, and pyrrhotite (Dentith et al., 2014), and lithological characterization is impossible if the geological complex contains other iron alteration minerals. On the other hand, HS can distinguish iron alteration minerals, but magnetite (considered the most magnetic mineral), despite a high iron content, does not show distinct absorption features in the reflective spectral range of the electromagnetic spectrum (Hunt et al., 1995; Till et al., 2018; Kuras et al., 2022b).

Although HSM Integration seems promising for mineral exploration in iron-rich geological complexes, especially in remote areas, little research has been done to combine optical remote sensing and magnetics, e.g., drone-based magnetics with multispectral images (Jackisch et al., 2022), drone-based magnetics, RGB, HS, and multispectral data (Jackisch et al., 2020), and drone-based magnetics and HS images (Jackisch et al., 2019).

2.5.1. Magnetics

As a geophysical technique, the magnetic investigation method is the oldest and most commonly used in Earth´s subsurface research (Hinze et al., 1990). The magnetic method measures Earth's magnetic field variations, including anomalies. This technique is applicable for mineral exploration since magnetics can retrieve depth information compared to other mapping tools (Clark 1997). Therefore, in survey areas with poor exposure, the magnetic method is an integral approach for subsurface mineral exploration (Dentith et al., 2014). Magnetic field measurement is a required method for investigating magnetic anomalies. Therefore, modern instruments – magnetometers measure the Total Magnetic Intensity (TMI), a vector quantity of the vertical and the two horizontal components of the magnetic field (Dentith et al., 2014).

Rock magnetism is determined by mineral content with magnetic properties. Magnetic properties of rocks reflect ferrous (Fe^{2+}) and ferric (Fe^{3+}) iron content in the rocks and depend on the chemical composition of iron-bearing minerals oxidation ratio of the iron (Fe^{3+}/Fe^{2+}) (Clark 1997). Generally, magnetic minerals are accessory minerals in intrusive rocks. Therefore, a magnetic property cannot directly relate to one lithology (rock type) (Dentith et al., 2014).

CHAPTER 3

Publications

3.1. Feature-Level Data Fusion

3.1.1. Hyperspectral-LiDAR data meet machine and deep learning in the urban environment

Before analysis began on HS and LiDAR data in the urban environment, we focused on related works and published our conclusions in the literature review in Paper A (Kuras et al., 2021). This scientific review presents ML and DL urban mapping methods focusing on airborne HS and LiDAR data and describes key characteristics of HS and LiDAR data and common urban land cover classes such as roads, buildings, and vegetation (Figure 8).
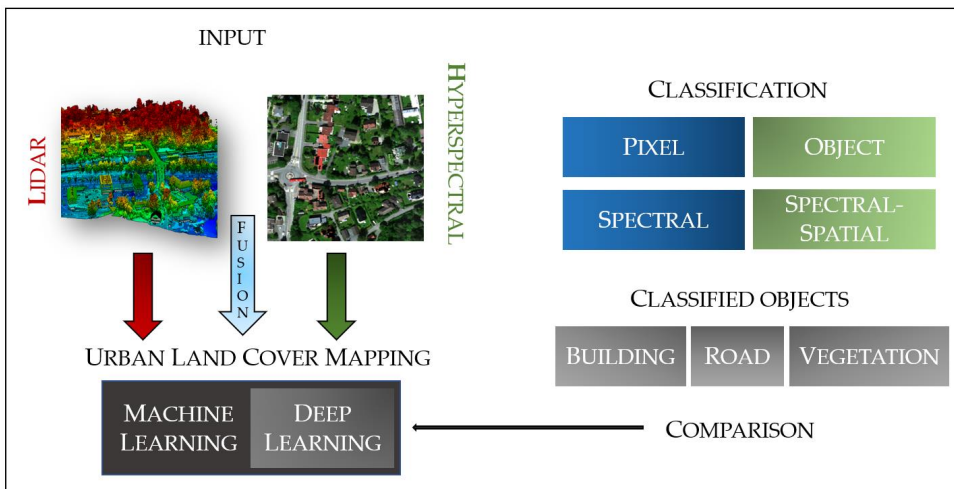
Figure 8. An illustrated overview of the scope of Paper A.

We provided the latest information on advances in mapping techniques based on HS and LiDAR data in urban environments based on the reflective spectral range (400 – 2500 nm). This work was intended to sum up classification methods in the urban environments for data scientists and remote sensing experts in the last decade, including Support Vector Machines (SVM), Random Forest (RF), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). Moreover, we pointed out the

advantages and challenges of applying each algorithm to HS and LiDAR data and categorized the comparison based on input data, domain (spectral, spatial), and extracted features expressed in an extensive summarizing table.

In the last decades, ML and DL have been revolutionizing the world of data processing of remote sensing data, such as HS and LiDAR data. Also, with the advancement of sensory technology, detailed information about the target of interest can be accessed by combining different methods. While shallow ML algorithms tend to operate pixel-based, DL algorithms function object-based, allowing simultaneous spectral features and contextual information extraction. This improves the accuracy of the results, and such classification reflects reality more closely.

Particular attention should be paid to DL implementations of multisensor data fusion, such as the HL-Fusion. An HL-Fusion can be the key to complex urban analysis, going beyond human knowledge to explore and preserve the urban environment. The airborne HL-Fusion has been studied for urban land cover classification combining spectral features from HS data and height features and their derivatives from LiDAR data (Brell et al., 2019; Dian et al., 2016; Debes et al., 2014; Morchhale et al., 2016). In addition, an increasing amount of analysis on HL-Fusion data includes the spatial context in HS and/or LiDAR part, especially when implementing DL algorithms (Ghamisi et al., 2016; Chen et al., 2016; Li et al., 2018; Hang et al., 2020; Feng et al., 2019; Zhang et al., 2020; Chen et al., 2017; Singh et al., 2022; Senchuri et al., 2021; Zhao et al., 2020).

This literature review was the basis for analyses of HL-Fusion data. With this work, not only did we select the appropriate algorithms for the analyses, but the key features of HS and LiDAR data were effectively extracted to align with the classification of selected urban land cover classes in **Paper B** (Chapter 3.1.2.) and **Paper C** (Chapter 3.1.3.).

3.1.2. Semantic segmentation in an urban environment based on airborne Hyperspectral-LiDAR fusion

In the present and the next section, we focus on the fusion of airborne-based HL-Fusion at the feature extraction level for urban analysis. The foundation for these two studies is an extensive literature review presented in **Paper A** (Chapter 3.1.). The purpose of this section is first to design an HL-Fusion at the feature extraction level, proposing for the first time a 3D convolutional operation based U-Net (Ronneberger et al., 2015) and ResU-Net (Zhang et al., 2018; Yang et al., 2019) (Figure 9) for HL-Fusion data in the urban environment.
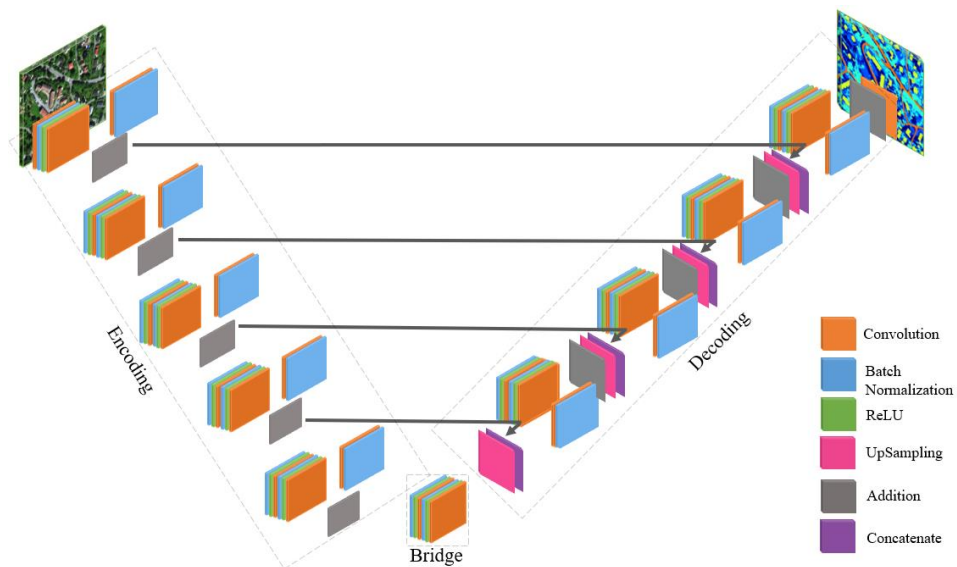
Figure 9. The ResU-Net architecture implemented as an extension of U-Net.

For this comparative study, the dataset from 2021 described in Chapter 2.1.1. was used. To facilitate the supervised segmentation in the complex urban environment, five main classes representing common urban land cover surfaces were selected, such as low and high vegetation, buildings, roads, and railways.

We chose to compare different segmentation models based on U-Net and ResU-Net architectures with 2D and 3D convolutional operations. All algorithms were tested with different loss functions, such as two statistical-based Categorical Cross-Entropy, Focal loss (Lin et al., 2018), and one geometric-based Jaccard loss (Jaccard 1912; Duque-Arias

et al., 2021). However, to get the best out of Focal and Jaccard losses, we combined them in an additive way (Focal-Jaccard loss) (Figure 10).
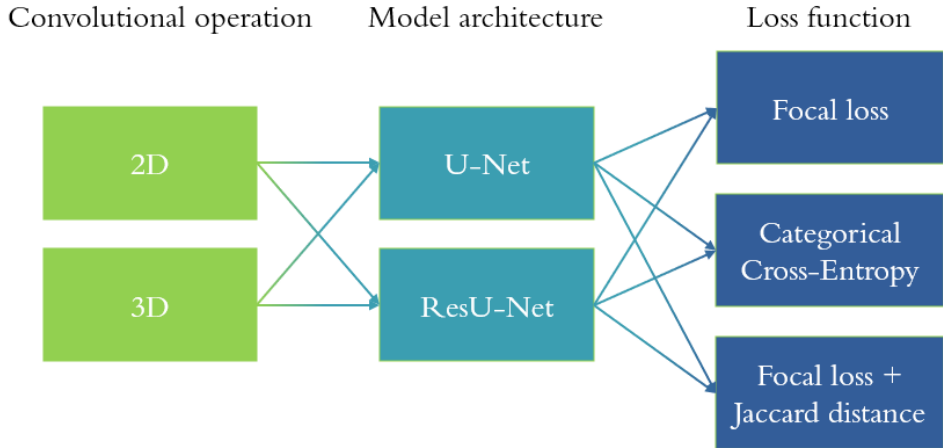
Figure 10. All semantic segmentation combinations implemented in this study.

Table 1. Segmentation accuracies obtained by Matthew's correlation coefficient (MCC). The gray shaded boxes indicate the highest accuracy score comparing all models for the five classes. Fields shaded in green represent the best accuracy result obtained by comparing loss functions within one model.

	2D U-Net			2D ResU-Net			3D U-Net			3D ResU-Net		
	Cross-Entropy	Focal	Focal-Jaccard	Cross-Entropy	Focal	Focal-Jaccard	Cross-Entropy	Focal	Focal-Jaccard	Cross-Entropy	Focal	Focal-Jaccard
low vegetation	0.82	0.59	0.64	0.71	0.63	0.62	0.61	0.56	0.63	0.54	**0.94**	0.61
high vegetation	0.95	0.89	0.91	0.59	0.94	**0.98**	0.88	0.82	0.91	0.82	0.86	**0.98**
building	0.94	**0.99**	**0.99**	0.89	0.97	**0.99**	0.97	**0.99**	0.98	0.16	0.95	**0.99**
road	0.90	0.97	0.96	0.83	0.95	0.94	0.96	0.95	**0.98**	0.21	0.92	0.97
railway	0.95	**0.99**	**0.99**	0.05	0.88	0.95	**0.99**	0.93	**0.99**	0.10	0.98	**0.99**
Overall accuracy (MCC)	**0.84**	0.80	0.82	0.62	0.80	**0.81**	0.78	0.73	**0.80**	0.31	**0.88**	0.84

Figure 11. a) HS image in RGB color representation, b) Ground-truth with five classes obtained semi-automatically, c) and d) best segmentation results such as 2D U-Net with Categorical Cross-Entropy, and 3D ResU-Net with Focal loss, respectively.

From Table 1, which demonstrates the experimental results, the following conclusions can be stated:

1. In the 3D segmentation, Focal (Figure 11d) and Focal-Jaccard loss functions yielded significantly better overall accuracy based on Matthew´s correlation coefficient (MCC) than the Categorical Cross-Entropy models. This is because Focal and Jaccard losses converge to a lower loss faster than Categorical Cross-Entropy, which helps classify classes that are difficult to predict (Kuras et al., 2022c). Moreover, we took advantage of the Focal-Jaccard loss function as Focal loss can handle imbalanced data, and Jaccard distance is known for recognizing objects.

2. Despite the simplicity, the 2D U-Net model with Categorical Cross-Entropy (Figure 11c) is stable and not sensitive regarding loss functions. In order to test the sufficiency of 2D segmentation against 3D segmentation, further research is carried out on more advanced segmentation tasks, intraclass segmentation, and the identification of materials in the urban environment.

3.1.3. Multitemporal analysis on fused airborne Hyperspectral and LiDAR data

For feature-level HL-Fusion, we have extended the scope of analysis by approaching more complex problems relying on the outcome from **Paper B**. In this work, we studied the same dataset as in **Paper B** but supplemented it by adding measurements from the same study area from August 2019. Inspired by the capabilities of multitemporal HL-Fusion and DL, we designed a generic workflow for multitemporal HL-Fusion analysis at the feature level (Figure 12), focusing on the following contributions:

   1. Performance of segmentation based on unsupervised endmember extraction and abundance map retrieval.

   2. Optimizing the segmentation results without parameter regularizations iterating the segmentation using multitemporal datasets.

   3. Generation of by-products, such as a local spectral library, to expand the local database.

   4. Automatic update of local maps based on available original maps of the region of interest.

The analysis began with endmember extraction, which for HS data referred to the extraction of the most representative spectral features in an unsupervised way (Kuras et al., 2022c). From the LiDAR point cloud, raster-based features were generated, such as point density, the intensity from the first return, slope, multiple returns, and nDSM. Extracted endmembers from HS and LiDAR were the basis for retrieving normalized abundance maps fed into 2D ResU-Net (Yang et al., 2019; Zhang et al., 2018) segmentation algorithm. Segmentation was carried out on data from 2019 and 2021, comparing the implementation of data augmentation and analysis without it. After the initial segmentation, endmembers features were updated and introduced into the second iteration of segmentation. After the second iteration, a spectral library based on HS data was automatically generated for each defined class considering intraclass variability. This work was completed by calculating change maps for each class for the 2019 and 2021 datasets.
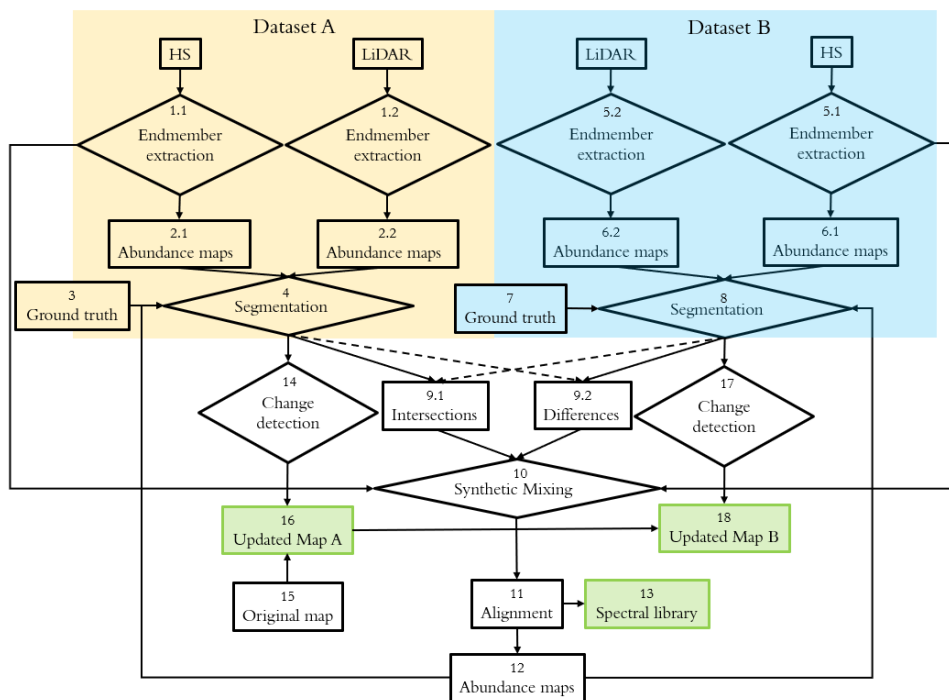
Figure 12. Generic workflow for multitemporal HL-Fusion analysis.

From Table 2, it can be inferred that optimization of endmember features before the second iteration of segmentation improved the final results for each class regardless of the implementation of data augmentation.

Table 2. Initial (I) and optimized (II) segmentation results based on MCC for each class and F1 score overall accuracies for the dataset from 2019 without and with data augmentation.

Color	Dataset	2019 without augmentation		2019 with augmentation	
	Segmentation	I	II	I	II
	Low vegetation	0.79	**0.81**	0.73	**0.75**
	High vegetation	0.92	0.92	0.94	**0.97**
	Building	0.88	**0.94**	0.97	**0.99**
	Road	0.78	**0.89**	0.92	**0.95**
	Railway	0.85	**1**	1	1
	F1	0.818	**0.831**	0.814	**0.843**

It is worth noting that the model with integrated data augmentation outperformed (Figure 13). The data augmentation not only diversified the training dataset, which is often limited but also allowed the model to learn context features, thanks to which the model easily located objects not marked on the original reference data.

a) HS scene from dataset from 2019

b) FKB Ground truth data

c) I segmentation with augmentation

d) II segmentation with augmentation
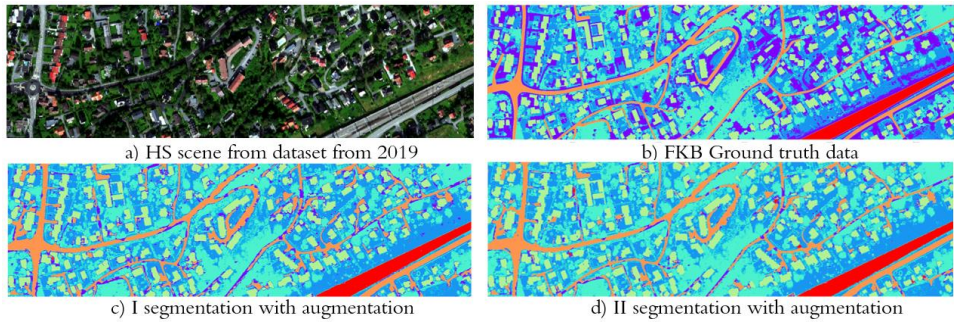
Figure 13. Segmentation maps for the dataset from 2019 presenting a) the HS scene in RGB color representation, b) the ground truth data, c) initial (I) and d) optimized (II) segmentation with data augmentation.

The best results from the optimized segmentation were used to retrieve change detection maps (Figure 14) from which it is possible to read changes associated with the addition or removal of a given object, i.e., an effective local update of maps in the urban environment.

a) Change detection of buildings

b) HS scene from dataset from 2019
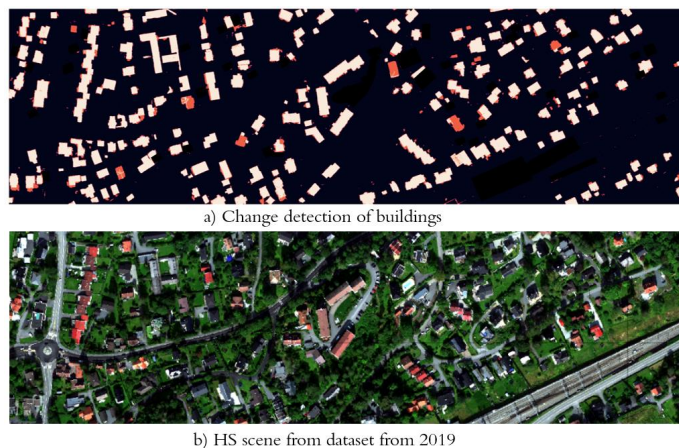
Figure 14. An example of change detection of buildings highlighting old (white) and new (orange) buildings.

3.2. Application-Level Data Fusion

3.2.1. Airborne Hyperspectral-Magnetic Fusion for geological characterization in harsh environments

To demonstrate the versatility of multisensor data fusion, we present another application of fusion of airborne HS data with another remote sensing sensor for material identification in harsh environments. This study aimed to evaluate a combination of airborne HS imaging and aeromagnetics (HSM Integration) with geochemistry support for the geological characterization of an ultramafic complex in harsh environments. Both methods complement each other since optical remote sensing delivers spectral information about the surface, while magnetics penetrates depths (Clark 1997). In this study, it was required to consider only magnetic responses related to the near surface to compare them with the optical data.

An HSM integration has great potential in characterizing minerals and rocks based on spectral analysis and their magnetic properties in three dimensions. The dataset includes airborne optical and magnetic data from the ultramafic complex described in Chapter 2.1.2.

This study developed a new Modified Absorption Feature Depth index (MIFD) for spectral analysis. The MIFD is based on the Iron Feature Depth Index (IFD) developed by Mielke (Mielke et al., 2014). The IFD is based on the subtraction of the absorption feature center (ferrous iron – 1000 nm or ferric iron – 650 nm) from the interpolation line between the right and the left shoulder of this feature (Mielke et al., 2014). The MIFD interpolated continuum (Equation 2) is divided by the absorption feature center, calculating the quotient between the interpolated line $line_{int}$ and absorption center r_{center} (Equation 3) (Figure 15) (Kuras et al., 2022b).

$$line_{int} = r_{left} + (r_{right} - r_{left})\frac{\lambda_{center}-\lambda_{left}}{\lambda_{right}-\lambda_{left}} \qquad (2)$$

$$MIFD = \frac{line_{int}}{r_{center}} \qquad (3)$$

The MIFD was calculated for two absorption bands, close to 650 and 1000 nm, associated with Fe^{3+} ($MIFD_{650}$) and Fe^{2+} ($MIFD_{1000}$), respectively (Gupta 2003).
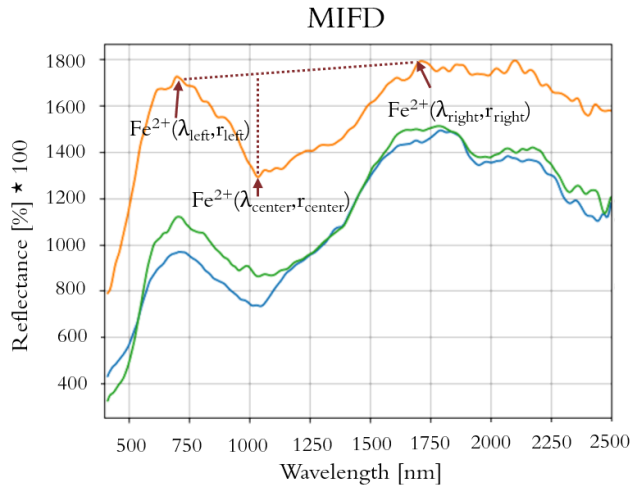
Figure 15. Schematic illustration of the $MIFD_{1000}$ on a spectrum example.

The magnetic analysis assumed that the ferrous and ferric MIFDs' ratio is directly linked to magnetic properties in the ultramafic complex. This is because only a certain Fe^{3+}/Fe^{2+} ratio in the mineral can retain magnetic properties. Therefore, we considered the ratio of $MIFD_{650}/MIFD_{1000}$ as "susceptibility" values, filled modeling cells along the surface with these values, and performed forward modeling to determine the magnetic responses for measurement locations of an aeromagnetic survey that was acquired across the complex.

The pseudo-magnetic responses from the iron ratio $MIFD_{650}/MIFD_{1000}$ were plotted against the magnetic information (residual magnetic anomaly) and the lithological information of the complex in Figure 16.

Our results provide a link between optical data and aeromagnetics. This study is an initial step in the geological analysis of complexes retaining magnetic properties based on multidisciplinary research. We focused on non-invasive mapping and integrating airborne-based HS imaging with complementary sensors. Our study presents the following key results:

   1. The integration of remote sensing methods is related to the chemical and physical properties of the investigated rocks, considering only the near-surface material.

2. At the airborne scale, the absorption feature depths of chemical elements, such as iron, are strongly correlated with the chemical amount of the element.

3. The regional lithology of the study area can explain the magnetic properties of the Niaqornarssuit Complex.

4. The main benefit of multisensor integration of airborne HS and magnetic data is a significantly higher resolution of the hybrid model and, thus, a better exploration possibility.
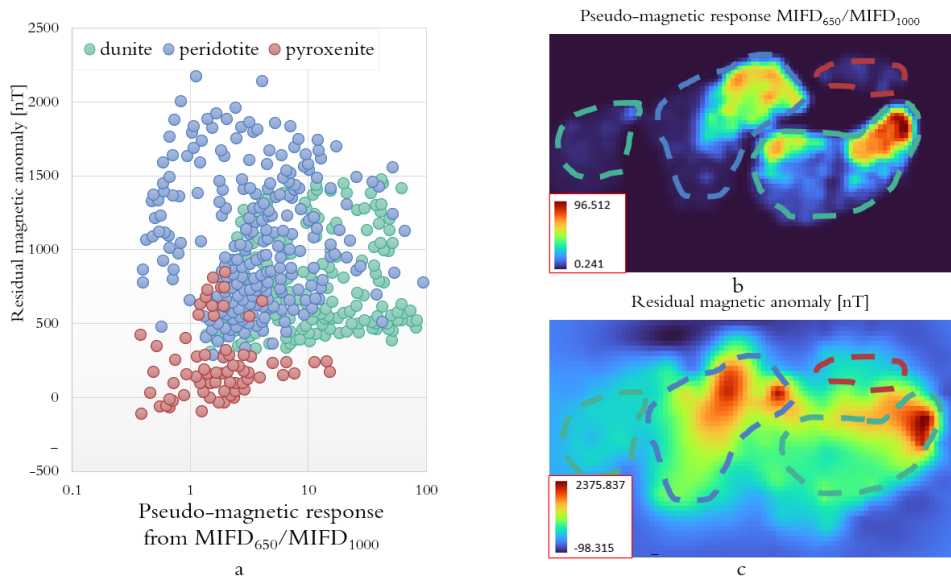
Figure 16. (**A**) Correlation plot between the magnetic data and the pseudo-magnetic responses calculated by forward modeling of MIFD650/MIFD1000-ratios with corresponding lithologies such as dunite (green), peridotite (blue), and pyroxenite (red). (**B**, **C**) show the calculated pseudo-magnetic response of the $MIFD_{650}/MIFD_{1000}$ ratio and the magnetic data, respectively.

CHAPTER 4

Conclusions

This dissertation is the product of a three-year PhD research program conducted at the Norwegian University of Life Sciences (NMBU) in Ås in Norway and was part of a project with Bærum municipality on ML and DL applications using airborne HS and LiDAR data to update local urban maps automatically. This doctoral thesis demonstrated scientific contributions to the potential of airborne HS data for multisensor data fusion for different applications. Detailed scientific contributions have been provided in feature-level-based HL-Fusion for land cover classification in the urban environment and application-level-based HSM Integration for geological characterization in remote areas.

Feature-level data fusion

The research on feature-level data fusion started with a publication of an extensive literature review on HL-Fusion for land cover analysis in the urban environment based on ML and DL classifications (**Paper A**, Chapter 3.1.1.). The primary purpose of this review was to deepen knowledge of the HS and LiDAR classification for urban analyses and report the most common ML and DL algorithms. We highlighted, in particular, that most HL-Fusion approaches are made on feature-level fusion using DL algorithms. Therefore we described in detail the key characteristics of HS and LiDAR data, comparing previous scientific work based on a classifier, input data (HS, LiDAR, HL-Fusion), domain (spectral, spatial), classes included in the analysis (building, vegetation, road), extracted features for classification, advantages, and limitations of each study. The most valuable outcome of this review was the obtained knowledge, and the work was the basis for our research on feature-level-based HL-Fusion in the urban environment.

Relying on Paper A, work on **Paper B** (Chapter 3.1.2.) began by selecting appropriate features extracted separately from HS and LiDAR data. HS data were subjected to unsupervised iterative endmember extraction, representing endmembers derived from pure pixels in the HS scene. LiDAR data were normalized by extracting five features: intensity from the first return, point density, multiple returns, slope, and nDSM. The new feature space vector thus created was used for segmentation based on 2D and 3D

convolutional operations testing various advanced loss functions, such as standard Categorical Cross-Entropy, Focal loss, and a combination of Focal-Jaccard loss. For the first time, we proposed a 3D residual U-Net (ResU-Net) for HL-Fusion data in the urban environment. Thanks to this comparative study, we could deduce that the final result of segmenting such high dimensional HS data with LiDAR features depends on the algorithm and parameters chosen for classification. The best models for our classification purpose proved to be 2D U-Net with standard Categorical Cross-Entropy loss function and 3D ResU-Net with Focal loss. Both models correctly classified the study area and maintained the correct geometry (without rounded building edges). Despite its simplicity, 2D U-Net with Categorical Cross-Entropy demonstrated that it is stable and non-sensitive to loss functions. On the other hand, 3D ResU-Net with Focal loss outperformed, and we are convinced that with more complex study areas and classification tasks, 3D models, despite requiring time and adequate computational resources, will become state-of-art in the near future.

Referring to the results of Paper B, **Paper C** (Chapter 3.1.3.) presents an advanced multitemporal urban analysis for fused HS and LiDAR data. In this study, we similarly extracted endmembers for supervised segmentation tasks as in Paper B. However, we focused on improving the available model through data augmentation and endmember feature optimization iterating the segmentation process. For this analysis, we applied bitemporal HL-Fusion data from 2019 and 2021. For optimization, we used features from both datasets, creating a stable and improved final segmentation and an automatically generated spectral library, which can be used as a by-product to expand a local database. Thanks to multitemporal analysis on HL-Fusion data, we generated a change detection map, which is the desired local map update as a final result.

Application-level data fusion

To demonstrate the versatility of airborne HS data applications, **Paper D** (Chapter 3.2.1.) presents the application-based integration of airborne HS data and helicopter-borne magnetic data for geological characterization in harsh environments. The main goal of this research study was to find a suitable physical link to integrate HS and magnetic data with the geological information of an iron-rich geological complex. This physical link in HSM Integration is iron, which not only has distinct absorption features in the reflective spectral range in HS data, but iron is the only material able to retain

magnetic properties. The work on the HSM Integration started by analyzing HS and magnetic data separately. For HS data, an innovative modified iron feature depth index (MIFD) was invented to effectively identify the ferrous and the ferric iron, considering their relative content. The result of the HS analysis was used for forward modeling with magnetic data as a starting model. The final phase of this study was based on comparing the total magnetic intensity map, pseudo-magnetic response with iron ratio, and geological information about the complex. With this study, we proved that multisensor data fusion usually involves improving the resolution of such a hybrid model, effectively enhancing the analysis results and diversifying them by adding missing parameters that would not be possible to acquire with a single sensor only.

4.1. Overall Conclusions

This dissertation focused on exploring the most effective feasibilities of using airborne HS data in fusion with other remote sensors. Based on the four articles in this PhD thesis, it can be concluded that each issue requires an individual approach, and there are undoubtedly many pathways to solving a given problem. Using my scientific background and creativity for my PhD thesis, I proposed methods in the analysis of multisensor data in the feature-level and application-level ranges, which solved the issues and confirmed the hypothesis that multisensor data fusion is profitable in terms of improved outcomes but also their variety, which can prove to be a gain in time even when the complexity of the analysis is considerably higher. HS and LiDAR fusion at the feature level is an effective way to integrate these two sensors in the analysis of the urban environment. It is worth noting that the potential and possibility of combining features extracted for analysis from both sensors, i.e., working in a common feature space, significantly facilitates the discrimination of urban land cover classes at different complexity scales. However, the fusion of HS data in geological characterization in harsh environments at the application level enables the exploitation of the capabilities of each sensor individually, generating single-sensor results. However, the fusion of those results allows for deepening the knowledge and comprehension of physical phenomena and problems that a single sensor cannot cover.

4.2. Limitations of the Multisensor Data Fusion

Computation cost

Analysis in an urban environment is very complex due to high material heterogeneity, material variance, and non-linear behavior within a defined object. To correctly perform such analysis with modern DL algorithms, it is necessary to use more input data to sufficiently train the algorithm, which is time-consuming and computationally expensive. In addition, the number of parameters increases, especially when working on high-dimensional data (HS data). Particularly time-consuming were analyses that relied on 3D convolutional operations. Therefore, we compared 2D and 3D segmentation methods in Paper B (Kuras et al., 2022c). However, we could not unequivocally state that 3D segmentation always achieves better results because many factors affect the algorithm's functionality in the algorithmic process.

Data quality

Multisensor fusion refers to the analysis of data from various sensors. These data are acquired in different ways, under different weather conditions, and at different times of the year. Since research on multisensor fusion is growing, the quality of the data should be high, avoiding, e.g., in HS data acquisition, low solar angles that cause low signal, clouds, and technical errors, which are caused by incorrect adjustment of integration time, altitude of the aircraft, and its speed.

Complexity of the methods

The operation of the sensors and data acquisition, as well as the preprocessing and processing of data from different sensors at an airborne scale, are highly complex. In order to combine HS images with magnetics in HSM integration and compare them, the spatial resolution of HS images had to be reduced twice to discretize magnetic modeling (Kuras et al., 2022b). The resolution of magnetics was increased using minimum curvature gridding, which may introduce new artifacts. Another example is the complex process in the multitemporal analysis of HL-Fusion data (Kuras et al., 2022a). This approach is very advanced and complicated, ranging from endmember extraction to segmentation and change detection framework.

Data acquisition costs

In order to analyze airborne-based remote sensing data, an expensive flight campaign must be arranged with a professional crew. Until now, HS imagery and magnetic data have been collected locally; thus, the available data are limited or confidential. In Norway, however, there are already publicly available LiDAR data at www.hoydedata.no/LaserInnsyn/, mainly for extracting elevation information and its derivatives.

Work organization – multidisciplinary research

This dissertation on multisensor data fusion was linked to various scientific fields such as physics, mathematics, building physics, remote sensing, geology, and geophysics, as well as various institutions, e.g., Norwegian municipalities, national mapping agency, and scientific institutes from different European countries, among others. Each group operated with different terminology and had varied priorities, which was undoubtedly a unique lesson but also required a lot of commitment and time.

Transferability/generalization

All analyses on data used for classification in the urban environment (Chapter 2.1.1.) and geological characterization (Chapter 2.1.2.) were performed locally. The analyses focused on the general process chain and small data set, which may provide good regional results; however, generalization and transferability were neglected in Paper B (Kuras et al., 2022c) and Paper D (Kuras et al., 2022b).

4.3. Future Perspectives

In the near future, the demand for multisensor data fusion is expected to grow due to technological advances, scientific curiosity to enhance comprehension of the functioning of the Earth, and the only possibility to introduce solutions that a single sensor cannot provide. In this Chapter, we will focus on a recommendation to overcome some of the limitations mentioned in Chapter 4.2.

In the feature-level fusion of HS data which in recent years is mainly based on (supervised) DL algorithms, it is advisable to continue research on unsupervised classification, which will not only save time but also avoid the errors made when

manually creating ground truths. As remote sensing methods become more accessible, more areas are being scanned multitemporal. Referring to this, multitemporal analysis with multisensor fused data from one study of interest over many years requires research in detecting mobile and static objects and high and low frequent changes in the urban environment. The optimal solution for HS fusion with LiDAR data would be an analysis based on a (hyperspectral) 3D point cloud. With such analysis, the classification task would gain more reliability, stability, accuracy, and closeness to reality. In this dissertation, multisensor data fusion was performed only at an airborne scale. Future studies based on application-level data fusion should include multiscale analysis considering satellite, aerial, drone scale, handheld, and laboratory measurements. This is particularly important in scene analysis where the complexity of classes varies, such as analyzing land cover surfaces combined with biochemical and biophysical parameters of the materials. Further research on multisensor data fusion, due to its great potential, versatility, and possibilities, should include physical parameters in the fusion process to create generic automatic methods for selected analysis tasks.

Bibliography

Asner, G.P., D.E. Knapp, J. Boardman, R.O. Green, T. Kennedy-Bowdoin, M. Eastwood, R.E. Martin, C. Anderson, and C.B. Field. 2012. 'Carnegie Airborne Observatory-2: Increasing science data dimensionality via high-fidelity multi-sensor fusion', *Remote Sensing of Environment*, 124: 454-65.

Bedini, E., and T.M. Rasmussen. 2018. 'Use of airborne hyperspectral and gamma-ray spectroscopy data for mineral exploration at the Sarfartoq carbonatite complex, southern West Greenland', *Geosciences Journal*, 22: 1-11.

Ben-Dor, E. 2001. 'Imaging spectrometry for urban applications.' in F.D. van der Meer and S.M. de Jong (eds.), *Imaging spectrometry* (Kluwer Academic Publishers: The Netherlands).

Brell, M., C. Rogass, K. Segl, B. Bookhagen, and L. Guanter. 2016. 'Improving sensor fusion: a parametric method for the geometric coalignment of airborne hyperspectral and LiDAR data', *IEEE Transactions on geoscience and remote sensing*, 54.

Brell, M., K. Segl, L. Guanter, and B. Bookhagen. 2019. '3D hyperspectral point cloud generation: fusing airborne laser scanning and hyperspectral imaging sensors for improved object-based information extraction', *ISPRS Journal of Photogrammetry and Remote Sensing*, 149: 200-14.

Camps-Valls, G., and L. Bruzzone. 2005. 'Kernel-based methods for hyperspectral image classification', *IEEE Transactions on geoscience and remote sensing*, 43.

Chen, Y., C. Li, P. Ghamisi, X. Jia, and Y. Gu. 2017. 'Deep fusion of remote sensing data for accurate classification', *IEEE Geoscience and Remote Sensing Letters*, 14: 1253-57.

Chen, Y., C. Li, P. Ghamisi, C. Shi, and Y. Gu. 2016. "Deep fusion of hyperspectral and LiDAR data for thematic classification " In *International Geoscience and Remote Sensing Symposium* Beijing, China: IEEE.

Clark, D.A. 1997. 'Magnetic petrophysics and magnetic petrology: aids to geological interpretation of magnetic surveys', *Journal of Australian Geology and Geophysics*, 17: 83-104.

Clark, R.N. 1999. 'Spectroscopy of rocks and minerals, and principles of spectroscopy.' in A.N. Rencz (ed.), *Manual of remote sensing, remote sensing for the earth sciences* (John Wiley and Sons: New York).

Cocks, T., R. Jenssen, A. Stewart, I. Wilson, and T. Shields. 1998. "The HyMap airborne hyperspectral sensor: The system, calibration and performance." In *1st EARSeL Workshop on Imaging Spectroscopy*. Zurich, Switzerland.

Dabbiru, L., S. Samiappan, R.A.A. Nobrega, J.A. Aanstoos, N.H. Younan, and R.J. Moorhead. 2015. "Fusion of synthetic aperture radar and hyperspectral imagery to detect impacts of oil spill in Gulf of Mexico." In *IEEE International Geoscience and Remote Sensing Symposium*, 1901-04. Milan, Italy: IEEE.

Dalponte, M., L. Bruzzone, and D. Gianelle. 2008. 'Fusion of hyperspectral and LiDAR remote sensing data for classification of complex forest areas', *IEEE Transactions on geoscience and remote sensing*, 46.

Debes, C., A. Merentitis, R. Heremans, J. Hahn, N. Frangiadakis, T. van Kasteren, W. Liao, R. Bellens, A. Pizurica, S. Gautama, W. Philips, S. Prasad, Q. Du, and F. Pacifici. 2014. 'Hyperspectral and LiDAR data fusion: outcome of the 2013 GRSS data fusion contest', *IEEE Journal of selected topics in applied earth observations and remote sensing*, 7.

Dentith, M., and S.T. Mudge. 2014. *Geophysics for the Mineral Exploration Geoscientist* (Cambridge University Press).

Dian, Y., Y. Pang, Y. Dong, and Z. Li. 2016. 'Urban tree species mapping using airborne LiDAR and hyperspectral data', *Indian Society of Remote Sensing*, 44: 595-603.

Duque-Arias, D., S. Valesco-Forero, J.E. Deschaud, F. Goulette, A. Serna, E. Decenciere, and B. Marcotegui. 2021. "On power Jaccard losses for semantic segmentation." In *16th International Conference on Computer Vision Theory and Applications*. Vienna, Austria.

Durrant-Whyte, H., and T.C. Henderson. 2008. 'Multisensor Data Fusion.' in B. Siciliano and O. Khatib (eds.), *Springer Handbook of Robotics* (Springer Berlin: Heidelberg).

Essa, K.S., and Z.E. Diab. 2022. 'Magnetic data interpretation for 2D dikes by the metaheuristic bat algorithm: sustainable development cases', *Scientific Reports*, 12: 1-29.

Feng, Q., D. Zhu, J. Yang, and B. Li. 2019. 'Multisource Hyperspectral and LiDAR Data Fusion for Urban Land-Use Mapping based on a Modified Two-Branch Convolutional Neural Network', *ISPRS International Journal of Geo-Information*, 8.

Geotech. 2012 "Report on a helicopter-borne versatile time-domain electromagnetic (VTEMplus) and horizontal magnetic gradiometer geophysical survey. Niaqomarssuit Block, Greenland." In.

Ghamisi, P., B. Höfle, and X.X. Zhu. 2016. 'Hyperspectral and LiDAR data fusion using extinction profiles and deep convolutional neural network ', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*

Gothenborg, J., and L. Keto. 1977. "Report on the aerial reconnaissance between Sukkertoppen Ice Calot and Nordenskiölds Gletscher. In archives of Geological Survey of Denmark and Greenland. GEUS Report File 20210." In. Copenhagen, Denmark: Kryolitselskabet Øresund A/S.

Gu, Y., Y. Zhang, D. You, and Y. Zhang. 2012. 'Representative multiple kernel learning for classification in hyperspectral imagery', *IEEE Transactions on geoscience and remote sensing*.

Gupta, R. P. 2003. *Remote sensing geology* (Springer-Verlag Berlin Heidelberg).

Hakala, T., J. Suomalainen, S. Kaasalainen, and Y. Chen. 2012. 'Full waveform hyperspectral LiDAR for terrestrial laser scanning', *Optics Express*, 20.

Hall, D.L., and J.Llinas. 1997. 'An introduction to multisensor fusion', *Proceedings of the IEEE*, 85: 6-23.

Hang, R., Y. Li, P. Ghamisi, D. Hong, G. Xia, and Q. Liu. 2020. 'Classification of Hyperspectral and LiDAR Data Using Coupled CNNs', *IEEE Transactions on geoscience and remote sensing*, 68: 4939-50.

Hasani, H., F. Samadzadegan, and P. Reinartz. 2017. 'A metaheuristic feautre-level fusion strategy in classification of urban area using hyperspectral imagery and LiDAR data', *European Journal of Remote Sensing*, 50: 222-36.

Hinze, W.J., and R.R.B. von Frese. 1990. 'Magnetics in geoexploration', *Proceedings of the Indian Academy of Sciences*: 515-47.

Hsu, S.M., and H.H.K. Burke. 2003. 'Multisensor fusion with hyperspectral imaging data: Detection and classification', *Lincoln Laboratory Journal*, 14.

Hughes, G.F. 1968. 'On the mean accuracy of statistical pattern recognizers', *IEEE Transactions on Information Theory*, 14: 55-63.

Hui, L., L. Di, H. Xianfeng, and L. Deren. 2008. "Laser intensity used in classification of LiDAR point cloud data." In *International Symposium on Geoscience and Remote Sensing*. Boston, MA, USA: IEEE.

Hunt, C.P., B.M. Moskowitz, and S.K. Banerjee. 1995. 'Magnetic properties of rocks and minerals', *American Geophysical Union*.

Jaccard, P. 1912. 'The distribution of the flora in the alpine zone', *New phytologist*, 11: 37-50.

Jackisch, R., B.H. Heincke, R. Zimmermann, E.V. Sørensen, M. Pirttijärvi, M. Kirsch, H. Salmirinne, S. Lode, U. Kuronen, and R. Gloaguen. 2022. 'Drone-based magnetic and multispectral surveys to develop a 3D model for mineral exploration at Qullissat, Disko Island, Greenland', *Solid Earth*, 13: 793-825.

Jackisch, R., S. Lorenz, M. Kirsch, R. Zimmermann, L. Tusa, M. Pirttijärvi, A. Saartenoja, H. Ugalde, Y. Madriz, M. Savolainen, and R. Gloaguen. 2020. 'Integrated Geological and Geophysical Mapping of a Carbonatite-Hosting Outcrop in Siilinjärvi, Finland, Using Unmanned Aerial Systems', *Remote Sensing*, 12.

Jackisch, R., Y. Madriz, R. Zimmermann, M. Pirrtijärvi, A. Saartenoja, B.H. Heincke, H. Salmirinne, J.P. Kujasalo, L. Andreani, and R. Gloaguen. 2019. 'Drone-Borne Hyperspectral and Magnetic Data Integration: Otnmäki Fe-Ti-V Deposit in Finland', *Remote Sensing*, 11.

Jürgens, C. 2001. "Urban and suburban growth assessment with remote sensing." In *OICC 7th International Seminar on GIS Applications in Planning and Sustainable Development*, 13-15. Cairo, Egypt.

Kaufman, J.R., M.T. Eismann, and M. Celenk. 2015. 'Assessment of spatial x2013:spectral feature level-fusion for hyperspectral target detection', *IEEE Journal of selected topics in applied earth observations and remote sensing*, 8: 2534-44.

Khaleghi, B., A. Khamis, F. Karray, and S.N. Razavi. 2013. 'Multisensor Data Fusion: A Review of the State-of-the-art', *Information Fusion*, 14.

Khodadadzadeh, M., J. Li, S. Prasad, and A. Plaza. 2015. 'Fusion of Hyperspectral and LiDAR Remote Sensing Data Using Multiple Feature Learning', *IEEE Journal of Selected Topics in Applied Earth Obersvations and Remote Sensing*, 8: 2971-83.

Kruse, F.A., J.W. Boardman, A.B. Lefkoff, J.M. Young, K.S. Kierein-Young, T.D. Cocks, R. Jenssen, and P.A. Cocks. 2000. "HyMap: an Australian hyperspectral sensor solving global problems-results from USA HyMap data acquisitions." In *the 10th Australasian Remote Sensing and Photogrammetry Conference*, 18-23.

Kuras, A., M. Brell, K.H. Liland, and I. Burud. 2022a. 'Multitemporal Feature-Level Fusion on Hyperspectral and LiDAR Data in the Urban Environment', *Submitted to Remote Sensing Journal*.

Kuras, A., M. Brell, J. Rizzi, and I. Burud. 2021. 'Hyperspectral and Lidar Data Applied to the Urban Land Cover Machine Learning and Neural-Network-Based Classification: A Review', *Remote Sensing*, 13.

Kuras, A., B.H. Heincke, S. Salehi, C. Mielke, N. Köllner, C. Rogass, U. Altenberger, and I. Burud. 2022b. 'Integration of Hyperspectral and Magnetic Data for Geological Characterization of the Niaqornarssuit Ultramafic Complex in West-Greenland', *Remote Sensing*, 14.

Kuras, A., A. Jenul, M. Brell, and I. Burud. 2022c. 'Comparison of 2D and 3D semantic segmentation in urban areas using fused hyperspectral and lidar data', *Journal of Spectral Imaging*, 11.

Li, H., P. Ghamisi, U. Soergel, and X.X. Zhu. 2018. 'Hyperspectral and LiDAR fusion using deep three-stream convolutional neural networks', *Remote Sensing*, 10.

Li, S., W. Song, L. Fang, Y. Chen, P. Ghamisi, and J.A. Benediktsson. 2019. 'Deep learning for hyperspectral image classification: an overview', *IEEE Transactions on geoscience and remote sensing*, 57: 6690-709.

Liao, W., A. Pizurica, R. Luo, and W. Philips. 2017. 'A comparison on multiple level features for fusion of hyperspectral and LiDAR data', *Joint Urban Remote Sensing Event (JURSE)*: 1-4.

Lin, T.Y., P. Goyal, R. Girshick, K. He, and P. Dollar. 2018. 'Focal loss for dense object detection', *Computer Vision and Pattern Recognition*.

Luo, R., W. Liao, H. Zhang, L. Zgang, P. Scheunders, Y. Pi, and W. Philips. 2017. 'Fusion of Hyperspectral and LiDAR Data for Classification of Cloud-Shadow Mixed Remote Sensed Scene ', *IEEE Journal of selected topics in applied earth observations and remote sensing*.

Man, Q., P. Dong, and H. Guo. 2015. 'Pixel- and feature-level fusion of hyperspectral and lidar data for urban land-use classification', *International Journal of Remote Sensing*, 36: 1618-44.

Manolakis, D., R. Lockwood, and T. Cooley. 2016. *Hyperspectral Imaging Remote Sensing: Physics, Sensors, and Algorithms* (Cambridge University Press: Cambridge).

Meer, F.D. van der, and S.M. de Jong. 2001. *Imaging spectrometry: basic principles and prospective applications* (Kluwer Academic Publishers).

Mielke, C., N.K. Boesche, C. Rogass, H. Kaufmann, C. Gauert, and M. de Wit. 2014. 'Spaceborne Mine Waste Mineralogy Monitoring in South Africa, Applications for Modern Push-Broom Missions: Hyperion/OLI and EnMAP/Sentinel-2', *Remote Sensing*, 6: 6790-816.

Morchhale, S., V.P. Pauca, R.J. Plemmons, and T.C. Torgersen. 2016. "Classification of pixel-level fused hyperspectral and LiDAR data using deep convolutional neural networks." In *8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 1-5. Los Angeles, CA.

Ortenberg, F. 2011. 'Hyperspectral Sensor Characteristics.' in A. Huete, J.G. Lyon and P.S. Thenkabail (eds.), *Fundamentals, sensor systems, spectral libraries, and data mining for vegetation* (CRC Press).

Østergaard, C. 2011. "21st North - 2010 Field Work Qaqortorsuaq (Ikertoq)." In, 99.

Paoletti, M.E., J.M. Haut, J. Plaza, and A. Plaza. 2019. 'Deep learning classifiers for hyperspectral imaging: a review', *ISPRS Journal of Photogrammetry and Remote Sensing*, 158: 279-317.

Ronneberger, O., P. Fischer, and T. Brox. 2015. "U-Net: Convolutional Networks for Biomedical Image Segmentation." In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234-41.

Ruffin, C., and R.L. King. 1999. "The analysis of hyperspectral data using Savitzky-Golay filtering-Theoretical basis (part 1)." In *IEEE 1999 International Geoscience and Remote Sensing Symposium*, edited by IGARSS'99 Proceedings: Remote Sensing of the System Earth—A Challenge for the 21st Century. Hamburg, Germany: IEEE.

Salehi, S., and S. M. Thaarup. 2017. 'Mineral mapping by hyperspectral remote sensing in West Greenland using airborne, ship-based and terrestrial platforms': 47-50.

Santara, A., K. Mani, P. Hatwar, A. Singh, A. Garg, K. Padia, and P. Mitra. 2016. 'BASS Net: band-adaptive spectral-spatial feature learning neural network for hyperspectral image classification', *IEEE Transactions on geoscience and remote sensing*.

Senchuri, R., A. Kuras, and I. Burud. 2021. "Machine Learning Methods for Road Edge Detection on Fused Airborne Hyperspectral and LIDAR Data." In *11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. Amsterdam, Netherlands: IEEE.

Sharma, R., V.I. Pavlovic, and T.S. Huang. 1998. 'Toward multimodal human-computer interface', *Proceedings of the IEEE*, 86: 853-69.

Signoroni, A., M. Savardi, A. Baronio, and S. Benini. 2019. 'Deep learning meets hyperspectral image analysis: a multidisciplinary review', *Journal of Imaging*, 5.

Singh, M.K.K., S. Mohan, and B. Kumar. 2022. 'Fusion of hyperspectral and LiDAR data using sparse stacked autoencoder for land cover classification with 3D-2D convolutional neural network', *Journal of Applied Remote Sensing*, 16: 034523.

Teo, T.A., and H.M. Wu. 2017. 'Analysis of land cover classification using multi-wavelength LiDAR system', *Applied Sciences*.

Till, J.L., and N. Nowaczyk. 2018. 'Authigenic magnetite formation from goethite and hematite and chemical remanent magnetization acquisition', *Geophysical Journal International*, 213: 1818-31.

Torabzadeh, H., F. Morsdorf, and M.E. Schaepman. 2014. 'Fusion of imaging spectroscopy and airborne laser scanning data for characterization of forest ecosystems', *ISPRS Journal of Photogrammetry and Remote Sensing*, 97: 25-35.

Tóvári, D., and T. Vögtle. 2012. 'Object classification in laserscanning data', *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 36.

Tukiainen, T., and L. Thorning. 2005. 'Detection of kimberlitic rocks in West Greenland using airborne hyperspectral data: the HyperGreen 2002 project', *Greenland Bulletin, Geological Survey of Denmark*, 7: 69-72.

Wehr, A., and U. Lohr. 1999. 'Airborne laser scanning - an introduction and overview', *ISPRS Journal of Photogrammetry and Remote Sensing*, 54: 68-82.

Wei, L., K. Wang, Q. Lu, Y. Liang, H. Li, Z. Wang, R. Wang, and L. Cao. 2021. 'Crops fine classification in airborne hyperspectral imagery based on multi-feature fusion and deep learning', *Remote Sensing*, 13: 2917.

Yan, W.Y., N. El-Ashmawy, and A. Shaker. 2015. 'Urban land cover classification using airborne LiDAR data: a review', *Remote Sensing of Environment*.

Yang, X., X. Li, Y. Ye, X. Zhang, H. Zhang, X. Huang, and B. Zhang. 2019. "Road detection via deep residual dense u-net." In *International Joint Conference on Neural Networks*. Budapest, Hungary.

Zhang, M., W. Li, Q. Du, L. Gao, and B. Zhang. 2020. 'Feature Extraction for Classification of Hyperspectral and LiDAR Data Using Patch-to-Patch CNN', *IEEE Transactions on cybernetics*, 50: 100-11.

Zhang, Z., Q. Liu, and Y. Wang. 2018. 'Road extraction by deep residual u-net', *IEEE Geoscience and Remote Sensing Letters*, 15.

Zhao, X., R. Tao, W. Li, H.C. Li, Q. Du, W. Liao, and W. Philips. 2020. 'Joint Classification of Hyperspectral and LiDAR Data Using Hierarchical Random Walk and Deep CNN Architecture', *IEEE Transactions on geoscience and remote sensing*, 58: 7355-70.

Zhou, Z., and J. Gong. 2018. 'Automated residential building detection from airborne LiDAR data with deep neural networks', *Advanced Engineering Informatics*, 36: 229-41.

Appendices

Paper A:

Hyperspectral and Lidar Data Applied to the Urban Land Cover Machine Learning and Neural-Network-Based Classification: A Review

Hyperspectral and Lidar Data Applied to the Urban Land Cover Machine Learning and Neural-Network-Based Classification: A Review

Agnieszka Kuras [1], Maximilian Brell [2], Jonathan Rizzi [3] and Ingunn Burud [1]

[1]   *Faculty of Science and Technology, Norwegian University of Life Sciences, PB 5003, 1430 Aas, Norway*

[2]   *Helmholtz Center Potsdam, GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany*

[3]   *Norwegian Institute of Bioeconomy Research, Raveien 9, 1430 Aas, Norway*

Hyperspectral and Lidar Data Applied to the Urban Land Cover Machine Learning and Neural-Network-Based Classification: A Review

Agnieszka Kuras [1,*], Maximilian Brell [2], Jonathan Rizzi [3] and Ingunn Burud [1]

[1] Faculty of Science and Technology, Norwegian University of Life Sciences, PB 5003, 1430 Aas, Norway; ingunn.burud@nmbu.no

[2] Helmholtz Center Potsdam, GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany; maximilian.brell@gfz-potsdam.de

[3] Norwegian Institute of Bioeconomy Research, Raveien 9, 1430 Aas, Norway; jonathan.rizzi@nibio.no

* Correspondence: agnieszka.kuras@nmbu.no

Abstract: Rapid technological advances in airborne hyperspectral and lidar systems paved the way for using machine learning algorithms to map urban environments. Both hyperspectral and lidar systems can discriminate among many significant urban structures and materials properties, which are not recognizable by applying conventional RGB cameras. In most recent years, the fusion of hyperspectral and lidar sensors has overcome challenges related to the limits of active and passive remote sensing systems, providing promising results in urban land cover classification. This paper presents principles and key features for airborne hyperspectral imaging, lidar, and the fusion of those, as well as applications of these for urban land cover classification. In addition, machine learning and deep learning classification algorithms suitable for classifying individual urban classes such as buildings, vegetation, and roads have been reviewed, focusing on extracted features critical for classification of urban surfaces, transferability, dimensionality, and computational expense.

Keywords: machine learning; deep learning; lidar; hyperspectral; remote sensing; urban environment; data fusion; sensor fusion; urban mapping; land cover classification

1. Introduction

Over the last few decades, global urbanization has grown rapidly. By 2050, around 68% of the world`s population will be living in urban areas [1]. This can cause environmental challenges, including ecological problems, poor air quality, deterioration of public health, microclimate changes leading to severe weather, higher temperatures, limited access to water, persistent vulnerability to natural hazards, and the release of toxic particles from fast industrialization into the atmosphere [2,3]. These challenges lead to difficulties in advanced urban analyses due to urban surfaces' spectral and structural diversity and complexity over a small area [4,5]. Therefore, constant monitoring of urban areas is often highly required. Systematic monitoring and updating of maps are critical in urban areas, where many objects are mobile (vehicles and temporary buildings), and the infrastructure, vegetation, and construction are constantly changing.

Spatiotemporal investigations of the urban regions are today provided by remote sensing technology advances [6]. Especially, airborne remote sensing is a powerful developing tool for urban analysis that offers time-efficient mapping of a city essential for diverse planning [7], management activities [8], and monitoring urban and suburban land uses [9]. It has been proven as a common technique for mapping urban land cover changes to investigate, e.g., social preferences, the regional ecosystem, urbanization change, and biodiversity [10]. Urban remote sensing, in particular, is widely used for the investigation

of three-dimensional urban geometry that is crucial for modeling urban morphology [11], identifying various objects, heterogeneous material, and mixtures. However, the growing challenges require a state-of-the-art technological solution in terms of sensors and analysis methods. Continuous development and improvement of remote sensing sensors increase interest in identifying urban land cover types based on spectral, spatial, and structural properties [12,13]. In urban mapping, lidar analyses (light detection and ranging), hyperspectral data (HS), and synthetic aperture radar (SAR) have become significant. Different portions of the electromagnetic spectrum are useful in analyzing urban environments from the reflective spectral range to the microwave radar [14]. The latter provide high-resolution images independent of the time of day and weather; however, due to the requirement of oblique illumination of the scene, occlusion and layover appear, making the analysis of dynamic urban areas difficult [15].

Urban land cover classification accuracy and interpretability based only on a single sensor in complex, dense urban areas are often insufficient [16]. The heterogeneity in the urban areas leads to high spectral variation within one land cover type, resulting in very complex analyses. The impervious surfaces (roofs, parking lots, roads, and pavements) notably vary in the spectral and spatial-structural manner. In addition, scale and spatial resolution are relevant for estimating urban heterogeneity. Scale defines heterogeneity, in which materials are taken into account analytically or absent or grouped into one class, e.g., individual trees, type versus forest, or vegetation in general [17]. Spatial resolution, on the other hand, determines the level of pixel mixing. However, high spatial resolution increases the physical material heterogeneity, increasing the complexity of analyses.

HS data provide spectral information about materials, differentiating them without elevation context. The challenge in the pure spectral analysis is the negligence of object identification, mostly built from various materials maintaining very high intra-object heterogeneity. By contrast, lidar data can distinguish between different land cover classes from the same material at a different height, such as asphaltic open parking lots and roads [18,19]. Furthermore, passive remote sensors, such as HS, are sensitive to atmospheric conditions and illumination, whereas lidar as an active sensor is less sensitive to these factors. This property of lidar enables, e.g., a physical correction of shadow and illumination purposes when combined with HS data [20–25] and intensity measurement for urban land cover mapping in shaded areas [26]. Regardless of the spatial and spectral resolution of airborne-based HS sensors, urban environments are characterized by spectral ambiguity and reduced spectral value under the shadow caused by topography changes, buildings, and trees, which can be overcome by adding lidar data as presented by [27]. In order to overcome the limitations of individual sensor capabilities, the recent technologies are based on multisensory fusion in the classification of urban surfaces, combining active and passive remote sensing, such as airborne-based lidar and hyperspectral data (HL-Fusion). Such an HL-Fusion can provide complementary information regarding the three-dimensional topography, spatial structure, and spectral information in the landcover classification purposes [19,28–31].

Moreover, a fusion of spectral, spatial, and elevation features provides robust and unique information relevant to the urban environment [30]. The airborne HL-Fusion has already been investigated for urban land cover classification purposes [30,32,33]. However, diverse combination methods are implemented on different data and product levels based on either physical or empirical approaches [34]. Furthermore, since all fusion processes are very complex, there is no defined framework for fusing these sensors. Therefore, a comprehensive summary of previous research on data fusion may enhance the understanding of fusion possibilities, challenges, and common issues that limit the classification results in the urban environment.

Machine learning (ML) techniques have been applied as classifiers for HS data [23–28]. Depending on the classification aim, different mapping methods are applied to achieve the goal. ML algorithms are in constant improvement, providing algorithms that can hierarchically extract more complex features. This ability is assigned to a subfield of

machine learning as deep learning (DL). DL has been proven as an effective technique for feature extraction of HS data on the spatio-spectral level [35–40]. Although ML and DL methods are considered relevant classification tools in remote sensing, different algorithms perform best, extracting different pixel- or object-based features. Choosing a classification algorithm for HS data requires knowledge about the features that can be extracted. Especially, DL has gained popularity, thanks to finding unique deep parameters in a pixelwise manner [41]. However, in the urban context, a per pixel classification can lead to noisy results considering high spatial distribution.

Moreover, classification results mainly depend on the number of training samples, limiting the performance and accuracy when the training dataset is insufficient for learning the network algorithm [42]. In order to improve the classification results and reduce the heterogeneity issue, the inclusion of contextual information around pixels and object-oriented classification [43,44] were considered, which allowed retrieving spatial information of HS data and extracting critical spatial patterns of urban land cover classes [45,46]. ML- and DL-based land cover classification in the urban environment from lidar is primarily directed to detect buildings or high vegetation [47]. This is due to the lidar`s ability to extract geometric features from objects, deriving their shape, elevation, and other properties that are useful for a classification purpose. Especially, lidar, in combination with HS, is a powerful tool for classifying urban materials. However, since the objects in the urban scene are complex, analyses with conventional classifiers achieve a low accuracy [48]. Commonly, the application of ML and DL algorithms for classification purposes in the urban environment outperforms traditional classifiers developing very quickly [49].

This review study presents the latest ML and DL urban mapping methods focusing on airborne HS and lidar data. The datasets cover the reflective spectral range of the electromagnetic spectrum (VNIR, SWIR). The paper focuses on ML and DL classification algorithms applied in the urban environment for land cover classes, such as buildings, roads, vegetation, and water analysis. We point out algorithms applicable for HS, lidar, and HL-Fusion and the challenges of applying each algorithm to hyperspectral and lidar data.

The structure of this review paper is arranged as follows (Figure 1): in Section 2, typical urban land cover classes are described in terms of their complexity in HS and lidar data analysis. Section 3 synthesizes the general characteristics of HS and lidar data, highlighting the automated and handcrafted features extracted from both sensors. In Section 4, classification algorithms for urban mapping purposes are described. Section 5 shows the results and discussion of the presented algorithms in urban environment classification. Finally, we point out conclusive remarks on the mapping methods, HL-Fusion potential, perspectives for further research, and recommendations for new research fields.
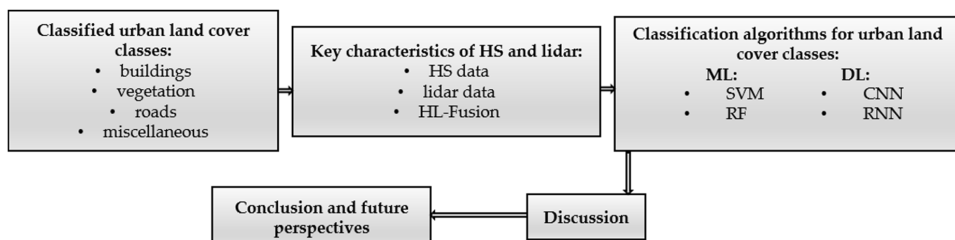
Figure 1. The structure of the review paper.

2. Classified Urban Land Cover Classes

The urban land cover consists of very complex physical materials and surfaces that are constantly having anthropological impacts. The urban surface types are a mosaic of seminatural surfaces such as grass, trees, bare soil, water bodies, and human-made mate-

rials of diverse age and composition, such as asphalt, concrete, roof tiles for energy conservation and fire danger [50], and generally impervious surfaces for urban flooding studies and pollution [51]. The complexity of urban analysis also depends on the scale chosen and its purpose. Many classifications refer to urban materials with fine spatial resolution deepening the heterogeneity, allowing a more detailed mapping result. The classification of urban objects, which consist of many different materials and variance within a class, although significant (e.g., in city map updates), becomes a challenge due to the highly nonlinear and heterogeneous composition of different objects surfaces and materials, and thus, there is the need to use more training data for classification purposes, which is time-consuming and computationally expensive.

2.1. Buildings

Buildings in an urban context can be recognized as shapes with planar surfaces and straight lines [52]. Building detection based on remote sensing methods plays a crucial role in many applications in the urban environment, such as in 3D monitoring of urban development in time [53], urban planning, telecommunication network planning, vehicle navigation [33], urban energy planning [53], city management, and damage assessment [54]. Many mapping techniques are based on shape identification, outlines, and preliminary model data [54]. Besides detecting buildings as objects, building roof extraction has recently been a hot topic within the remote sensing community. Building roofs are defined by planarity properties and height derivatives based on elevation. A 3D visualization of buildings is of great importance for infrastructure management and modeling, 3D city mapping, simulations, change detection, and more [55]. Both airborne-based optical and lidar data have been used recently to map buildings. A common way to detect buildings is to use a digital surface model (DSM) [56,57], a normalized DSM (nDSM) [58,59], or a point cloud extracted from lidar data [60–63]. Lidar is capable of extracting building heights and planar roof faces [33]. It is beneficial for spatiotemporal assessment and investigation of building density for sustainability study and residential development in cities [53].

By contrast, airborne-based HS data can better distinguish between materials at the roof surfaces due to their spectral differences [33]. However, not including the elevation information from the lidar scanner, the classification of buildings and their roofs can be too complex without human expertise. One example is a building surrounded by an arid lawn with open soil, a grass rooftop, a building with an asphaltic parking lot, or bitumen roofing surrounded by asphaltic parking at the building's ground-level high vegetation (trees) overhanging buildings [64]. Therefore, an HL-Fusion can improve the building classification results offering high accuracy on a spectral and spatial basis.

2.2. Vegetation

Vegetation is recognized by its geometrical complexity, defined by parameters such as the roughness, point density ratio measure [65], and chlorophyll spectral feature. In the last decade, active (Sentinel-1, LiDAR, and radar) and passive (Quickbird, Worldview, Sentinel-2, Landsat, and MODIS) remote sensing has been widely applied to vegetation detection. Lidar data are used to generate virtual 3D tree models [66], map low and high vegetation [67], and, using multispectral lidar, assess vegetation variety regarding its health and density [68], as well as extract vegetation indices, e.g., NDVI [69] for monitoring changes caused by urbanization, anthropogenetic activities, and harvesting applying wavelet transform [70,71]. However, vegetation detection is not a straightforward approach. The analysis is often complex and detailed due to the increasingly finer spatial resolution of remote sensing devices, such as distinguishing photosynthetic and nonphotosynthetic vegetation [72]. Vegetation is often not defined as a whole but as groups, for example, as low vegetation (grass), middle vegetation (shrubs), and high vegetation (trees). One of the more complex challenges is the similar morphology of low/young trees and shrubs, causing misclassification of shrubs as high trees [73]. HS data are also used to

detect vegetation on a spectral basis (chlorophyll reflectance), differentiating between vegetation types and healthiness. More biophysical parameters can be defined due to more spectral bands than multispectral lidar (usually 2–3 wavelengths), such as the leaf area index, fractional cover, and foliage biochemistry [74]. Both sensors have been fused in many studies, e.g., for canopy characterization for biomass assessment and estimation of risk of natural hazards [75] and urban tree species mapping [76].

2.3. Roads

Road detection from airborne-based HS and lidar data is essential in remote sensing applications, e.g., a road navigation system, urban planning and management, and geographic information actualization [77,78]. The elevation feature derived from lidar data has been proven as a significant parameter to extract time-efficient road methods compared to optical methods [79]. DSM distinguishes more precise boundaries of surfaces, even in occluded regions [80]. However, only lidar-data-based classification is limited when roads are at the same elevation but made of different materials, such as asphalt, concrete, or other impervious materials [18]. Therefore, HS imaging can differentiate between different materials and their conditions to complement road classification purposes. It has already been proven by Herold et al. [81] for the following uses: map alteration, degradation, and structural damages of road surfaces based on spectral analysis. Usually, to detect roads, texture information is implemented [82]. In addition, lane marks can be used as an indicator for new roads; however, this approach is illumination sensitive [83]. HS data classification without topographic information is challenging when differentiating between two objects made from the same material: differentiation between a parking lot, parking at the ground level, cycleway, and a road [30].

2.4. Miscellaneous

Apart from the above-described land cover classes, the urban environment consists of more complex thematic classes. They commonly cannot be chemically or physically described by a single hyperspectral absorption feature or other single features, such as height or shape, which are, however, extracted from contextual information. Thus, spatial context is critical and necessary for identifying industrial areas, commercial or residential buildings, playgrounds, and harbors in coastal cities. The combination of spectral and spatial features from HS and lidar data shows potential, allowing identifying thematic class and assessing its condition in terms of quality and materials.

3. Key Characteristics of Hyperspectral and Lidar Data

In urban land cover classification, handcrafted feature engineering plays an important role in standard shallow ML algorithms, such as support vector machines (SVM) and random forest (RF). Features are manually derived from remotely sensed data and defined to describe an object of interest, starting from spectral bands through, for example, spectral indices and contextual information, which are generally very useful in defining important biophysical parameters, e.g., for vegetation [84]. However, manually derived features may not sufficiently represent the highly complex and unique urban environment [85]. Depending on the classification objective and classified objects, different features are required. However, in DL, the feature engineering process is simplified as features are extracted during the training step [86]. These automatic high-level features can represent complex spatial correlations and nonlinear relationships. Examples of handcrafted features for both HS and lidar data are described below in this section.

3.1. Hyperspectral (HS) Images

HS data retrieved from an imaging spectrometer are a three-dimensional cube that includes two-dimensional spatial information (x, y) with spectral information at each pixel

position x_iy_j [87]. Each pixel in the obtained digital data contains a nearly continuous spectrum covering the reflective spectral range of the visible, near-infrared (VNIR: 400–1000 nm) and short-wave infrared (SWIR: 1000–2500 nm) [88,89]. HS as a passive system is dependent on the given lighting conditions resulting in high intraclass (within a class) spectral variability. In these wavelength ranges of the electromagnetic spectrum, particular absorption features and shapes make it possible to identify the material`s chemical and physical properties [90]. For example, in urban land cover classification, the reflective spectral range is often used to map diverse soils [91], vegetation [92], rooftop materials [93,94], and other complex physical materials [12,95–97].

A high spectral resolution characterizes airborne-based HS applications at the expense of spatial resolution since the HS sensor's spatial resolution linearly depends on the flight altitude and the instantaneous field of view (IFOW) [98]. However, due to technology development, the spatial resolution of HS is enhanced. Spectrometers with high spectral and spatial resolution have been used to identify detailed urban materials [12,13,94,99]. With a higher spatial resolution of the hyperspectral camera, it is more likely that the spectral signals are less mixed, producing pure pixels and thus detecting materials in the urban environment with high geometric detail and material accuracy. However, a high resolution can lead to difficulties, detecting more diverse materials within a single object, thus increasing heterogeneity and making object-based classification on a coarser scale more challenging. Especially in urban remote sensing, the spatial complexity of the objects and their heterogeneity have been an issue for limited spatial resolution in many studies [94,100]. When within a single pixel, the spectral mixture is very complex, the different spectral properties of individual urban materials are lost, making classification at the level of relevant urban materials challenging [101]. Therefore, a high spatial resolution of hyperspectral sensors has become a crucial parameter in urban mapping.

Land cover classification based on HS data is affected by spatial and spectral resolution, classification purposes (scale and defined land cover classes), mapping methods, and data acquisition and preprocessing. The latter can be the optical geometry, integration time, and other parameters during the acquisition [102]. Especially in airborne-based HS imaging, the sensor experiences altitude variation, which results in geometric distortions in the HS scene [103]. It is always a compromise between off-nadir distortion, spatial resolution, mixed pixels, and SNR (signal-to-noise ratio). Therefore, the strategy and flight scheme must be adapted to the level of the classification target in an advanced way. The flight line's swath width is reduced at a lower altitude, which requires more flight lines to be flown to cover the target area with changing light conditions due to long integration time [104] and leads to higher off-nadir distortions [105]. However, there are challenges for flying at higher altitudes, such as a high degree of mixed pixels due to a low spatial resolution [106]. In addition, the short integration time at lower altitudes results in lower SNR and decreased sensor sensitivity, producing a more elevated noise floor.

3.1.1. Spectral features

Within one material, spectral features can vary due to color, coating, degradation, alteration, roughness, the illumination of material, data acquisition, location of the material, and preprocessing data (Figure 2) [97,107,108]. These variations within a material are more and more investigated, generating spectral libraries of complex urban materials [12,109,110] and normalization based on advanced preprocessing. HS images result in high-dimensional data leading to computationally expensive analyses. For this reason, the first common step of the classification process of the HS data is very often a spectral dimensionality reduction to the relevant components applying linear spectral transformations without losing important spectral information [111]. Standard techniques for dimensionality reduction are often statistically based, such as principal component analysis (PCA) [112], linear discriminant analysis (LDA) [113], multivariate curve resolution (MCR) [114], and other unsupervised classification methods. Such data compression saves computing time, reduces noise, and retains needed information [115]. They are often

based on the individual image statistic, and thus they are not directly transferable to other flight lines or flight campaigns. In addition, quantification procedures based on the spectral signature are no longer possible. Statistical calculations have been applied to the spectral features of the urban materials, such as continuum removal [116,117]. The continuum-removal algorithm is applied to identify spectral absorption features by their wavelength positions and shapes, removing the overall albedo of the reflectance curve and reducing the searched material's superimposition [118]. However, the general shape of an absorption feature is relevant for material identification and quantification. Continuum removal may prove effective only for limited studies, excluding the original shape of the spectra. Some handcrafted target-specific features can be calculated from optical remote sensing data, such as normalized difference vegetation index (NDVI) for vegetation detection [8,18,119,120], new impervious index, road detection index, new roof extraction index for the detection of built-up, roads, and roofs [121], normalized difference built-up index [122], visible red and green near-infrared built-up indices [123], road extraction index [124], and hyperspectral difference water index for the detection of urban water bodies [125].
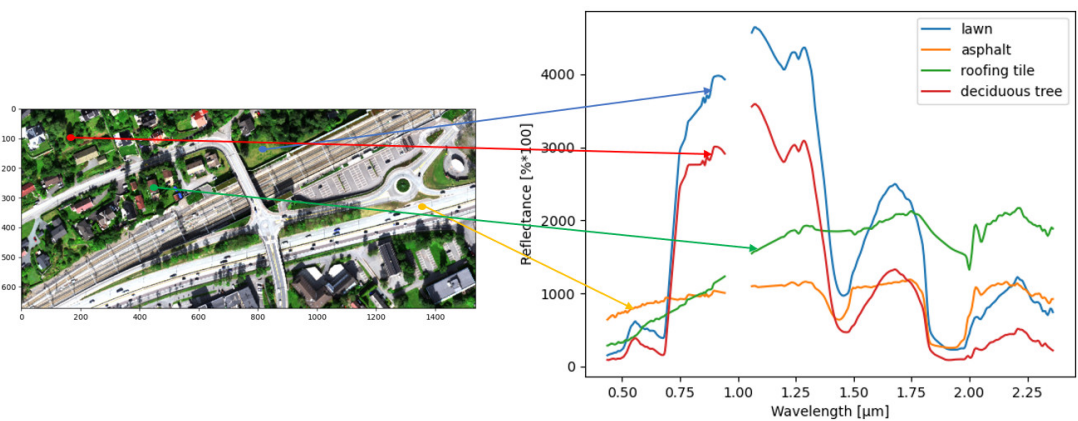
Figure 2. At surface reflectance of some urban surfaces (HySpex sensors VNIR-1800 and SWIR-384). The hyperspectral dataset was acquired by the Terratec AS Company in August 2019 over Baerum municipality, Oslo, Norway.

3.1.2. Spatial Information

Spatial-context information is widely used to achieve robust and accurate classification maps considering the neighborhood in the target pixel. While spectral features are the most relevant features in material-based classification, adding spatial features to object classification makes it easier to group pixels with some spectral variance into one class representing an object or land cover type [126] (see Section 3.3). In addition, the spatial noise of the classification results can be reduced [127,128]. In [129], the authors proposed a context-sensitive semisupervised SVM classification technique using contextual information without assumptions about the labeling of contextual pixels. In [130,131], the authors also added the contextual features into hyperspectral image classification, including the information in the classification map generation step. Spatial information is often incorporated in hyperspectral classification problems applying Markov random field where a predefined neighborhood of a pixel assumes that the central pixel belongs to the same class [36,132,133]. Contextual features can also be extracted considering texture (see Section 3.3.1), morphological features (see Section 3.3.2), and image segmentation.

3.2. Lidar Data

Lidar data is a three-dimensional point cloud (x, y, z) which delivers by default information about elevation, multiple-return, the reflected intensity, texture, and waveform-derived feature spaces from the object hit by laser pulse [31,134]. As an active sensor, a lidar system emits radiation from one bandwidth (more in the case of multiwavelength lidar scanners) to the object surface at high repetition rates. Lidar scanners are whiskbroom-type instruments and typically use the monochromatic laser in visible—532 (bathymetric/coastal mapping)—and near-infrared—1064 and 1550 nm—for example, for vegetation detection and differentiation between asphaltic and nonasphaltic roads [135] which can be used as an additional intensity feature in land cover mapping in the reflective spectral range [31]. The advantage of using airborne lidar is insensitivity to relief displacement and illumination conditions [31], retaining full 3D geometry of data.

3.2.1. Height Features and Their Derivatives (HD)

The height feature is used to calculate the three-dimensional coordinates (x,y,z) that generate a gridded 2.5-dimensional topographical profile of the area of interest [31]. Especially in the urban environment, the z value height is crucial for precise contour generation of elevated objects [31]. In addition, the height difference between the lidar return and the lowest point in cylindrical volume has been investigated and proven as an important feature in discriminating ground and nonground points [136,137]. Moreover, a digital surface model (DSM) (Figure 3A) is extracted from the height information applying interpolation of 3D points onto a 2D grid. From a DSM, a surface roughness layer [138] and a normalized DSM (nDSM) (Figure 3C) are calculated, subtracting the digital terrain model (DTM) (Figure 3B) from the DSM [31]. The overlapping of the building height information and the terrain height information is thus excluded. The object representation heterogeneity is therefore reduced, which helps the classification procedure.
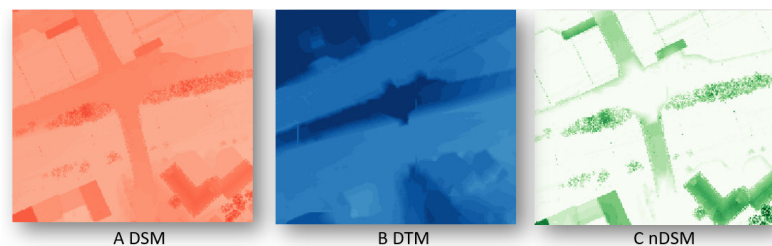
Figure 3. Examples of DSM (**A**), DTM (**B**), and nDSM (**C**) from Riegl VG-1560i LiDAR scanner acquired by the Terratec AS Company in August 2019 over Baerum municipality, Oslo, Norway.

The nDSM represents the above-ground points that correspond to the actual heights of the object, omitting information about the objects which could complicate the classification, for example, the differentiation of buildings in lowland or hilly regions. The height information from lidar data helps differentiate between high and low vegetation [139], tree-level characterization applying the canopy height model (CHM) [140], and roads and buildings in the urban environment [8]. In addition, slope calculation (first derivative of any elevation product) and surface curvature (second derivative of the elevation surface) have been applied for detecting surface roughness [141,142] and changes in the normal vectors of the surface [143]. Moreover, calculated skewness and kurtosis models from the lidar elevation data were applied by Antonarakis et al. [144] to determine planted and natural riparian forests and their ages [32]. In the classification approaches, Charaniya et al. [145] included height variation, Bartels and Wei [146] calculated mean variance and standard derivation of the height in the first echo from lidar to measure the roughness, and Im et al. [147] added homogeneity, contrast, and entropy of height as feature spaces after image segmentation (Figure 4).

3.2.2. Intensity Data

Intensity values extracted from lidar data correspond to the peak amplitudes from the illuminated object [31]. Applying intensity as a feature space, Song et al. [148] presented an approach to determine asphalt roads, grass, house roofs, and trees. However, trees' diverse intensity values undermine the classification due to the canopies` complex geometry [149]. Moreover, lidar-based intensity can differentiate between low vegetation and impervious surfaces, such as built-up areas. MacFacen et al. [150] applied the estimated mean intensity values from a lidar dataset in an object-based image classification approach. Intensity data are unstable and contain artifacts in the overlapping regions of single strips and eccentricity caused by the gain response, sensor scanning, and environmental factors [151–153]. To remove the noise from the intensity data, interpolation, filtering methods, and radiometric calibration are commonly used [148,154]. Additionally, the influence of flying altitude variations, topography, and atmospheric conditions can be corrected, adjusting intensity values, which is called range compensation [155].
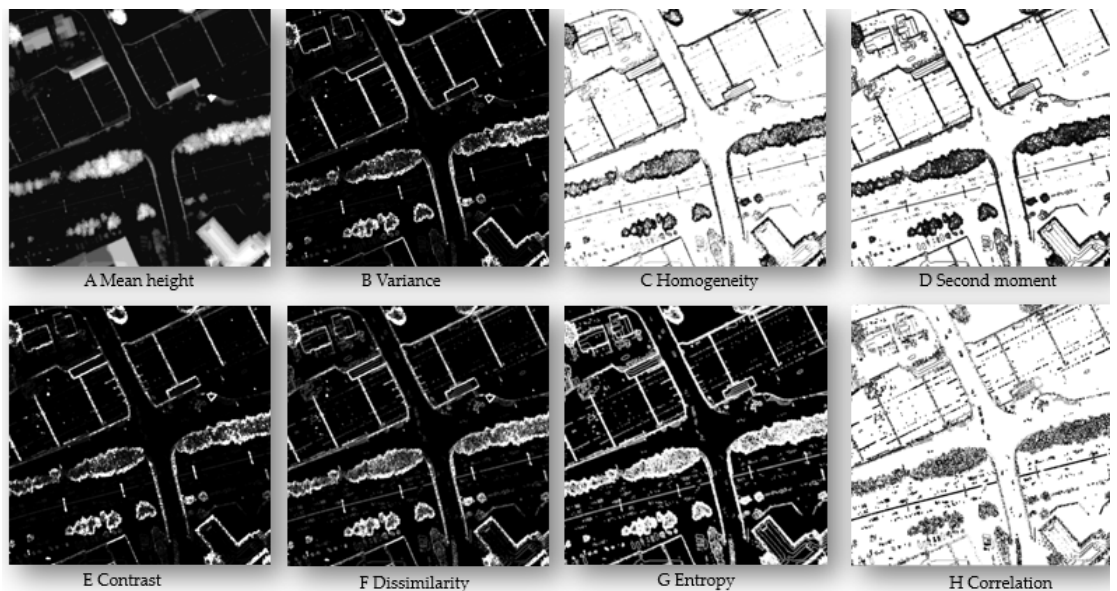
Figure 4. Features derived from the height information from lidar data.

3.2.3. Multiple-Return

A lidar-based laser pulse can split into multiple laser returns if it hits a permeable object such as a tree canopy and obtains a response from, e.g., branches, leaves, stems, and the ground [31]. Multiple-return data has been recently used as an additional feature space in the urban mapping in the commercial building, small house, and tree determination [146]. Charaniya et al. [145] and Samadzadegan et al. [48] extracted the first, the last, and the normalized difference (NDI) between these returns to investigate roads and buildings. However, multiple returns occur as well if the laser pulse reaches building edges [156].

3.2.4. Waveform-Derived Features

Full-waveform lidar scanners can retrieve the entire signal of the backscattered laser pulse as a 1D signal profile in the chronological sequence [134,156,157]. A full-waveform

lidar system can better correct the intensity values than the discrete systems, such as accurate estimation of the surface slope [158], eliminating the assumption of Lambertian reflectors [159]. However, before using any classification approach, proper radiometric calibration is needed to adjust waveform data from different flight campaigns. Such a radiometric calibration should include preflight, onboard, and vicarious calibration, as presented by Wagner [155]. The waveform-derived features extracted from the gaussian decomposition function have been tested for urban mapping purposes [47,136,160,161]. They include the waveform amplitude, (normalized) number of echoes, their width (Gaussian standard deviation), the difference between the first and the last return, echo shape, and echo cross-section. The latter provides high values for buildings, medium values for vegetation, and low values for roads [137]. For building facades and vegetation that meet multiple echoes, the normalized number of echoes feature is, therefore, relevant [137]. Jutzi and Stilla [162] extracted linear features on roofs based on full-waveform data. Chehata et al. [136] provided that by adding echo width as a feature, the classification results improved for low vegetation. Echo shape was investigated by [137,163], providing low values to roofs and high values to vegetation. It has been proven that the waveform geometry helps to differentiate between trees and built-up areas [136,156,164], determine tree species [165,166], and segment lidar point clouds in an urban area [167]. The waveform amplitude depends on the target. High amplitudes were observed by Chehata et al. [136] for rooftops, gravel, cars, bare soil, and grass, and low amplitudes for asphalt, tar street, and water. Mallet and Bretar [156] observed high amplitudes for grass and bare earth and found that the spread in the pulse and low amplitudes can be assigned to flat surfaces by increasing the incident angle. The echo waveform classification has been applied by Lin and Mills [168] and Doneus et al. [169]. The terrain echoes were separated from echoes from bushes and low vegetation. The echo pulse is wider on the canopy surface and plowed field than on the meadow and street [156]. High point density in full-waveform lidar data helps to detect vegetation types and states [170].

3.2.5. Eigenvalue-Based Features

The eigenvalues are calculated based on the covariance matrix of x, y, and z dimensions of the 3D point cloud as λ_1, λ_2, and λ_3. Eigenvalues as features help detect geometrical parameters, such as plane, edge, and corner [171]. The following structure features have been applied to lidar data: omnivariance, anisotropy, planarity, sphericity, linearity, and eigenentropy for features for context-driven target detection [172] building detection [171]. Some of them are shown in Figure 5. The planarity feature is proven relevant for road classification or other flat surfaces and sphericity for building and natural ground (low vegetation) detection [136].

A Omnivariance　　　　　　　　　　　　　　　　B Linearity
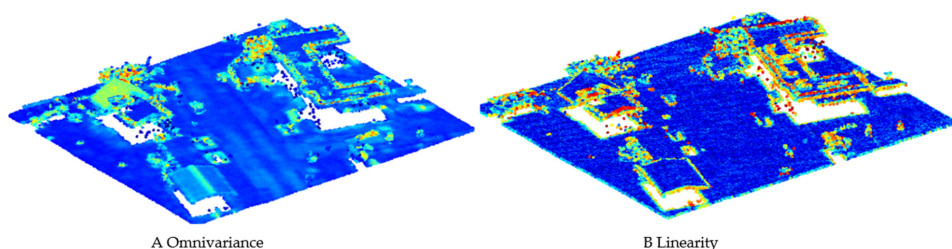
Figure 5. Structure features derived from lidar data: omnivariance (**A**) and linearity (**B**) from [171].

3.3. Common Features—HS and Lidar

3.3.1. Textural Features

Besides spectral information of hyperspectral sensors, pixel-wise spatial features are relevant for image content, such as textural features. The textural attributes in a hyperspectral scene can be extracted by the local binary patterns (LBP) operator proposed by [173], providing information about the surface granularity [174]. To include spatial information in the classification purposes, the textural operators are window based. Peng et al. [175] extracted them as rotation-invariant features for urban classification purposes except for spectral features and Gabor features [176]. The latter are frequential filters interpreting the texture of the hyperspectral bands used by [177,178]. The texture can be analyzed by applying the gray-level co-occurrence matrix (GLCM) measures [53,179]. GLCM measures, first proposed by Haralick et al. [180], consist of energy, contrast, correlation, entropy, and homogeneity. GLCM dissimilarity, entropy, homogeneity, and second-moment help to detect building edges and height differences. However, contrast, correlation, and variance do not improve building classification and temporal change [53]. Texture features have been used to classify urban materials for pattern recognition in lidar, satellite, and airborne data [48,181–184]. Samadzadegan et al. [48] calculated four measures: mean, entropy, standard deviation, and homogeneity to classify trees, buildings, and grounds. Huang et al. [181] applied, except for homogeneity and entropy, the angular second moment and dissimilarity from the DSM in the classification approach.

3.3.2. Morphological Features

Mathematical morphology contains operators such as erosion, dilation, opening, closing, rank filters, top hat, and other derived transforms. Mainly, these operators are applied on panchromatic images from hyperspectral sensors, binary or greyscale images with isotropic and geodesic metrics with a structural element [185]. For example, the opening operator focuses on the bright spots, removing objects smaller than the structural element, whereas the closing operator acts on the dark objects (Figure 6). Morphological features with a structural element contain information about the minimum size of the target being investigated [18]. They help reduce shape noise, enhance edges, interpret the texture and extract structures on images regarding their shapes, orientation, and sizes [185–188]. In image processing, morphological features are based on both spectral and spatial information involving pixels in the neighborhood. They are widely used in hyperspectral image classification [178,187–191], noise reduction in lidar [192], building detection [193], and HL-Fusion-based classification [18]. It has been proven that the inclusion of morphological features improves the accuracy in differentiation between roads and buildings [8].
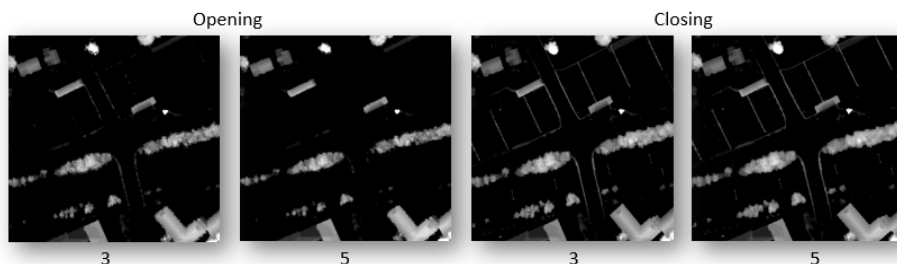
Figure 6. Opening and closing operations on lidar dataset with different kernel sizes (3 and 5) of the structural element.

3.4. Hyperspectral-Lidar Data Fusion

HL-Fusion combines spectral-contextual information obtained by an HS sensor and a lidar scanner's spectral-spatial-geometrical information. Even if the active and passive sensors characterize different physics, their features can be combined from both sensors. Both sensors cover the reflective spectral range intersecting either in the VIS (532 nm) or the SWIR (1064, 1550 nm) wavelength regions. More rarely, multi-spectral lidar systems are used, which overlap in several of the three common wavelengths, allowing the identification of materials or objects using spectral properties [194]. Under laboratory conditions, prototypical hyperspectral lidar systems are being developed [69,195,196]. The combination of HS and lidar sensors significantly impacts remote sensing, opening up possibilities for fully three-dimensional target analysis [196]. Examples include civil engineering, historical preservation, geomorphological studies, and material processing. However, it is not only the classification concerning 3D geometry determined by sensor fusion. Most rely on geometric simplification of high-dimensional data, reducing both lidar data and HS data to 2.5 grids, where geometrically aligned lidar and HS data are classified based on raster data.

HL-Fusion is usually conducted by adjusting the spatial resolution of one sensor to another (HL to lidar), empirically correcting for geometric errors. Such fusion does not consider the different sensor characteristics (e.g., scan, view, or incidence angles). This kind of fusion also fails when the scene has low-contrast areas, as it is very sensitive to illumination, losing information about details important in complex and heterogeneous urban environments. Despite the dimensional degradation, HL-Fusion has great potential for achieving enhanced results in land cover classification rather than using single sensors, especially when combining spectral and spatial features. In the last decade, fusion has been attempted in this way, for example, by adding to the spectral features extracted from HS data, elevation information, intensity, and other lidar-derived features, which allowed one to upgrade the level of the classification from pixel- to object-based analyses.

Spectral-spatial-based classification on fused data often improves the certainty of a pixel's belonging to a class. On the other hand, an increasing number of features extracted for classification purposes from different sensors can lead to a curse of dimensionality, especially when the training data are limited [197]. HL-Fusion can also be performed physically, taking into account sensor parameters, measuring principles, quantities, illumination sources, the position of the sensors, and attitude in the preprocessing phase [198,199]. Intensity values can describe the physical link between the spectral and spatial responses of both sensors' overlapping wavelengths [199]. However, single studies provided HL-Fusion based on fitting spectral data to the first return from lidar data, thus preserving full 3D geometry and structure, improving the scale of analysis and its performance and robustness [200].

4. Classification of Urban Land Cover Classes

Urban land cover classification based on remote sensing data has been carried out on a pixel or object-based classification. Pixel-by-pixel analysis assigns only one of the defined classes to each pixel without considering neighboring pixel decisions [201]. In remote sensing, pixel-based classification relies on the spectral properties of each pixel from the scene. However, pixel-based approaches for high-dimensional remotely sensed HS and lidar data were assumed to be inaccurate for reliable classification purposes [202,203]. Therefore, object-based classification has become relevant, reconstructing reality more truthfully, managing fine spatial resolution data, and suppressing noise. Object-based methods include spatial, textural, contextual, topological, and spectral information [204,205], where objects are defined as classification units [43]. Moreover, the object-based analysis consists of image segmentation, grouping spectrally homogeneous regions, and classification, assigning the segments to the corresponding classes with various properties [206]. Both pixel and object-based classification can be driven in the unsupervised, for

example, deep belief networks (DBN) [207–209], and stacked autoencoder (SA) [41,210–212]) or supervised (RF, SVM) matter.

Analyses on the unsupervised basis separate classification units relying on their common features without providing reference data. This kind of classification is helpful if the knowledge about the study area is limited. In addition, unique classes can be recognized that may have been overlooked applying supervised classification. However, the control over the generated classes is limited, or the final results do not present the analytics intentions, for example, if the desired class is not directly correlated. Supervised classification identifies unknown pixels/objects, validating the accuracy by reference classes assigned to known pixels/objects [213]. One of the advantages of using supervised classification is controlling the number and name of the class labels, which are then assigned to the classification units in the final step [214]. However, supervised classification requires human expertise and the preparation of such reference or ground-truth data adequate for selected area and classification purposes. Such ground-truth sampling includes the removal of outliers and remains representative samples for overall input [215]. This can be accomplished by applying active learning [208,216], random sampling, or stratified random sampling [217].

The ground-truth labeling often requires an equal number of instances assigned to a class. Therefore, a class imbalance issue leads in (multiclass) HS classification to decreased accuracy of many standard algorithms such as decision trees, k-nears neighbor, neural networks, and SVM [218]. Especially for high-dimensional data (HS) and ML/DL-based multiclass problems, the minority classes are neglected or misclassified [219]. Various strategies can be applied to overcome imbalance class issues partially: simplification of the network architecture [38], data augmentation for minority classes, and random sampling for equal class distribution [220].

Complex urban land cover mapping is mainly based on spectral and spatial features of remote sensing data, implemented in classification algorithms. Such an analysis is mainly limited to comparing classification approaches, a general classification scheme, or a small data set, which provides high-accuracy results on local space, excluding generalization and transferability aspects [221]. Often, the evaluation of the classification approach is complicated since the training data may not be representative enough for independent testing data set. In addition, urban land cover analysis usually depends on human expertise at a local scale [84].

Various ML and DL algorithms have been recently explored in feature extraction, pattern recognition, and image classification to deal with high-dimensional space [49,88]. Feature extraction in remote sensing analysis contains mainly shallow supervised and unsupervised and deep feature extraction [222]. In HS data, spectral feature extraction is applied to reduce the high dimensionality and to avoid redundant bands preserving only relevant spectral information. This strategy can also help in increasing separability between different classes. However, spatial feature extraction (texture and morphology) finds the contextual relationship of adjacent pixels improving the only spectral-based classification [132,133,179,209]. In DL, automated extraction of features is common and outperforms shallow ML if the training data fed to the algorithm are not limited.

Aiming to analyze the complexity and improve the DL algorithm learning process quality, a thorough understanding is required of the filter function in the DL architecture [223]. One way to do this is to visualize the parameters of the entire algorithm architecture. However, studies on urban land cover classification based on HS and lidar rarely focus on explaining how the DL algorithms work. As the limited amount of high-dimensional remotely sensed data is fed as input to DL classifiers, there is a probability that the hyperparameter tuning causes overfitting. To avoid this issue, e.g., data augmentation, adding noise, model regularization methods (max-pooling and dropout [224]), and simplifying the model are used. Data augmentation helps diversify training data without new labeling costs, thus leading to more robust classification and adequate classification. In remotely sensed-based classification, training data have been flipped and rotated [225,226],

mirrored across horizontal, vertical, and diagonal axes on HS [226,227] and lidar data [228], mixup strategy [229], and generation of virtual training samples through Generative Adversarial Networks (GANs) [230] on HS data. In addition, noise is proven to be suited as a data augmentation type. Haut et al. [231] added random occlusion data augmentation (rectangular figures of different sizes) in various HS image patches. Many studies applied Gaussian white noise during simulation to improve the robustness of the classification and reduce the model's dependence on local attributes in HL-Fusion [80] and HS data [232].

Apart from overfitting issues, the time-expensive DL algorithms deal with vanishing gradient problems where the learning is unstable and saturates the activations [233]. This problem can be solved by implementing data normalization between each network layer (e.g., local response normalization [234], batch normalization [235], and layer normalization [236]), choosing proper optimizers and nonlinear activation functions [45].

The following section describes the most common ML and DL algorithms for the classification in the urban environment, such as SVM, RF, CNN (convolutional neural network), and RNN (recurrent neural network) (Table 1). Nevertheless, there are many more ML and DL classification algorithms that are not included in this review. Starting with ML algorithms, over time and with technology development, they have become more advanced. Urban analysis with conventional learning-based classifiers was based on interpreting handcrafted low-level features, linear classifiers and nonlinear classifiers, and binary and multiclass classification [88]. Examples are statistical learning on HS data [237], logistic regression on HS data [133], and maximum likelihood classification on lidar data [146]. However, the DL algorithms evolve in classifying urban objects on a larger scale, automatically extracting high-level features. In addition, DL can handle the issue of the complex spatial distribution of spectral information. Automatically derived features in DL rely on a mathematical basis, tuning the model by changing the parameters and neglecting its standard implementations the physical aspect of remote sensing data. In addition to CNNs and RNNs, which have been included in this article, many different DL network frameworks show promising potential for further analysis and a deeper understanding of DL, primarily for HS data. Some of these algorithms are DBN [207–209] with SA [41,210–212] and GAN [35]. However, these algorithms are in the initial phase of implementation and were not applied until 2019 to HL-Fusion data.

Table 1. Overview of classification methods on different urban land cover classes.

Classifier	Input	Domain	Class	Features	Advantages	Limitations	Study
SVM	HS	spectral	building, vegetation, road	spectral	High accuracy among classes with low material variations	Low accuracy among classes with high material variations (synthetic grass, tennis court) or similar material classes (road, highway)	[40]
					Insensitive to noisy data, high accuracy (vegetation, water)	Spectral similarities of materials (misclassification of roofs and other impervious surfaces, impervious and non-vegetated pervious surfaces)	[238]

		vegetation, road		High accuracy among classes with low material variations (metal sheets, vegetation)	Misclassified bricks as gravel and asphalt as bricks	[239]
				Accurate classification with limited training data set		[240]
	spectral-spatial	vegetation, road	spectral, spatial	Adding spatial information improves overall accuracy and generalization	Misclassification of bricks requires knowledge about spatial features (maybe not available in the spectral library)	[241]
				Integration of spatial and spectral features (contextual SVM)		[242]
SVM	Lidar	building, vegetation, road	HD, intensity	Robust and accurate classification	Misclassified small isolated buildings, rounded building edges	[64]
		building, vegetation	full-waveform	Can handle geometric features of 3D point cloud	Not balanced classes lead to misclassification (grass and sand)	[243]
			multiple-return, intensity, morphology, texture	Fusion of single SVM classifiers and textural features improve the final results	Misclassification (building classified as tree class) due to limited training data	[48]
		building, vegetation, road	HD, intensity, spectral	Spectral features performed better than geometrical features in classification based on multispectral lidar	Geometrical features cannot discriminate among low height classes: grass, road	[120]
		building	HD, intensity, texture, spatial	GLCM features (mean and entropy) improve building classification	The magnitude of temporal change of buildings cannot be achieved using SVM, misclassification between trees and buildings	[53]
		building, vegetation, road	HD, intensity, morphology, spectral	Morphological features with nDSM improve road and building classifi-	nDSM provided misclassification between grass and trees	[53]

				cation based on multispectral lidar		
			building, vegetation, road	HD, full-waveform	Dual-wavelength lidar improves land cover classification, especially low and high vegetation, and soil and low vegetation	Very low accuracy of low and high vegetation applying SVM on single wavelength lidar	[67]
SVM	HL-Fusion	spectral-spatial vs. object-based	roof, vegetation, road	HS: spectral Lidar: HD, intensity	The hyperspectral point cloud is robust and provides better results for vegetation and tin roof than grid-based fusion	Accuracy of hyperspectral point cloud classification depends on the proportion between point density of lidar and spatial resolution of HS, very complex in processing (in comparison to grid data)	[200]
		spectral-spatial	vegetation	HS: spectral Lidar: HD	Overall accuracy increased, adding spatial to spectral features	Spatial features introduced misclassification errors in individual tree species	[76]
RF	HS	spectral	vegetation, road	spectral	High classification accuracy of vegetation, good robustness, insensitive to noise	Cascaded RF provides more generalization performance than standard RF	[244]
RF	Lidar		building, vegetation	full-waveform, HD, eigenvalue-based, multi-return	The ability of RF to select important features	Misclassification of grass (natural ground) and roads (artificial ground)	[136]
			building, vegetation, road	HD, intensity, texture	Overall high accuracy, multispectral lidar especially promising for ground-level classes (roads, low vegetation)	Misclassification of gravel and asphalt	[245]
RF	HL-Fusion		building, vegetation, road	HS: spectral Lidar: HD	The ability of RF to select essential features		[18]

Method	Data	Approach	Classes	Format	Advantages	Disadvantages	Ref
			building, road		High overall accuracy with original raw data	Single-class low accuracy (highway, railway), limited training data	[40]
CNN	HS	spectral-spatial	vegetation, road	raw	Very high overall accuracy, insensitive to noise [42,239], CNN in combination with Markov Random Fields improves overall accuracy taking into account complete spectral and spatial information [36], spectral and spatial features extracted simultaneously (full advantage of structural properties) [248]	The model achieved worse overall accuracy on other datasets (Indian pines), computationally expensive, misclassification of bricks and gravel, requires larger data set than standard ML [42,239], time-consuming, limited training data [36]	[36,39, 42,219, 225,229, 230,239, 246–258]
CNN	Lidar	object-based	building	HD	Applicable to large-scale point cloud data sets due to a low number of input features [54] overall high accuracy with applying multiview rasters of roofs [55]	Misclassified buildings as vegetation (especially buildings with complex roof configuration) due to limited and too homogeneous training data, sparse point density [54], height derived features are not sufficient to extract various roof types, require a large training data set [55]	[54,55]
			building, vegetation, road	multi-wavelength intensity, HD	Time-effective due to the simplicity of the model	Trajectory data, strip registration and radiometric correction not included	[259]
		pixel-based		HD	Automatic design of CNN for robust features extraction and high accuracy	Time-expensive search and training	[260]
CNN	HL-Fusion	spectral-spatial		HS: spectral Lidar: HD, spatial	Generalization capability, improved accuracy when	Not efficient in handling high-dimensional data	[16]

		building, vegetation, road		fusing HS and LiDAR	compared to standard ML classifiers	
			HS: spectral, spatial Lidar: HD		Oversmoothing problems in classification results	[29,261, 262]
			HS: spectral, spatial Lidar: HD, spatial	Effective extraction of essential features, reduced noise		[30,263]
	spectral-spatial	vegetation, road		Improved accuracy of fused data, deep neural network used for feature fusion improved the classification results [265]		[80,264]
	pixel-based	building, vegetation, road	HS: spectral Lidar: HD		Remarkable misclassification of objects made from similar materials (parking lots, roads, highway)	[265]
CRNN	HS	spectral-spatial	building, vegetation, road	spectral, spatial	Does not require fixed input length, effectively extracted contextual information	Big training data set required	[266]
			vegetation, road				[39]
RNN	HS	spectral	building, vegetation, road	spectral	Performs better than standard ML algorithms and CNNs	Issues with differentiation of asphalt/concrete made objects (roads, parking lot, highway) requires a longer calculation time	[37]
			vegetation, road				[267,268]
		spectral-spatial	vegetation, road	texture, morphology, spatial	Adding spatial features to the classification improves the overall and class accuracy, high level features can represent complex geometry	Computational time and memory-expensive	[256,269, 270]

4.1. Support Vector Machines (SVM)

SVM is a supervised ML algorithm that performs the classification of locating a hyperplane between two classes [241]. Such a hyperplane separates two groups in the training dataset, finding the largest margin between the support vectors from different groups [271]. The SVM approach is widely used in pattern recognition, regression, and solving linear equations [271]. It has been proven to be a classifier that can handle the high-dimensional HS data being insensitive to noisy samples [272–276]. Moreover, SVM can deal with smaller training datasets more efficiently than artificial neural networks and maximum likelihood classification algorithms [53]. The decision function of the SVM can be specified by different kernels such as radial basis function (RBF), spectral-based [277], and Gaussian

function [19], which classify only in the spectral domain, and composite kernels that include contextual information to the classification [241,278]. The kernel-based methods define the segments by applying the nonlinear geometrical separators [272]. The spectral-based kernel uses the spectral angle of the support vectors to define the hyperplane between them, while for each pixel, spatial information is derived and combined with spectral features in kernel composition. Deep SVM has been implemented with exponential radial basis function, gaussian radial basis function, and neural and polynomial kernel functions, achieving better robustness than conventional classifiers [279].

4.1.1. Buildings

In recent years, a multiwavelength lidar scanner has become an interesting mapping device that can differentiate objects with the same height, such as buildings and trees, based on pseudonormalized difference vegetation index (pseudoNDVI) [68] and geometrical features, e.g., roughness (curvature) [120]. Teo and Wu [15] provided a case study where curvature, intensity, and nDSM were used on multispectral lidar. They applied these lidar features as input for image clustering and found that especially geometric features are suitable for building detection. Huo et al. [8] applied the SVM algorithm with RBF kernel on multispectral lidar data. In the paper, the authors focused, among other things, on building extraction using the combination of nDSM, morphological profiles, novel hierarchical morphological profiles (HMP) [186], pseudoNDVI, and intensity values. Intensity values only extracted from lidar can lead to misclassification of building asphalt roof (parking lot) and a road with similar spectral properties. Shirowzhan and Trinder [53] provided the SVM classification method for building extraction, including DSM, nDSM, and intensity map. However, the results provided a misclassification between roads and buildings in the hilly or vegetation-rich area. A pixel-based classification method is often not able to separate buildings and vegetation boundaries. Samadzadegan et al. [48] proposed a multiclass SVM on building extraction. The authors used first- and last-pulse intensities, first- and last-pulse ranges, entropy, standard deviation, homogeneity, and other geometrical features and showed that texture features improve the final results for building detection. In building analysis based on HS data, the spectral classifie's domain has limitations in the classification of building roofs (roofing tiles, bitumen, concrete, fiber cement, metals, and slates) [97,238]. To overcome the limitations of single sensor applications, HL-Fusion can complete robust building analysis using spatio-spectral-elevation information. Spectral features from HS data can exclude vegetation growing around and on buildings and differentiate between roof materials. By contrast, lidar data provide shape information that can help determine roof types and building types.

4.1.2. Vegetation

SVM classifier is a standard algorithm in vegetation detection in the urban environment. The authors of [48] suggested a multiclass framework for lidar data, analyzing the normalized difference between the first and the last laser pulse. High vegetation class (trees) was falsely classified as buildings due to limited training data. Teo et al. [120] stated that lidar penetration improves the overall accuracy of vegetation analysis. However, by splitting vegetation into high and low vegetation, lidar data cannot distinguish low height classes such as roads and grass. Huo et al. [8] applied SVM on multispectral lidar data calculating the NDVI and pseudoNDVI [178] and improving the overall classification accuracy, however, having challenges in distinguishing between low and high vegetation. Wang et al. [67] addressed a similar problem in the study and compared single- and dual-wavelength lidar by applying, among others, full-waveform data that were not included in previous studies. The authors showed that dual-wavelength improves the accuracy of low and high vegetation and bare soil and low vegetation compared to single-wavelength lidar. In HS analysis, spectral features are still more accurate in chlorophyll detection than lidar, mainly when a class is characterized by low material variations [40,239]. In addition,

HS has been proven to characterize fraction coverage of photosynthetic vegetation, non-photosynthetic vegetation, and soil [72]. Furthermore, by adding spatial features to the hyperspectral analysis, vegetation detection becomes facilitated [241,242]. Spatial information is also used in HL-Fusion in object-based classification, being able to classify different types of vegetation (tree species) [32,76] and also, in the case of generating hyperspectral point clouds, maintain higher reality factors such as full 3D geometry, generic and robust characteristics [200].

4.1.3. Roads

Huo et al. [8] and Teo et al. [120] applied SVM on multispectral lidar data to detect roads. Achieving high accuracy classification, Huo et al. [8] referred to the misclassification of roads as lawn and bare soil, which can be easily solved by adding HS to the lidar data due to access to more detailed spectral information than lidar only. One of the causes can be similar spectral signatures and insufficient distinctive spectral-spatial features to differentiate between objects. Teo et al. [120] mentioned classification issues applying geometrical features among grass, road, and soil due to similar height. However, spectral features from multiwavelength lidar can overcome the challenge.

In contrast, spectral features in HS analysis applying SVM are often insufficient for achieving robust and accurate results of road classification [40,239]. This is due to considering only spectral information without contextual information and remarkable spectral similarity between physical material belonging to different classes. SVM has also been widely used in road classification on fused HS and lidar data. Brell et al. [200] generated an HS point cloud, where they classified different road materials such as concrete and asphalt. The challenge in distinguishing concrete and asphalt is the influence of shadow deteriorating discrimination between different road materials. The spectral properties of those materials can vary locally based on aging, deterioration, contamination, roughness properties, and other conditions [200].

4.2. Random Forest (RF)

RF is a nonparametric ensemble learning algorithm based on a combination of binary decision tree classifiers [280]. A decision tree in the ensemble is independent of other trees and is trained with random variables by bootstrap sampling [77]. For classification purposes, each tree gives a class prediction as an output. The class that most trees have chosen is considered to be the final result [281]. RF has become a widely used classification algorithm in HS imaging due to its high accuracy and high processing speed [282]. Moreover, RF can handle high-dimensional data selecting redundant spectral bands without overfitting [18,77]. RF has also been applied to airborne-based lidar data as a classifier solving multiclass problems and selecting the essential features for urban mapping [136].

4.2.1. Buildings

Niemeyer et al. [283] proposed a new building classification method based on the 3D point cloud from lidar data. The classification technique transforms the RF classifier into a conditional random field (CRF) framework [218] and provides high-accuracy results for large buildings over 50 m². However, misclassification occurs at building facades and dormers. In addition, various features derived by lidar have been tested by Chehata et al. [136]. In the paper, multiecho, full-waveform, different height-based, local plane-based, and eigenvalue-based features have been applied to classify buildings. However, confusion errors occurred for transition points between buildings and the ground class.

Further, echo-based features did not have any influence on classification results. Debes et al. [18] presented a fusion framework consisting of unsupervised classification that supports the supervised classification on ensemble learning. They showed that lidar elevation information is required to differentiate between buildings and vegetation or different building types in addition to HS spectral data.

4.2.2. Vegetation

Niemeyer et al. [283] applied an RF classification framework with conditional random fields on lidar data to discriminate vegetation and buildings from each other. Chehata et al. [136] applied RF on lidar data experiencing issues in the classification between vegetation and artificial (roads) and natural ground (grass), respectively. However, applying intensity, height, and texture features of multispectral lidar is very promising for ground-level classes, for example, low vegetation [245]. In HS analysis, spectral features fed to RF classifier provide high vegetation accuracy, good robustness, and insensitivity to noise [244]. Debes et al. [18] chose an RF algorithm on HL-Fusion with elevation features from lidar and NDVI from HS data that outperformed urban area classification [18].

4.2.3. Roads

Niemeyer et al. [283] proposed an RF classification framework for lidar data described in Section 3.1, where one of the classes was asphalt considered a road. However, other objects apart from roads are also made of asphalt, such as roof parking lots, making the analysis difficult, e.g., using only HS data. Jackson et al. [284] mentioned this issue clarifying that the road class pixels are contaminated by other materials and objects such as gravel, puddles, and cars. In addition, vehicles appearing in the image usually have highly reflective properties, making road classification difficult for RF classifiers [8]. Recently, lidar point cloud intensity data have been proposed for road landmark inventory with active learning [285].

4.3. Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN)

CNN is a DL algorithm that has become an important HS, lidar, and HL-Fusion classification method. The network's deep convolutional architecture can effectively deal with complex remote sensing data solving nonlinear issues [286] with an example architecture in Figure 7.
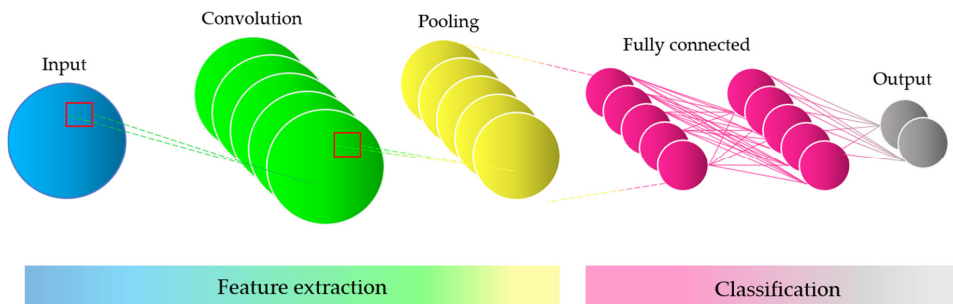
Figure 7. Convolutional neural network architecture.

CNN has two characteristics different from other DL algorithms, such as local connections and share weights. Local connections help find the data's spatial relationship, and share weights reduce the number of parameters needed for training purposes and generate a filter [16,286]. CNN architectures can be trained in an unsupervised or supervised way. The unsupervised method is the greedy layer-wise pretraining of hyperspectral data [287–290]. The supervised method is the standard backpropagation [234,286,291,292]. However, CNNs require a high number of model parameters. The high dimensionality and limited training samples of the remotely sensed data can lead to overfitting and longer processing time than other classification techniques [37,39,293]. An advantage of applying CNN is that the input data must not be preprocessed. CNN is capable of automatically learning abstract features and detecting high-level objects [54]. CNN is

used for land cover mapping based on HS data [35,37–39] and HL-Fusion [16,30,265,293]. Moreover, CNN has been recently used in combination with other algorithms such as MRF [37] to extract HS pixel vectors on a spatial and spectral basis and extinction profiles [30], capable of effectively reducing noise and improving classification accuracy.

RNN is a DL algorithm widely used to work on sequential data [37]. RNN is a compound of successive recurrent layers capable of extracting contextual parameters at consecutive time steps (Figure 8) [266]. An advantage of RNN is that the input sequences may have different lengths [266]. However, RNN requires a longer processing time than standard ML algorithms, such as SVM or RF [37]. RNN has already been used to extract contextual information of HS data [37,266] and recognize temporal changes of objects measured by the lidar scanner [294].

Figure 8. Recurrent neural network architecture.

4.3.1. Buildings

Zhou and Gong [54] focused on building detection in different conditions in damage assessment. Their approach relies on roof object extraction, challenging for lidar data due to sparse points in the boundaries between rooftops and the vertical facades of buildings damaged. In addition, very complex roof configurations cannot be distinguished using lidar data when the training data are too homogeneous. However, DL algorithms such as CNN provided accurate classifica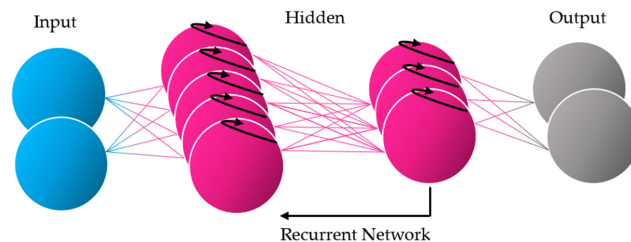tion results of pre- and post-disaster data with minimal required preprocessing of lidar data and time consumption. Shahajan et al. [55] provided a DL approach that extracts the lidar points from different views assigned to roofs applying height-derived features. However, like the previous study, height-derived features are insufficient in roof type differentiation, and the CNN algorithm requires a large training data set. CNN classifier has also been used in HS data analysis due to its relevant spectral-spatial domain [40]. However, CNN on HS data is time-consuming even if the preprocessing of the fed input is minimized, requires a more extensive data set than shallow ML classifiers, and is not transferable with the same model parameters to other independent test data. Li et al. [30] proposed a DL framework based on spatial and elevation features extracted from extinction profiles, spectral, spatial, and elevation features extracted from the CNN model to classify buildings, among others, on HL-Fusion. Extinction profiles were also used to derive spectral, spatial, and elevation features from HS and lidar data. These features were applied as input for CNN classification on buildings [16]. Morchhale et al. [265] have proven that CNN-based classification on HL-Fusion can distinguish between commercial buildings and highways and between residential buildings and parking lots, improving generalization capability.

Wu et al. [266] introduced deep RNN for HS data classification combining with CNN and creating a convolutional recurrent neural network (CRNN). This framework enabled the extraction of hidden feature representations and provided highly accurate results for building detection. For HS image classification, Mou et al. [37] provided an RNN framework with a GRU activation function that maintains a constant error, helping the network learn more effectively in a high-dimensional space. As a result, his classifier achieved very high accuracy in recognizing commercial and residential areas in the urban environment.

Even though only spectral features without contextual information were considered, RNN outperformed standard ML algorithms and CNNs.

4.3.2. Vegetation

CNN is used in HS analysis for vegetation detection [36,39,42,225,239,246–248]. Li et al. [248] simultaneously extracted spectro-spatial features of HS data benefitting structural properties needed for detailed vegetation interpretation. However, more and more algorithms for vegetation classification are based on HL-Fusion data. Ghamisi et al., Morchhale et al., and Li et al. [16,30,265] proposed different frameworks based on CNNs. Chen et al. [264] created a CNN framework used to extract the spectral-spatial features of HS data and the elevation features of lidar data. He applied a fully connected DNN to fuse the derived features from both sensors, ending the classification approach with the logistic regression to generate the final classification map [264]. Deep RNN introduced by Mou et al. [37] has been used for vegetation classification. Although RNN resulted in high overall accuracy, the most significant challenges occurred in classifying different grass class types, such as healthy grass, stressed grass, and synthetic grass [37].

4.3.3. Roads

CNN algorithms have already been widely applied as an initial framework for road classification as objects or materials, e.g., gravel, concrete, and asphalt. Santara et al. [38] compared different ML and DL algorithms, including the CNN framework. CNN classified roads as asphalt and gravel with high accuracy only on HS data. Recently, much more often, CNN is used as a classifier for HL-Fusion. Morchhale et al. [265] compared CNN on HS data and HL-Fusion. The classification and differentiation accuracy between road, parking lot, and highway increased in the HL-Fusion. Li et al. [28] proposed that he focuses on classification challenges between similar spectral characteristics of road materials, e.g., asphalt and concrete, and the similar height of different objects, such as grass and asphalt road. Ghamisi et al. [25] applied the CNN classifier with logistic regression and mentioned the challenge of similar spectral signatures of roofs and roads for HS data classification. Wu et al. [266] and Yang et al. [39] proposed CRNNs for HS image classification. Mou et al. [37] presented a different framework—the deep RNN. In both network frameworks, the road was grouped into road and highway. The deep RNN outperformed other conventional classifiers in differentiating similar objects, such as road, highway, and railway [37].

5. Discussion

Airborne HS and lidar data-based classification in the urban environment over the last 20 years has increased significantly since 2016, as shown in the annual number of articles reviewed in this paper found up to 2021 (Figure 9). Therefore, it can be assumed that the interest in HS and lidar remote sensing, advances in sensor technology, computing power, and easy access to remote sensing-based datasets are relevant factors paving the way for large-scale urban environment analysis. However, it has to be noted that the HS-based land cover classification far exceeds lidar and HL-Fusion analyses. Since 2016, the scientific production of urban classification methods based on ML and DL has significantly increased for HS, lidar, and HL-Fusion. This is due to the availability of more advanced computer infrastructures, less expensive sensors with higher resolution, and more accessible data for HS and lidar.

Nevertheless, the HS continued to be widely used. Firstly, this may have been due to the lack of data of the same study area from the two sensors acquired simultaneously. Secondly, most land cover classification approaches are based on physical material classification, which relies significantly on spectral analysis. Sometimes, therefore, it is not necessary to fuse two sensors for some purposes to improve classification by a small fraction with much more effort and time spent fusing the sensors. However, assuming that urban

analysis is a highly complex task, one of the fusion application arguments may be that HS and lidar complement each other in spectral and spatial analysis with the addition of elevation information and active and passive sensor characteristics.
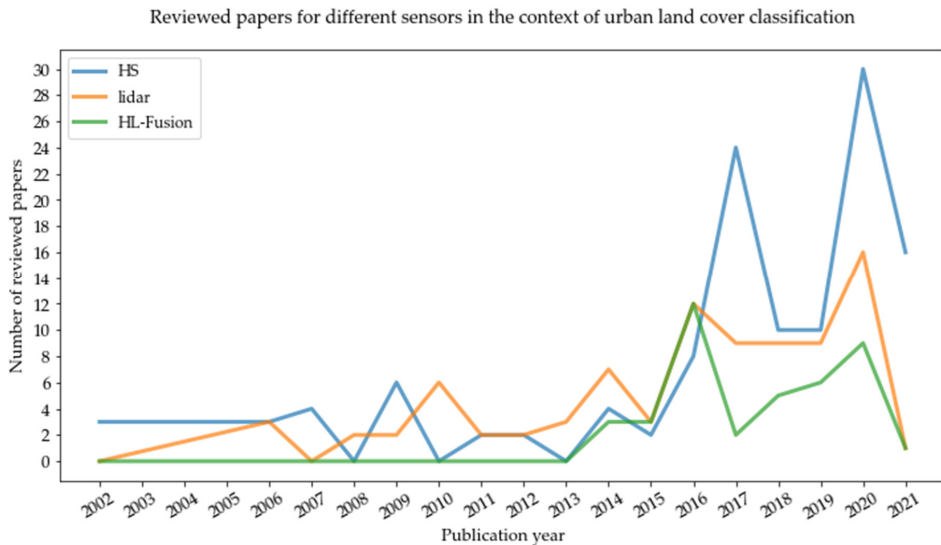
Figure 9. Each color is assigned to a different sensor, such as HS, lidar, and HL-Fusion.

5.1. HS-Based Classification

SVM and RF have been proven to be insensitive to noise in HS-based analysis, providing very high accuracy for material classification with limited training data. The only spectral-based SVM classifier can quickly identify object-based classification classes with low material variations. However, for complex urban land cover classification, the spectral domain of the SVM is not sufficient to capture the heterogeneity of the objects or land cover classes built from various materials, for example, identification of impervious and nonvegetated pervious surfaces [238]. For such an analysis, contextual information is necessary. The spatial features can be added to the SVM classifier by applying the composite kernel, improving the accuracy and generalization capabilities.

RF also applies spectral-domain only for HS data. As for SVM, the land cover classes having high material variations within a class are often misclassified, such as road materials (concrete, asphalt, and gravel). The difference and advantage over the SVM classifier is the capability to select important features. This aspect is also advantageous for the DL-based classifiers since shallow ML-based algorithms use handcrafted features controlled and transferred to other independent and unknown test areas.

On the other hand, when the classification objective is focused on a smaller study area, automated features of the DL algorithm may prove to be a better solution for high-dimensional HS data. One factor is that the relationships between objects or land cover classes are not linear in a complex urban environment. By extracting the handcrafted features, we have control and knowledge about them. In contrast, the automated features can obtain high-level features that may allow a much better classification result by recognizing complex relationships that cannot be analyzed by applying shallow ML at the expense of generalizability and transferability. The advantage of the CNN is in its spectral-spatial domain, which searches for high-level features, e.g., by simultaneously extracting spectral and spatial features. As features in CNN are retrieved during the algorithm and the the

original, however, normalized, HS data are fed into the model. This saves time for pre-processing, which is necessary for SVM or RF classifiers.

On the other hand, normalization of the extraction of high-level features is notably more time-consuming than classification with shallow ML algorithms. In totally, the increase in dimensions is an enormous challenge in DL classification. The hidden complex relationships are not universally and globally representative. The easiest way to influence this in DL is to ensure the global representativeness of the training and test data which may be hardly possible in remote sensing. Another way is the support of handcrafted features that underrepresent the local properties. In addition, assuming that the algorithm extracts the most important features for correct classification, these features may vary depending on the complexity and diversity of the training data. Therefore, the transferability and generalizability of a model, which is critical, e.g., automatic map updating, is limited. In addition, DL algorithms require a larger training data set, which may not be feasible due to a lack of data and computationally expensive DL algorithms, such as CNN and RNN. The RNN has been proven to outperform even in the spectral domain only compared to standard ML algorithms and CNNs [37]. However, for single classes such as asphalt or concrete-made objects (roads, parking lots, highway), RNN may not solve misclassification only in the spectral domain. The RNN requires more computative time than the CNN, and an extensive training data set is needed. Since RNN considers the temporal domain, this classifier shows greater robustness, transferability, and generalization.

5.2. Lidar-Based Classification

Lidar-based urban land cover classification is not a straightforward approach due to the complexity of the urban environment, where different classifiers with different derived features can identify different land cover classes. Nevertheless, the SVM is a common approach for lidar-based classification. In particular, those land cover classes are distinguished by their unique geometry and where the material composition is not essential. For example, building detection requires the capture of complex geometry, including roofs (planar surfaces) and facades (vertical surfaces). For this purpose, full-waveform data and geometric features are commonly used [243]. However, depending on the specific purpose of the building classification, different features may play an important role. For example, in the analysis of various roof types, the focus was on height-derived features that were insufficient when the roof had a very complex geometry [55]. However, the problem may lie in too low resolution, too sparse point density of the lidar system, or the CNN classifier, which needs a much larger training data set considering heterogeneity and complexity of the objects of interest. In addition, the transition from 3D point cloud to 2.5D representation is challenging to preserve inherent point cloud information.

Raster (2.5D) processing is more efficient in data handling as soon as it comes to spectral-spatial neighborhood analyses and is therefore preferred by most classifiers. The SVM is mainly used for building detection, differentiate between low and high vegetation, and distinguish trees and buildings. The differentiation between low and high vegetation is still a problem. It appears that using height-derived features and full-waveform data from single-wavelength lidar is not sufficient. However, using the same features with a dual-wavelength lidar scanner significantly improved low and high vegetation classification results. Therefore, it can be concluded that spectral features play a significant role in the detailed classification of land cover classes [67]. This assumption of the importance of spectral features in lidar-based classification applying the SVM was also mentioned in a study where spectral features were more critical than geometrical features in classification on a multispectral lidar scanner [120]. In this study, it was found that geometrical features are not able to detect ground-level classes such as roads and grass, which on the other hand, is possible using the spectral features of the multispectral lidar.

Ground-level classes cause many problems also when using full-waveform data. For example, in one study, incorrect classification of grass and sand was caused because the training data contained no balanced classes in the SVM classification [243]. On the other

hand, a similar problem of incorrect RF-based classification (this time of low vegetation and roads) also appeared when applying full-waveform data [136]. Therefore, it can be concluded that the last return from lidar is not sufficient for differentiating between ground-level classes. However, this problem has been solved by applying the RF classifier using multispectral lidar data but adding texture features to the elevation and spectral information [245]. However, neither SVM nor RF on single or multispectral lidar can differentiate very heterogeneous classes, such as asphaltic, concrete, or gravel road. This issue can only be solved by adding hyperspectral information on the material using the little available hyperspectral lidar, including reflected intensity information, or integrating lidar with hyperspectral imaging.

5.3. HL-Fusion Classification

HL-Fusion aims to combine the two different sensors with improving the classification result. In urban land cover classification based on HL-Fusion, DL turns out to be the most commonly used method (Table 1). One reason for the DL selection could be the intentional neglect of the more complex preprocessing. Thereby, however, there is a risk of losing transferability and generalizability. This is especially critical for optical data, e.g., if the atmosphere is not corrected according to physical models or the shadow has not been corrected, training data must cover all atmospheric conditions and represent the existing urban heterogeneities. However, both the enormous, rapid development of DL, combined with the progress of sensor technology and multisensory fusion, are becoming an interesting field for further scientific research in the near future. Especially in the analysis of complex urban environments, only a single sensor is usually insufficient for classifying urban land cover classes correctly. Besides material characteristics, spatial correlation is essential and full 3D geometry and topography information. Using context in a more spacious neighborhood for classification purposes, training time increases significantly, especially for DL algorithms (CNN). For shallow ML and DL algorithms, spectral-spatial classification with handcrafted features has been proven to always be more accurate, with the capability of transferability and generalization [16,18,29,265]. HL-Fusion with SVM classifier improved the classification result, but the limited studies did not include the variation between different features derived from HS and lidar. In this case, the application of spectral features from HS and height and derivatives and intensity data proved accurate. Unfortunately, fusing two different sensors also come with some challenges. Adding to the already high-dimensional HS data more dimensions, one can meet the curse of dimensionality problems. By limited training data, high-dimensional feature spaces are often insufficient to recognize desired patterns due to the low ratio of training data to the high dimensional features [295]. More dimensions in source data mean more necessary training and test data due to increased heterogeneity and the number of features, and the need for more computational power and storage.

Although object-based classification is much more comprehensive than pixel-based classification, objects or land cover have become important in classification because they reflect reality much more closely. Spectral features from HS data are reliable for material classification, even in complex urban environments. However, the lack of topographic and geometric information makes accurate results based on only one sensor impossible. Lidar for providing these needed features is very promising in complex urban land cover classification. Lidar complements HS data to add height information to vegetation detection enabling identification of individual trees (full 3D geometry), bushes, low vegetation. In addition, in the detection of the road (edges), lidar provides refined features providing precise boundaries [296]. Thanks to HL-Fusion, there is no need to limit oneself to classify land cover classes, monitor the urbanization processes, and study the urban environment. The potential capabilities of the two sensors enable urban analysis in a holistic, multi-aspect and multidisciplinary way.

6. Conclusions and Future Perspectives

ML and DL revolutionize digital processing of remotely sensed data such as HS, lidar, and HL-Fusion. A significant factor influencing such a great advance in technology is the variety of information obtained time-efficiently by remote sensing systems. Both HS and lidar-based data are used for urban analysis by applying ML and DL algorithms. This review provides the latest information on advances in mapping techniques based on HS and lidar data in urban environments based on the reflective spectral range (400–2500 nm). This multidisciplinary research described in this article was intended to summarize urban land cover classification for ML and DL experts and remote sensing specialists. Particular attention should be paid to DL implementations in HL-Fusion, which may be the key to classifying land cover classes in a complex urban environment. DL is a promising tool for extracting spectral-spatial features and more complex features than classical ML algorithms, which usually improves the accuracy of the classification results. One of the main challenges related to DL's use is the need for a globally representative dataset for the model training purposes and the availability of annotated lidar data to make it generalizable and transferable: this might require extensive manual work that can be costly but may be overcome applying data augmentation strategy. The HL-Fusion-based classification opens up a new dimension of urban analysis, approximating ML and DL classification results to the reality and going beyond human expertise to discover and care for the urban environment.

The growing trend of using DL in classification will probably remain unchanged over the next few years, discovering new network algorithms, which are already implemented in single case studies. However, as the technology continues to improve, HL-Fusion, despite its high dimensionality, should be considered in analyzing complex urban environments. Crucial is the transferability and generalization aspect, one of the biggest concerns since DLs are usually valid only locally. Inferring from this, it does not allow, for example, the significant updates of city maps.

Abbreviations

Abbreviation	Explanation
CHM	Canopy Height Model
CRF	Conditional Random Field
CNN	Convolutional Neural Network
CRNN	Convolutional Recurrent Neural Network
DBN	Deep Belief Networks
DL	Deep Learning
DSM	Digital Surface Model
DTM	Digital Terrain Model
GAN	Generative Adversarial Network
GLCM	Gray-Level Co-Occurrence Matrix
HD	Height features and their Derivatives

HS	Hyperspectral
HL-Fusion	Hyperspectral-Lidar fusion
IFOW	Instantaneous Field of View
Lidar	Light Detection and Ranging
LDA	Linear Discriminant Analysis
LBP	Local Binary Patterns
ML	Machine Learning
MCR	Multivariate Curve Resolution
NDI	Normalized Difference Index
NDVI	Normalized Difference Vegetation Index
nDSM	normalized Digital Surface Model
PCA	Principal Component Analysis
psuedoNDVI	Pseudo Normalized Difference Vegetation Index
RBF	Radial Basis Function
RF	Random Forest
RNN	Recurrent Neural Network
SAR	Synthetic Aperture Radar
SWIR	Shortwave-Infrared
SNR	Signal to Noise Ratio
SA	Stacked Autoencoder
SVM	Support Vector Machines
VNIR	Visible and Near-Infrared
VIS	Visible light

References

1. United Nations. *2018 Year in Review*; United Nations: New York, NY, USA, 2018.
2. Chen, F.; Kusaka, H.; Bornstein, R.; Ching, J.; Grimmond, C.S.B.; Grossman-Clarke, S.; Loridan, T.; Manning, K.W.; Martilli, A.; Miao, S. The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems. *Int. J. Climatol.* **2011**, *31*, 273–288.
3. Lee, J.H.; Woong, K.B. Characterization of urban stormwater runoff. *Water Res.* **2000**, *34*, 1773–1780.
4. Forster, B.C. Coefficient of variation as a measure of urban spatial attributes, using SPOT HRV and Landsat TM data. *Int. J. Remote Sens.* **1993**, *14*, 2403–2409.
5. Sadler, G.J.; Barnsley, M.J.; Barr, S.L. Information extraction from remotely-sensed images for urban land analysis. In Proceedings of the 2nd European GIS conference (EGIS'91), Brussels, Belgium, 2–5 April 1991; pp. 955–964.
6. Carlson, T. Applications of remote sensing to urban problems. *Remote Sens. Environ.* **2003**, *86*, 273–274.
7. Coutts, A.M.; Harris, R.J.; Phan, T.; Livesley, S.J.; Williams, N.S.G.; Tapper, N.J. Thermal infrared remote sensing of urban heat: Hotspots, vegetation, and an assessment of techniques for use in urban planning. *Remote Sens. Environ.* **2016**, *186*, 637–651.
8. Huo, L.Z.; Silva, C.A.; Klauberg, C.; Mohan, M.; Zhao, L.J.; Tang, P.; Hudak, A.T. Supervised spatial classification of multispectral LiDAR data in urban areas. *PLoS ONE* **2018**, *13*, doi:10.1371/journal.pone.0206185
9. Jürgens, C. Urban and suburban growth assessment with remote sensing. In Proceedings of the OICC 7th International Seminar on GIS Applications in Planning and Sustainable Development, Cairo, Egypt, 13-15 Feburary, 2001; pp. 13–15.
10. Hepinstall, J.A.; Alberti, M.; Marzluff, J.M. Predicting land cover change and avian community responses in rapidly urbanizing environments. *Landsc. Ecol.* **2008**, *23*, 1257–1276.
11. Batty, M.; Longley, P. *Fractal Cities: A Geometry of Form and Function*; Academic Press: London, UK; San Diego, CA, USA, 1994.
12. Ben-Dor, E.; Levin, N.; Saaroni, H. A spectral based recognition of the urban environment using the visible and near-infrared spectral region (0.4-1.1 μm). A case study over Tel-Aviv, Israel. *Int. J. Remote Sens.* **2001**, *22*, 2193–2218.
13. Herold, M.; Gardner, M.E.; Roberts, D.A. Spectral resolution requirements for mapping urban areas. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1907–1919.
14. Brenner, A.R.; Roessing, L. Radar Imaging of Urban Areas by Means of Very High-Resolution SAR and Interferometric SAR. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 2971–2982.
15. Soergel, U. Review of Radar Remote Sensing on Urban Areas. In *Radar Remote Sensing of Urban Areas*, Soergel, U., Ed.; Springer: Berlin, Germany, 2010; pp. 1–47.
16. Ghamisi, P.; Höfle, B.; Zhu, X.X. Hyperspectral and LiDAR data fusion using extinction profiles and deep convolutional neural network *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.***2016**, *10*.

17. Benz, U.C.; Hofmann, P.; Willhauck, G.; Lingenfelder, I.; Heynen, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS J. Photogramm. Remote Sens.* **2004**, *58*, 239–258.

18. Debes, C.; Merentitis, A.; Heremans, R.; Hahn, J.; Frangiadakis, N.; Kasteren, T.v.; Liao, W.; Bellens, R.; Pizurica, A.; Gautama, S.; et al. Hyperspectral and LiDAR data fusion: outcome of the 2013 GRSS data fusion contest. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 550

19. Dalponte, M.; Bruzzone, L.; Gianelle, D. Fusion of hyperspectral and LiDAR remote sensing data for classification of complex forest areas. *IEEE Trans. Geosci. Remote Sens.* **2008**. Available online: https://rslab.disi.unitn.it/papers/R59-TGARS-Dalponte.pdf (Accessed on 2 May 2021)

20. Sohn, H.-G.; Yun, K.-H.; Kim, G.-H.; Park, H.S. Correction of building height effect using LIDAR and GPS. In Proceedings of the International Conference on High Performance Computing and Communications, Sorrento, Italy, 21-23 September, 2005; pp. 1087–1095.

21. Guislain, M.; Digne, J.; Chaine, R.; Kudelski, D.; Lefebvre-Albaret, P. Detecting and correcting shadows in urban point clouds and image collections. In Proceedings of the Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 537–545.

22. George, G.E. *Cloud Shadow Detection and Removal from Aerial Photo Mosaics Using Light Detection and Ranging (LIDAR) Reflectance Images*; The University of Southern Mississippi: Hattiesburg, MS, USA, 2011.

23. Brell, M.; Segl, K.; Guanter, L.; Bookhagen, B. Hyperspectral and Lidar Intensity Data Fusion: A Framework for the Rigorous Correction of Illumination, Anisotropic Effects, and Cross Calibration. *IEEE Trans. Geosci. Remote Sens.* **2017**. Available online: https://www.researchgate.net/publication/313687025_Hyperspectral_and_Lidar_Intensity_Data_Fusion_A_Framework_for_the_Rigorous_Correction_of_Illumination_Anisotropic_Effects_and_Cross_Calibration (Accessed on 2 May 2021)

24. Hui, L.; Di, L.; Xianfeng, H.; Deren, L. Laser intensity used in classification of LiDAR point cloud data. In Proceedings of the International Symposium on Geoscience and Remote Sensing, Boston, MA, USA, 8–11 July, 2008.

25. Liu, W.; Yamazaki, F. Object-based shadow extraction and correction of high-resolution optical satellite images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 1296–1302.

26. Zhou, W.; Huang, G.; Troy, A.; Cadenasso, M.L. Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: a comparison study. *Remote Sens. Environ.* **2009**, *113*, 1769–1777.

27. Priem, F.; Canters, F. Synergistic use of LiDAR and APEX hyperspectral data for high-resolution urban land cover mapping. *Remote Sens.* **2016**, *8*, 787.

28. Luo, R.; Liao, W.; Zhang, H.; Zgang, L.; Scheunders, P.; Pi, Y.; Philips, W. Fusion of Hyperspectral and LiDAR Data for Classification of Cloud-Shadow Mixed Remote Sensed Scene *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**. Available online: https://telin.ugent.be/~wliao/Luo_JSTARS2017.pdf (Accessed on 2 May 2021)

29. Chen, Y.; Li, C.; Ghamisi, P.; Shi, C.; Gu, Y. Deep fusion of hyperspectral and LiDAR data for thematic classification In Proceedings of the International Geoscience and Remote Sensing Symposium, Beijing, China, 10–14 July, 2016.

30. Li, H.; Ghamisi, P.; Soergel, U.; Zhu, X.X. Hyperspectral and LiDAR fusion using deep three-stream convolutional neural networks. *Remote Sens.* **2018**, *10*, 787.

31. Yan, W.Y.; El-Ashmawy, N.; Shaker, A. Urban land cover classification using airborne LiDAR data: a review. *Remote Sens. Environ.* **2015**.

32. Alonzo, M.; Bookhagen, B.; Roberts, D.A. Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sens. Environ.* **2014**, *148*, 70–83.

33. Kokkas, N.; Dowman, I. Fusion of airborne optical and LiDAR data for automated building reconstruction. In Proceedings of the ASPRS Annual Conference, Reno, Nevada, 1–5 May, 2006.

34. Torabzadeh, H.; Morsdorf, F.; Schaepman, M.E. Fusion of imaging spectroscopy and airborne laser scanning data for characterization of forest ecosystems. *ISPRS J. Photogramm. Remote Sens.* **2014**, *97*, 25–35.

35. Zhu, L.; Chen, Y.; Ghamisi, P.; Benediktsson, J.A. Generative adversarial networks for hyperspectral image classification *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 5046–5063.

36. Cao, X.; Zhou, F.; Xu, L.; Meng, D.; Xu, Z.; Paisley, J. Hyperspectral image classification with markov random fields and a convolutional neural network. *IEEE Trans. Image Process.* **2017**, *27*, 2354–2367.

37. Mou, L.; Ghamisi, P.; Zhu, X.X. Deep recurrent neural networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**. Available online: https://www.semanticscholar.org/paper/Deep-Recurrent-Neural-Networks-for-Hyperspectral-Mou-Ghamisi/5a391667242b4a631acdd5917681b16a86523987 (Accessed on 4 May 2021)

38. Santara, A.; Mani, K.; Hatwar, P.; Singh, A.; Garg, A.; Padia, K.; Mitra, P. BASS Net: band-adaptive spectral-spatial feature learning neural network for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2016**. Available online: https://arxiv.org/pdf/1612.00144.pdf (Accessed on 8 May 2021)

39. Yang, X.; Li, X.; Lau, R.Y.K.; Zhang, X. Hyperspectral image classification with deep learning models. *IEEE Trans. Geosci. Remote Sens.* **2018**, *99*, 1–16.

40. Li, S.; Song, W.; Fang, L.; Chen, Y.; Ghamisi, P.; Benediktsson, J.A. Deep learning for hyperspectral image classification: an overview. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6690–6709.

41. Chen, Y.; Lin, Z.; Zhao, X.; Wang, G.; Gu, Y. Deep learning-based classification of hyperspectral data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2094-2107

42. Hu, W.; Huang, Y.; Wei, L.; Zhang, F.; Li, H. Deep convolutional neural networks for hyperspectral image classification. *J. Sens.* **2015**.

43. Yu, Q.; Gong, P.; Clinton, N.; Biging, G.; Kelly, M.; Schirokauer, D. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 799–811.

44. Zhou, W.; Troy, A.; Grove, J.M. Object-based land cover classification and change analysis in the Baltimore metropolian area using multi-temporal high resolution remote sensing data. *Sensors* **2008**, *8*, 1613–1636.

45. Paoletti, M.E.; Haut, J.M.; Plaza, J.; Plaza, A. Deep learning classifiers for hyperspectral imaging: a review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *158*, 279–317.

46. Jiménez, L.O.; Rivera-Medina, J.L.; Rodríguez-Díaz, E.; Arzuaga-Cruz, E.; Ramírez-Vélez, M. Integration of spatial and spectral information by means of unsupervised extraction and classification for homogenous objects applied to multispectral and hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 844–851.

47. Niemeyer, J.; Wegner, J.; Mallet, C.; Rottensteiner, F.; Soergel, U. Conditional random fields for urban scene classification with full waveform LiDAR data. In *Photogrammetric Image Analysis*, Stilla, U., Rottensteiner, F., Mayer, H., Jutzi, B., Butenuth, M., Eds.; Springer Berlin/Heidelberg, Germany, 2011; Volume 6952, pp. 233–244.

48. Samadzadegan, F.; Bigdeli, B.; Ramzi, P. A multiple classifier system for classification of LiDAR remote sensing data using multi-class SVM. In *Multiple Classifier Systems*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 254–263.

49. Giampouras, P.; Charou, E. Artificial neural network approach for land cover classification of fused hyperspectral and LiDAR data. In Proceedings of the Artificial Intelligence Applications and Innovations, Paphos, Cyprus, 30 September–2 October, 2013; pp. 255–261.

50. Medina, M.A. Effects of shingle absorptivity, radiant barrier emissivity, attic ventilation flowrate, and roof slope on the performance of radiant barriers. *Int. J. Energy Res.* **2000**, *24*, 665–678.

51. Ridd, M.K. Exploring a V-I-S- (vegetation—impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. *Int. J. Remote Sens.* **1995**, *16*, 2165–2185.

52. Haala, N.; Brenner, C. Extraction of buildings and trees in urban environments. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 130–137.

53. Shirowzhan, S.; Trinder, J. Building classification from LiDAR data for spatial-temporal assessment of 3D urban developments. *Procedia Eng.* **2017**, *180*, 1453–1461.

54. Zhou, Z.; Gong, J. Automated residential building detection from airborne LiDAR data with deep neural networks. *Adv. Eng. Inform.* **2018**, *36*, 229–241.

55. Shajahan, D.A.; Nayel, V.; Muthuganapathy, R. Roof classification from 3-D LiDAR point clouds using multiview CNN with self-attention. *IEEE Geosci. Remote Sens. Lett.* **2019**, *99*, 1–5.

56. Matikainen, L.; Hyyppa, J.; Hyyppa, H. Automatic detection of buildings from laser scanner data for map updating. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Dresden, Germany, 8–10 October, 2003.

57. Hug, C.; Wehr, A. Detecting and identifying topographic objects in imaging laser altimetry data. In Proceedings of the International Archives of the Photogrammetry and Remote Sensing, Stuttgart, Germany, 1997; pp. 16–29. (Accessed on 8 May 2021)

58. Maas, H.G. The potential of height texture measures for the segmentation of airborne laserscanner data. In Proceedings of the 4th International Airborne Remote Sensing Conference and Exhibition and 21st Canadian Symposium on Remote Sensing, Ottawa, ON, Canada, 21-24 June, 1999; pp. 154–161.

59. Tóvári, D.; Vögtle, T. Object classifiaction in laserscanning data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2012**, *36*, Available online: https://www.researchgate.net/publication/228962142_Object_Classification_in_LaserScanning_Data (Accessed on 8 May 2021)

60. Galvanin, E.A.; Poz, A.P.D. Extraction of building roof contours from LiDAR data using a markov-random-field-based approach. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 981-987

61. Vosselmann, G. Slope based filtering of laser altimetry data. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Amsterdam, The Netherlands, 16–22 July, 2000; pp. 935–942.

62. Lohmann, P.; Koch, A.; Schaeffer, M. Approaches to the filtering of laser scanner data. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Amsterdam, The Netherlands, 16–22 July, 2000.

63. Tarsha-Kurdi, F.; Landes, T.; Grussenmeyer, P.; Smigiel, E. New approach for automatic detection of buildings in airborne laser scanner data using first echo only. In Proceedings of the ISPRS Comm. III Symposium, Photogrammetric Computer Vision, Bonn, Germany, 20–22 September, 2006; pp. 25–30.

64. Lodha, S.; Kreps, E.; Helmbold, D.; Fitzpatrick, D. Aerial LiDAR data classification using support vector machines (SVM). In Proceedings of the Third International Symposium on 3D Data Processing, Visualization, and Transmission, Chapel Hill, NC, USA, 14–16 June, 2006; pp. 567–574.

65. Rutzinger, M.; Höfle, B.; Pfeifer, N. Detection of high urban vegetation with airborne laser scanning data. In Proceedings of the Forestsat, Montpellier, France, 5–7 November, 2007, pp. 1–5.

66. Morsdorf, F.; Nichol, C.; Matthus, T.; Woodhouse, I.H. Assessing forest structural and physiological information content of multi-spectral LiDAR waveforms by radiative transfer modelling. *Remote Sens. Environ.* **2009**, *113*, 2152–2163.

67. Wang, C.K.; Tseng, Y.H.; Chu, H.J. Airborne dual-wavelength LiDAR data for classifying land cover. *Remote Sens.* **2014**, *6*, 700–715.

68. Wichmann, V.; Bremer, M.; Lindenberger, J.; Rutzinger, M.; Georges, C.; Petrini-Monteferri, F. Evaluating the potential of multispectral airborne LiDAR for topographic mapping and land cover classification. In Proceedings of the ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, La Grande Motte, France, 28 September–3 October 2015.

69. Puttonen, E.; Hakala, T.; Nevalainen, O.; Kaasalainen, S.; Krooks, A.; Karjalainen, M.; Anttila, K. Artificial target detection with a hyperspectral LiDAR over 26-h measurement. *Opt. Eng.* **2015**. Available online: https://www.spiedigitallibrary.org/jour-nals/optical-engineering/volume-54/issue-01/013105/Artificial-target-detection-with-a-hyperspectral-LiDAR-over-26-h/10.1117/1.OE.54.1.013105.full?SSO=1 (Accessed on 8 May 2021)

70. Ghaderpour, E.; Abbes, A.B.; Rhif, M.; Pagiatakis, S.D.; Farah, I.R. Non-stationary and unequally spaced NDVI time series analyses by the LSWAVE software. *Int. J. Remote Sens.* **2020**, *41*, 2374–2390.

71. Martinez, B.; Gilabert, M.A. Vegetation dynamics from NDVI time series analysis using the wavelet transform. *Remote Sens. Environ.* **2009**, *113*, 1823–1842.

72. Okin, G.S. Relative spectral mixture analysis—A multitemporal index of total vegetation cover. *Remote Sens. Environ.* **2007**, *106*, 467–479.

73. Yang, H.; Chen, W.; Qian, T.; Shen, D.; Wang, J. The Extraction of Vegetation Points from LiDAR Using 3D Fractal Dimension Analyses. *Remote Sens.* **2015**, *7*, 10815–10831.

74. Widlowski, J.L.; Pinty, B.; Gobron, N.; Verstraete, M.M. Detection and characterization of boreal coniferous forests from remote sensing data. *J. Geophys. Res.* **2001**, *106*, 33405–33419.

75. Koetz, B.; Sun, G.; Morsdorf, F.; Ranson, K.J.; Kneubühler, M.; Itten, K.; Allgöwer, B. Fusion of imaging spectrometer and LIDAR data over combined radiative transfer models for forest canopy characterization. *Remote Sens. Environ.* **2007**, *106*, 449–459.

76. Dian, Y.; Pang, Y.; Dong, Y.; Li, Z. Urban tree species mapping using airborne LiDAR and hyperspectral data. *J. Indian Soc. Remote Sens.* **2016**, *44*, 595–603.

77. Zhang, Z.; Liu, Q.; Wang, Y. Road extraction by deep residual u-net. *IEEE Geosci. Remote Sens. Lett.* **2018**. Available online: https://arxiv.org/abs/1711.10684 (Accessed on 8 May 2021)

78. Yang, X.; Li, X.; Ye, Y.; Zhang, X.; Zhang, H.; Huang, X.; Zhang, B. Road detection via deep residual dense u-net. In Proceedings of the International Joint Conference on Neural Networks, Budapest, Hungary, 14-19 July, 2019.

79. Miliaresis, G.; Kokkas, N. Segmentation and object-based classification for the extraction of the building class from LiDAR DEMs. *Comput. Geosci.* **2007**, *33*, 1076–1087.

80. Zhao, X.; Tao, R.; Li, W.; Li, H.C.; Du, Q.; Liao, W.; Philips, W. Joint Classification of Hyperspectral and LiDAR Data Using Hierarchical Random Walk and Deep CNN Architecture. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 7355–7370.

81. Herold, M.; Roberts, D.; Smadi, O.; Noronha, V. Road condition mapping with hyperspectral remote sensing. In Proceedings of the Airborne Earth Science Workshop, Pasadena, CA, USA, 31 March–2 April, 2004.

82. Kong, H.; Audibert, J.Y.; Ponce, J. General Road Detection from a Single Image. *IEEE Trans. Image Process.* **2010**. Available online: https://www.di.ens.fr/willow/pdfs/tip10b.pdf (Accessed on 7 May 2021)

83. Wu, P.C.; Chang, C.Y.; Lin, C. Lane-mark extraction for automobiles under complex conditions. *Pattern Recognit.* **2014**, *47*, 2756–2767.

84. Lin, Y.-C.; Lin, C.; Tsai, M.-D.; Lin, C.-L. Object-based analysis of LiDAR geometric features for vegetation detection in shaded areas. In Proceedings of the XXIII ISPRS Congress, Prague, Czech Republic, 12–19 July, 2016.

85. Poux, F.; Hallot, R.P.; Neuville, R.B. Smart point cloud: Definition and remaining challenges. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2016**, *42*, 119-127.

86. Arief, H.A.; Strand, G.H.; Tveite, H.; Indahl, U.G. Land cover segmentation of airborne LiDAR data using Stochastic Atrous Network. *Remote Sens.* **2018**, *10*, 973.

87. Clark, R.N. Spectroscopy of rocks and minerals, and principles of spectroscopy. In *Manual of Remote Sensing, Remote Sensing for the Earth Sciences*, Rencz, A.N., Ed.; John Wiley and Sons: New York, NY, USA, 1999; Volume 3.

88. Signoroni, A.; Savardi, M.; Baronio, A.; Benini, S. Deep learning meets hyperspectral image analysis: a multidisciplinary review. *J. Imaging* **2019**, *5*.

89. Ben-Dor, E. Imaging spectrometry for urban applications. In *Imaging spectrometry*, van der Meer, F.D., de Jong, S.M., Eds.; Kluwer Academic Publishers: Amsterdam, The Netherlands, 2001; pp. 243–281.

90. Ortenberg, F. Hyperspectral Sensor Characteristics. In *Fundamentals, Sensor Systems, Spectral Libraries, and Data Mining for Vegetation*, 2nd ed.; Huete, A., Lyon, J.G., Thenkabail, P.S., Eds.; Hyperspectral remote sensing of vegetation Volume I.; CRC Press: Boca Raton, FL, USA, 2011; p. 449.

91. Rossel, R.A.V.; McGlynn, R.N.; McBratney, A.B. Determining the composition of mineral-organic mixes using UV—vis—NIR diffuse reflectance spectroscopy. *Geoderma* **2006**, *137*, 70–82.

92. Adam, E.; Mutanga, O.; Rugege, D. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation:a review. *Wetl. Ecol. Manag.* **2010**, *18*, 281-296.

93. Heiden, U.H.W.; Roessner, S.; Segl, K.; Esch, T.; Mueller, A. Urban structure type characterization using hyperspectral remote sensing and height information. *Landsc. Urban. Plan.* **2012**, *105*, 361–375.

94. Roessner, S.; Segl, K.; Heiden, U.; Kaufmann, H. Automated differentiation of urban surfaces based on airborne hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 1525-1532.

95. Townshend, J.; Justice, C.; Li, W.; Gumey, C.; McManus, J. Global land cover classification by remote sensing: present capabilities and future possibilities. *Remote Sens. Environ.* **1991**, *35*, 243–255.

96. Heiden, U.; Roessner, S.; Segl, K.; Kaufmann, H. Analysis of Spectral Signatures of Urban Surfaces for their Identification Using Hyperspectral HyMap Data. In Proceedings of the IEEE/ISPRS Joint Workshop on Remote Sensing nd Data Fusion over Urban Areas, Rome, Italy, 8–9 November 2001; pp. 173–177.

97.   Heiden, U.; Segl, K.; Roessner, S.; Kaufmann, H. Determination of robust spectral features for identification of urban surface materials in hyperspectral remote sensing data. *Remote Sens. Environ.* **2007**, *111*, 537–552.

98.   Meer, F.v.d.; Jong, S.d.; Bakker, W. Imaging Spectrometry: Basic Analytical Techniques. In *Imaging Spectrometry*; van der Meer, F.D., de Jong, S.M., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2001; pp. 17–61.

99.   Franke, J.; Roberts, D.A.; Halligan, K.; Menz, G. Hierarchical Multiple Endmember Spectral Mixture Analysis (MESMA) of hyperspectral imagery for urban environments. *Remote Sens. Environ.* **2009**, *113*, 1712–1723.

100.   Hepner, G.F.; Housmand, B.; Kulikov, I.; Bryant, N. Investigation of the integration of AVIRIS and IFSAR for urban analysis. *Photogramm. Eng. Remote Sens.* **1998**, *64*, 813–820.

101.   Linden, S.v.d.; Okujeni, A.; Canters, F.; Degerickx, J.; Heiden, U.; Hostert, P.; Priem, F.; Somers, B.; Thiel, F. Imaging Spectroscopy of Urban Environments. *Surv. Geophys.* **2019**, *40*, 471–488.

102.   Pillay, R.; Picollo, M.; Hardeberg, J.Y.; George, S. Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems. *Sensors* **2020**, *20*, 3812.

103.   Yao, H.; Tian, L.F. Practical methods for geometric distortion correction of aerial hyperspectral imagery. *Appl. Eng. Agric.* **2004**, *20*, 367–375.

104.   Lulla, V.; Jensen, R.R. Best Practices for Urban Hyperspectral Remote Sensing Data Acquisition and Processing. In *Urban Sustainability: Policy and Praxis*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 43–54.

105.   Galbraith, A.E.; Theiler, J.; Thome, K.; Ziolkowski, R. Resolution Enhancement of Multilook Imagery for the Multispectral Thermal Imager. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 1964–1977.

106.   Pepe, M.; Fregonese, L.; Scaioni, M. Planning airborne photogrammetry and remote-sensing missions with modern platforms and sensors. *Eur. J. Remote Sens.* **2018**, *51*, 412–436.

107.   Heiden, U.; Segl, K.; Roessner, S.; Kaufmann, H. Determination and verification of robust spectral features for an automated classification of sealed urban surfaces. In Proceedings of the EARSeL Workshop on Imaging Spectroscopy, Warsaw, Poland, 27–29 April, 2005.

108.   Lacherade, S.; Miesch, C.; Briottet, X.; Men, H.L. Spectral variability and bidirectional reflectance behavior of urban materials at a 20 cm spatial resolution in the visible and near-infrared wavelength. A case study over Toulouse (France). *Int. J. Remote Sens.* **2005**, *26*, 3859–3866.

109.   Herold, M.; Roberts, D.A.; Gardner, M.E.; Dennison, P.E. Spectrometry for urban area remote sensing - development and analysis of a spectral library from 350 to 2400 nm. *Remote Sens. Environ.* **2004**, *91*, 304–319.

110.   Ilehag, R.; Schenk, A.; Huang, Y.; Hinz, S. KLUM: An Urban VNIR and SWIR Spectral Library Consisting of Building Materials. *Remote Sens.* **2019**, *11*.

111.   Manolakis, D.; Lockwood, R.; Cooley, T. *Hyperspectral Imaging Remote Sensing: Physics, Sensors, and Algorithms*; Cambridge University Press: Cambridge, UK, 2016.

112.   Pearson, K. On lines and planes of closest fit to systems of points in space. *Philos. Mag. Lett.* **1901**, *2*, 559–572.

113.   Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **2000**, *155*, 945–959.

114.   Lawton, W.H.; Sylvestre, E.A. Self modeling curve resolution. *Technometrics* **1971**, *13*, 617–633.

115.   Vidal, M.; Amigo, J.M. Pre-processing of hyperspectral images. Essential steps before image analysis. *Chemom. Intell. Lab. Syst.* **2012**, *117*, 138–148.

116.   Pandey, D.; Tiwari, K.C. Spectral library creation and analysis of urban built-up surfaces and materials using field spectrometry. *Arab. J. Geosci.* **2021**, *14*.

117.   Miller, D.L.; Alonzo, M.; Roberts, D.A.; Tague, C.L.; McFadden, J.P. Drought response of urban trees and turfgrass using airborne imaging spectroscopy. *Remote Sens. Environ.* **2020**, *40*.

118.   Clark, R.N.; Swayze, G.A.; Livo, K.E.; Kokaly, R.F.; Sutley, S.H.; Dalton, J.B.; McDougal, R.R.; Gent, C.A. Imaging spectroscopy: earth and planetary remote sensing with the USGS Tetracorder and expert systems. *J. Geophys. Res.* **2003**, *108*.

119.   Yongyang, X.; Liang, W.; Zhong, X.; Zhanlong, C. Building extraction in very high resolution remote sensing imagery using deep learning and guided filters. *Remote Sens.* **2018**, *10*.

120.   Teo, T.A.; Wu, H.M. Analysis of land cover classification using multi-wavelength LiDAR system. *Appl. Sci.* **2017**, *7*, https://doi.org/10.3390/app7070663

121.   Pandey, D.; Tiwari, K.C. New spectral indices for detection of urban built-up surfaces and its sub-classes in AVIRIS-NG hyperspectral imagery. *Geocarto Int.* **2020**.

122.   Zha, Y.; Gao, J.; Ni, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* **2003**, *24*, 583–594.

123.   Estoque, R.C.; Zhang, W.; Tan, K.; Liu, Y.; Liu, S. Multiple classifier system for remote sensing image classification: A review. *Sensors* **2012**, *12*, 4764–4792.

124.   Shahi, K.; Shafri, H.Z.M.; Taherzadeh, E.; Mansor, S.; Muniandy, R. A novel spectral index to automatically extract road networks from WorldView-2 satellite imagery. *Egypt. J. Remote Sens. Space Sci.* **2015**, *18*, 27–33.

125.   Xie, H.; Luo, X.; Xu, X.; Tong, X.; Jin, Y.; Haiyan, P.; Zhou, B. New hyperspectral difference water index for the extraction of urban water bodies by the use of airborne hyperspectral images. *J. Appl. Remote Sens.* **2014**, *8*.

126.   Xue, J.; Zhao, Y.; Bu, Y.; Liao, W.; Chan, J.C.-W.; Philips, W. Spatial-Spectral Structured Sparse Low-Rank Representation for Hyperspectral Image Super-Resolution. *IEEE Trans. Image Process.* **2021**, *30*, 3084–3097.

127. Rasti, B.; Scheunders, P.; Ghamisi, P.; Licciardi, G.; Chanussot, J. Noise Reduction in Hyperspectral Imagery: Overview and Application. *Remote Sens.* **2018**, *3*.
128. Gómez-Chova, L.; Alonso, L.; Guanter, L.; Camps-Valls, G.; Calpe, J.; Moreno, J. Correction of systematic spatial noise in push-broom hyperspectral sensors: application to CHRIS/PROBA images. *Appl. Opt.* **2008**, *47*, 46–60.
129. Bruzzone, L.; Marconcini, M.; Persello, C. Fusion of spectral and spatial information by a novel SVM classification technique. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–27 July, 2007.
130. Bovolo, F.; Bruzzone, L. A Context-Sensitive Technique Based on Support Vector Machines for Image Classification. In Proceedings of the International Conference on Pattern Recognition and Machine Intelligence, Kolkata, India, 20-22 December, 2005; pp. 260–265.
131. Farag, A.A.; Mohamed, R.H.; El-Baz, A. A unified framework for MAP estimation in remote sensing image segmentation. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 1617–1634.
132. Sun, L.; Wu, Z.; Liu, J.; Xiao, L.; Wei, Z. Supervised spectral-spatial hyperspectral image classification with weighted Markov Random Fields. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1490–1503.
133. Li, J.; Bioucas-Dias, J.M.; Plaza, A. Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and markov random fields. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*.
134. Wehr, A.; Lohr, U. Airborne laser scanning - an introduction and overview. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 68–82.
135. Clode, S.; Rottensteiner, F.; Kootsookos, P.; Zelniker, E. Detection and vectorization of roads from LiDAR data. *Photogramm. Eng. Remote Sens.* **2007**, *73*, 517–535.
136. Chehata, N.; Guo, L.; Mallet, C. Airborne LiDAR feature selection for urban classification using random forests. *Laserscanning* **2009**, *38*.
137. Guo, L.; Chehata, N.; Mallet, C.; Boukir, S. Relevance of airborne LiDAR and multispectral image data for urban scene classification using random forests. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 56–66.
138. Priestnall, G.; Jaafar, J.; Duncan, A. Extracting urban features from LiDAR digital surface models. *Comput. Environ. Urban. Syst.* **2000**, *24*, 65–78.
139. Hecht, R.; Meinel, G.; Buchroithner, M.F. Estimation of urban green volume based on single-pulse LiDAR data. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*.
140. Alonso, L.; Picos, J.; Bastos, G.; Armesto, J. Detection of Very Small Tree Plantations and Tree-Level Characterization Using Open-Access Remote-Sensing Databases. **2020**, *12*, 2276.
141. Grohmann, C.H.; Smith, M.J.; Riccomini, C. Muli-scale Analysis of Topographic Surface Roughness in the Midland Valley, Scotland. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 1200–1213.
142. Brubaker, K.M.; Myers, W.L.; Drohan, P.J.; Miller, D.A.; Boyer, E.W. The Use of LiDAR Terrain Data in Characterizing Surface Roughness and Microtopography. *Appl. Environ. Soil Sci.* **2013**, *2013*.
143. Brenner, C. Dreidimensionale Gebäuderekonstruktion aus digitalen Oberflächenmodellen und Grundrissen. Ph.D. Thesis, Stuttgart University, Stuttgart, Germany, 2000.
144. Antonarakis, A.S.; Richards, K.S.; Brasington, J. Object-based land cover classification using airborne LiDAR. *Remote Sens. Environ.* **2008**, *112*, 2988–2998.
145. Charaniya, A.P.; Manduchi, R.; Lodha, S.K. Supervised parametric classification of aerial LiDAR data. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Washington, DC, USA, 27 June–2 July, 2004.
146. Bartels, M.; Wei, H. Maximum likelihood classification of LiDAR data incorporating multiple co-registered band. In Proceedings of the 4th International Workshop on Pattern Recognition in Remote Sensing in conjunction with the 18th International Conference on Pattern Recognition, Hong Kong, 20–24 August, 2006.
147. Im, J.; Jensen, J.R.; Hodgson, M.E. Object-based land cover classification using high-posting-density LiDAR data. *GIsci. Remote Sens.* **2008**, *45*, 209–228.
148. Song, J.H.; Han, S.H.; Yu, K.Y.; Kim, Y.I. Assessing the possibility of land-cover classification using LiDAR intensity data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2002**, *34*, 259–262.
149. Yoon, J.-S.; Lee, J.-I. Land cover characteristics of airborne LiDAR intensity data: a case study. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 801–805.
150. MacFaden, S.W.; O'Neil-Dunne, J.P.M.; Royar, A.R.; Lu, J.W.T.; Rundle, A.G. High-resolution tree canopy mapping for New York City using LiDAR and object-based image analysis. *J. Appl. Remote Sens.* **2012**, *6*.
151. Yan, W.Y.; Shaker, A. Reduction of striping noise in overlapping LiDAR intensity data by radiometric normalization. In Proceedings of the XXIII ISPRS Congress, Prague, Czech Republic, 12-19 July, 2016.
152. Nobrega, R.A.A.; Quintanilha, J.A.; O'Hara, C.G. A noise-removal approach for LiDAR intensity images using anisotropic diffusion filtering to preserve object shape characteristics. In Proceedings of the ASPRS Annual Conference, Tampa, FL, USA, 7–11 May, 2007.
153. Minh, N.Q.; Hien, L.P. Land cover classification using LiDAR intensity data and neural network. *J. Korean Soc. Surv. Geodesy Photogramm. Cartogr.* **2011**, *29.4*, 429–438.
154. Brennan, R.; Webster, T.L. Object-oriented land cover classification of LiDAR-derived surfaces. *Can. J. Remote Sens.* **2006**, *32*, 162–172.
155. Wagner, W. Radiometric calibration of small-footprint full-waveform airborne laser scanner measurements: basic physical concepts. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 505–513.
156. Mallet, C.; Bretar, F. Full-waveform topographic LiDAR: state-of-the-art. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 1–16.

157. Bretar, F.; Chauve, A.; Mallet, C.; Jutzi, B. Managing full waveform LiDAR data: a challenging task for the forthcoming years. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2008**, *XXXVII*, 415–420.
158. Kirchhof, M.; Jutzi, B.; Stilla, U. Iterative processing of laser scanning data by full waveform analysis. *ISPRS J. Photogramm. Remote Sens.* **2008**, *63*, 99–114.
159. Höfle, B.; Pfeifer, N. Correction of laser scanning intensity data: data and model-driven approaches. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 1415–1433.
160. Alexander, C.; Tansey, K.; Kaduk, J.; Holland, D.; Tate, N.J. Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 423–432.
161. Neuenschwander, A.L.; Magruder, L.A.; Tyler, M. Landcover classification of small-footprint, full-waveform LiDAR data. *J. Appl. Remote Sens.* **2009**, *3*.
162. Jutzi, B.; Stilla, U. Waveform processing of laser pulses for reconstruction of surfacer in urban areas. *Meas. Tech.* **2005**, Available online: https://www.researchgate.net/publication/43136634_Waveform_processing_of_laser_pulses_for_reconstruction_of_surfaces_in_urban_areas
163. Chauve, A.; Mallet, C.; Bretar, F.; Durrieu, S.; Pierrot-Deseilligny, M.; Puech, W. Processing full-waveform LiDAR data: modelling raw signals. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2007**, *36*, 102–107.
164. Gross, H.; Jutzi, B.; Thoennessen, U. Segmentation of tree regions using data of a full-waveform laser. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2007**, *36*, 57–62.
165. Reitberger, J.; Krzystek, P.; Stilla, U. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. *Int. J. Remote Sens.* **2008**, *29*, 1407–1431.
166. Rutzinger, M.; Höfle, B.; Hollaus, M.; Pfeifer, N. Object-based point cloud analysis of full-waveform airborne laser scanning data for urban vegetation classification. *Sensors* **2008**, *8*, 4505–4528.
167. Melzer, T. Non-parametric segmentation of ALS point clouds using mean shift. *J. Appl. Geod.* **2007**, *1*, 158–170.
168. Lin, Y.-C.; Mills, J.P. Factors influencing pulse width of small footprint, full waveform airborne laser scanning data. *Photogramm. Eng. Remote Sens.* **2010**, *76*, 49–59.
169. Doneus, M.; Briese, C.; Fera, M.; Janner, M. Archaeological prospection of forested areas using full-waveform airborne laser scanning. *J. Archaeol. Sci.* **2008**, *35*, 882–893.
170. Harding, D.; Lefsky, M.; Parker, G. Laser altimeter canopy height profiles. Methods and validation for closed canopy, broadleaf forests. *Remote Sens. Environ.* **2001**, *76*, 283–297.
171. Gross, H.; Thoennessen, U. Extraction of lines from laser point clouds. In Proceedings of the ISPRS Conference Photogrammetric Image Analysis (PIA), Bonn, Germany, 20-22 September, 2006; pp. 87–91.
172. West, K.F.; Webb, B.N.; Lersch, J.R.; Pothier, S.; Triscari, J.M.; Iverson, A.E. Context-driven automated target detection in 3-D data. In Proceedings of the Automatic Target Recognition XIV, Orlando, FL, USA, 13-15 April, 2004; pp. 133–143.
173. Ojala, T.; Pietikainen, M.; Maenpaa, T.T. Multi resolution gray scale and rotation invariant texture classification with local binary pattern. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987.
174. Ge, C.; Du, Q.; Sun, W.; Wang, K.; Li, J.; Li, Y. Deep Residual Network-Based Fusion Framework for Hyperspectral and LiDAR Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 2458–2472.
175. Peng, B.; Li, W.; Xie, X.; Du, Q.; Liu, K. Weighted-Fusion-Based Representation Classifiers for Hyperspectral Imagery. *Remote Sens.* **2015**, *7*, 14806–14826.
176. Manjunath, B.S.; Ma, W.Y. Texture features for browsing and retrieval of image data. *IEEE Trans. Pattern Anal. Mach. Intell.* **1996**, *18*, 837–842.
177. Rajadell, O.; García-Sevilla, P.; Pla, F. Textural Features for Hyperspectral Pixel Classification. In Proceedings of the Iberian Conference on Pattern Recognition and Image Analysis, Póvoa de Varzim, Portugal, 10-12 June, 2009; pp. 208–216.
178. Aksoy, S. Spatial techniques for image classification. In *Signal and Image Processing for Remote Sensing*; CRC Press: Boca Raton, FL, USA, 2006; pp. 491–513.
179. Zhang, G.; Jia, X.; Kwok, N.M. Spectral-spatial based super pixel remote sensing image classification. In Proceedings of the 4th International Congress on Image and Signal Processing, Shanghai, China, 15-17 October, 2011; pp. 1680–1684.
180. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. *IEEE Trans. Syst. Man Cybern. Syst.* **1973**, *3*, 610–621.
181. Huang, X.; Zhang, L.; Gong, W. Information fusion of aerial images and LiDAR data in urban areas: vector-stacking, re-classification and post-processing approaches. *Int. J. Remote Sens.* **2011**, *32*, 69–84.
182. Puissant, A.; Hirsch, J.; Weber, C. The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *Int. J. Remote Sens.* **2005**.
183. Zhang, Y. Optimisation of building detection in satellite images by combining multispectral classification and texture filtering. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 50–60.
184. Huang, X.; Zhang, L.; Li, P. An Adaptive Multiscale Information Fusion Approach for Feature Extraction and Classification of IKONOS Multispectral Imagery Over Urban Areas. *IEEE Geosci. Remote Sens. Lett.* **2007**, *4*.
185. Pesaresi, M.; Benediktsson, J.A. A New Approach for the Morphological Segmentation of High-Resolution Satellite Imagery. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 309–320.
186. Soille, P.; Pesaresi, M. Advances in mathematical morphology applied to geoscience and remote sensing. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*.

187. Benediktsson, J.A.; Palmason, J.A.; Sveinsson, J.R. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 480–491.
188. Benediktsson, J.A.; Pesaresi, M.; Amason, K. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1940–1949.
189. Jouni, M.; Mura, M.D.; Comon, P. Hyperspectral Image Classification Based on Mathematical Morphology and Tensor Decomposition. *Math. Morphol. Theory Appl.* **2020**, *4*, 1–30.
190. Mura, M.D.; Benediktsson, J.A.; Waske, B.; Bruzzone, L. Extended profiles with morphological attribute filters for the analysis of hyperspectral data. *Int. J. Remote Sens.* **2010**, *31*, 5975–5991.
191. Aptoula, E.; Mura, M.D.; Lefevre, S. Vector attribute profiles for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3208–3220.
192. Sayed, W.M. Processing of LiDAR Data using Morphological Filter. *Int. J. Adv. Res.* **2014**, *2*, 361–367.
193. Rottensteiner, F.; Briese, C. A new method for building extraction in urban areas from high-resolution LiDAR data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2002**, *34*, 295–301.
194. Morsy, S.S.A.; El-Rabbany, A. Multispectral LiDAR Data for Land Cover Classification of Urban Areas. *Sensors* **2017**, *17*.
195. Suomalainen, J.; Hakala, T.; Kaartinen, H.; Räikkönen, E.; Kaasalainen, S. Demonstration of a virtual active hyperspectral LiDAR in automated point cloud classification. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 637–641.
196. Hakala, T.; Suomalainen, J.; Kaasalainen, S.; Chen, Y. Full waveform hyperspectral LiDAR for terrestrial laser scanning. *Opt. Express* **2012**, *20*.
197. Hughes, G.F. On the mean accuracy of statistical pattern recognizers. *IEEE Trans. Inf. Theory* **1968**, *14*, 55–63.
198. Asner, G.P.; Knapp, D.E.; Boardman, J.; Green, R.O.; Kennedy-Bowdoin, T.; Eastwood, M.; Martin, R.E.; Anderson, C.; Field, C.B. Carnegie Airborne Observatory-2: Increasing science data dimensionality via high-fidelity multi-sensor fusion. *Remote Sens. Environ.* **2012**, *124*, 454–465.
199. Brell, M.; Rogass, C.; Segl, K.; Bookhagen, B.; Guanter, L. Improving sensor fusion: a parametric method for the geometric coalignment of airborne hyperspectral and LiDAR data. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*.
200. Brell, M.; Segl, K.; Guanter, L.; Bookhagen, B. 3D hyperspectral point cloud generation: fusing airborne laser scanning and hyperspectral imaging sensors for improved object-based information extraction. *ISPRS J. Photogramm. Remote Sens.* **2019**, *149*, 200–214.
201. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Feitosa, R.Q.; Meer, F.v.d.; Werff, H.v.d.; Coillie, F.v.; et al. Geographic Object-Based Image Analysis - Toward a new paradigm. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 180–191.
202. Campagnolo, M.L.; Cerdeira, J.O. Contextual classification of remotely sensed images with integer linear programming. In Proceedings of the Computational Modeling of Objects Represented in Images: Fundamentals, Methods, and Applications, Niagara Falls, NY, USA, 2007; pp. 123–128.
203. Jong, S.M.D.; Hornstra, T.; Maas, H. An integrated spatial and spectral approach to the classification of Mediterranean land cover types: the SSC method. *Int. J. Appl. Earth Obs. Geoinf.* **2001**, *3*, 176–183.
204. Bhaskaran, S.; Paramananda, S.; Ramnarayan, M. Per-pixel and object-oriented classification methods for mapping urban features using IKONOS satellite data. *Appl. Geogr.* **2010**, *30*, 650–665.
205. Baker, B.A.; Warner, T.A.; Conley, J.F.; McNeil, B.E. Does spatial resolution matter? A multi-scale comparison of object-based and pixel-based methods for detection change associated with gas well drilling operations. *Int. J. Remote Sens.* **2013**, *34*, 1633–1651.
206. Johnson, B.; Xie, Z. Unsupervised image segmentation evaluation and refinement using a multi-scale approach. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 473–483.
207. Zhong, P.; Gong, Z.; Li, S.; Schönlieb, C.B. Learning to diversify deep belief networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, 55.
208. Liu, P.; Zhang, H.; Eom, K.B. Active deep learning for classification of hyperspectral images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*.
209. Chen, Y.; Zhao, X.; Jia, X. Spectral-spatial classification of hyperspectral data based on deep belief network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*.
210. Lin, Z.; Chen, Y.; Zhao, X.; Wang, G. Spectral–Spatial Classification of Hyperspectral Image Using Autoencoders. In Proceedings of the 9th International Conference on Information, Communications Signal Processing, Tainan, Taiwan, 10-13 December, 2013; pp. 1–5.
211. Tao, C.; Pan, H.; Li, Y.; Zhou, Z. Unsupervised spectral-spatial feature learning with stacked sparse autoencoder for hyperspectral imagery classification. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 2438–2442.
212. Yue, J.; Mao, S.; Li, M. A deep learning framework for hyperspectral image classification using spatial pyramid pooling. *Remote Sens. Lett.* **2016**, *7*, 875–884.
213. Campbell, J.B. *Introduction to Remote Sensing*, 3 Guilford Press Ed.; New York, NY, USA, 2002; p. 621.
214. Enderle, D.I.M.; Weih, R.C., Jr. Integrating supervised and unsupervised classification methods to develop a more accurate land cover classification. *J. Ark. Acad. Sci.* **2005**, *59*, 65–73.
215. Shabbir, S.; Ahmad, M. Hyperspectral Image Classification - Traditional to Deep Models: A Survey for Future Prospects. *arXiv:2101.06116*, **2021**.
216. Liu, C.; He, L.; Li, Z.; Li, J. Feature-driven active learning for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 341–357.

217. Richards, J.A. *Remote Sensing Digital Image Analysis: An. Introduction*, 3rd ed.; Springer-Verlag New York, Eds.; Springer: Secaucus, NJ, USA, 1999.
218. Garcia, S.; Zhang, Z.L.; Altalhi, A.; Alshomrani, S.; Herrera, F. Dynamic ensemble selection for multi-class imbalanced datasets. *Inf. Sci.* **2018**, *445-446*, 22-37.
219. Lv, Q.; Feng, W.; Quan, Y.; Dauphin, G.; Gao, L.; Xing, M. Enhanced-Random-Feature-Subspace-Based Ensemble CNN for the Imbalanced Hyperspectral Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 3988–3999.
220. Paing, M.P.; Pintavirooj, C.; Tungjitkusolmun, S.; Choomchuay, S.; Hamamoto, K. Comparison of sampling methods for imbalanced data classification in random forest. In Proceedings of the 11th Biomedical Engineering International Conference, Chaing Mai, Thailand, 2018; pp. 1–5.
221. Momeni, R.; Aplin, P.; Boyd, D.S. Mapping Complex Urban Land Cover from Spaceborne Imagery: The Influence of Spatial Resolution, Spectral Band Set and Classification Approach. *Remote Sens.* **2016**, *8*.
222. Rasti, B.; Hong, D.; Hang, R.; Ghamisi, P.; Kang, X.; Chanussot, J.; Benediktsson, J.A. Feature Extraction for Hyperspectral Imagery: The Evolution from Shallow to Deep (Overview and Toolbox). *IEEE Geosci. Remote Sens. Lett.* **2020**, *8*, 60–88.
223. Rauber, P.E.; Fadel, S.G.; Falcao, A.X.; Telea, A.C. Visualizing the hiden activity of artificial neural networks. *IEEE Trans. Vis. Comput. Graph.* **2017**, *23*, 101–110.
224. Paoletti, M.E.; Haut, J.M.; Plaza, J.; Plaza, A. A new deep convolutional neural network for fast hyperspectral image classification. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 120–147.
225. Yu, S.; Jia, S.; Xu, C. Convolutional neural networks for hyperspectral image classification. *Neurocomputing* **2017**, *219*, 88–98.
226. Zhou, W.; Kamata, S. Multi-Scanning Based Recurrent Neural Network for Hyperspectral Image Classification. In Proceedings of the 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10-15 January, 2021.
227. Lee, H.; Kwon, H. Contextual deep CNN based hyperspectral classification. In Proceedings of the Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10-15 July, 2016; pp. 3322–3325.
228. Seidel, D.; Annighöfer, P.; Thielman, A.; Seifert, Q.E.; Thauer, J.H.; Glatthorn, J.; Ehbrecht, M.; Kneib, T.; Ammer, C. Predicting Tree Species From 3D Laser Scanning Point Clouds Using Deep Learning. *Front. Plant. Sci.* **2021**, *12*.
229. Huang, L.; Chen, Y. Dual-Path Siamese CNN for Hyperspectral Image Classification with Limited Training Samples. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 518–522.
230. Neagoe, V.E.; Diaconescu, P. CNN Hyperspectral Image Classification Using Training Sample Augmentation with Generative Adversarial Networks. In Proceedings of the 13th International Conference on Communications (COMM), Bucharest, Romania, 18-20 June 2020.
231. Haut, J.M.; Paoletti, M.E.; Plaza, J.; Plaza, A.; Li, J. Hyperspectral Image Classification Using Random Occlusion Data Augmentation. *IEEE Geosci. Remote Sens. Lett.* **2019**.
232. Yu, C.; Han, R.; Song, M.; Liu, C.; Chang, C.I. A Simplified 2D-3D CNN Architecture for Hyperspectral Image Classification Based on Spatial-Spectral Fusion. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*.
233. Hochreiter, S. The Vanishing Gradient Problem During Learning Recurrent Neural Net and Problem Solutions *Int. J. Uncertain. Fuzziness Knowledge Based Syst.* **1998**, *6*, 107–116.
234. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*.
235. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on International Conference on Machine Learning, Lille, France, 6-11 July, 2015.
236. Ba, J.L.; Kiros, J.R.; Hinton, G.E. Layer normalization. *arXiv* **2016**, arXiv:1607.06450.
237. Camps-Valls, G.; Tuia, D.; Bruzzone, L.; Benediktsson, J.A. Advances in hyperspectral image classification: Earth monitoring with statistical learning methods. *IEEE Signal. Process. Mag.* **2014**, *31*, 45–54.
238. Linden, S.v.d.; Janz, A.; Waske, B.; Eiden, M.; Hostert, P. Classifying segmented hyperspectral data from a heterogeneous urban environment using support vector machines. *J. Appl. Remote Sens.* **2007**, *1*.
239. Li, W.; Wu, G.; Zhang, F.; Du, Q. Hyperspectral image classification using deep pixel-pair features. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 844–853.
240. Fauvel, M.; Chanussot, J.; Benediktsson, J.A. Evaluation of kernels for multiclass classification of hyperspectral remote sensing data. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Toulouse, France, 14-19 May, 2006; pp. II-813–II-816.
241. Chen, Y.; Nasrabadi, N.M.; Tran, T. D. "Hyperspectral image classification using dictionary-based sparse representation," IEEE Geosci. Remote Sens. vol. 49, pp.3973-3985, 2011. https://doi.org/10.1109/TGRS.2011.2129595
242. Plaza, A.; Benediktsson, J.A.; Boardman, J.W.; Brazile, J.; Bruzzone, L.; Camps-Valls, G.; Chanussot, J.; Fauvel, M.; Gamba, P.; Gualtieri, A.; et al. Recent advances in techniques for hyperspectral image processing. *Remote Sens. Environ.* **2009**, *13*, 110–122.
243. Mallet, C.; Soergel, U.; Bretar, F. Analysis of full-waveform LiDAR data for classification of urban areas. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2008**, *37*, 85–92.
244. Zhang, Y.; Cao, G.; Li, X.; Wang, B. Cascaded random forest for hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**.
245. Matikainen, L.; Karila, K.; Hyyppä, J.; Litkey, P.; Puttonen, E.; Ahokas, E. Object-based analysis of multispectral airborne laser scanner data for land cover classification and map updating. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 298–313.

246. Mei, S.; Ji, J.; Bi, Q.; Hou, J.; Du, Q.; Li, W. Integrating spectral and spatial information into deep convolutional neural networks for hyperspectral classification. In Proceedings of the International Geoscience and Remote Sensing Symposium, Beijing, China, 10-15 July, 2016; pp. 5067–5070.

247. Ran, L.; Zhang, Y.; Wei, W.; Zhang, Q. A hyperspectral image classification framework with spatial pixel pair features. *Sensors* **2017**, *17*, 2421.

248. Li, Y.; Zhang, H.; Shen, Q. Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sens.* **2017**, *9*, 67.

249. Vaddi, R.; Manoharan, P. Hyperspectral image classification using CNN with spectral and spatial features integration. *Infrared Phys. Technol.* **2020**, *107*.

250. Ge, Z.; Cao, G.; Li, X.; Fu, P. Hyperspectral Image Classifiacation Method Based on 2D-3D CNN and Multibranch Feature Fusion. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 5776–5788.

251. Vaddi, R.; Manoharan, P. CNN based hyperspectral image classification using unsupervised band selection and structure-preserving spatial features. *Infrared Phys. Technol.* **2020**, *110*.

252. Guo, H.; Liu, J.; Xiao, Z.; Xiao, L. Deep CNN-based hyperspectral image classification using discirmnative multiple spatial-spectral feature fusion. *Remote Sens. Lett.* **2020**, *11*.

253. Wang, J.; Song, X.; Sun, L.; Huang, W.; Wang, J. A Novel Cubic Convolutional Neural Network for Hyperspectral Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 4133–4148.

254. Gong, H.; Li, Q.; Li, C.; Dai, H.; He, Z.; Wang, W.; Li, H.; Han, F.; Tuniyazi, A.; Mu, T. Multiscale Information Fusion for Hyperspectral Image Classification Based on Hybrid 2D-3D CNN. *Remote Sens.* **2021**, *13*.

255. Kutluk, S.; Kayabol, K.; Akan, A. A new CNN training approach with application to hyperspectral image classification. *Digit. Signal. Process.* **2021**, *113*.

256. Yin, J.; Qi, C.; Chen, Q.; Qu, J. Spatial-Spectral Network for Hyperspectral Image Classification: A 3-D CNN and Bi-LSTM Framework. *Remote Sens.* **2021**, *13*.

257. He, X.; Chen, Y.; Lin, Z. Spatial-Spectral Transformer for Hyperspectral Image Classification. *Remote Sens.* **2021**, *13*, 498.

258. Rao, M.; Tang, P.; Zhang, Z. A Developed Siamese CNN with 3D Adaptive Spatial-Spectral Pyramid Pooling for Hyperspectral Image Classification. *Remote Sens.* **2020**, *12*, 1964.

259. Pan, S.; Guan, H.; Chen, Y.; Yu, Y.; Goncalves, W.N.; Junior, J.M.; Li, J. Land-cover classification of multispectral LiDAR data using CNN with optimized hyper-parameters. *ISPRS J. Photogramm. Remote Sens.* **2020**, *166*, 241–254.

260. Xie, J.; Chen, Y. LiDAR Data Classification Based on Automatic Designed CNN. *IEEE Geosci. Remote Sens. Lett.* **2020**.

261. Hang, R.; Li, Y.; Ghamisi, P.; Hong, D.; Xia, G.; Liu, Q. Classification of Hyperspectral and LiDAR Data Using Coupled CNNs. *IEEE Trans. Geosci. Remote Sens.* **2020**, *68*, 4939–4950.

262. Feng, Q.; Zhu, D.; Yang, J.; Li, B. Multisource Hyperspectral and LiDAR Data Fusion for Urban Land-Use Mapping based on a Modified Two-Branch Convolutional Neural Network. *ISPRS Int. J. Geoinf.* **2019**, *8*.

263. Zhang, M.; Li, W.; Du, Q.; Gao, L.; Zhang, B. Feature Extraction for Classification of Hyperspectral and LiDAR Data Using Patch-to-Patch CNN. *IEEE Trans. Cybern.* **2020**, *50*, 100–111.

264. Chen, Y.; Li, C.; Ghamisi, P.; Jia, X.; Gu, Y. Deep fusion of remote sensing data for accurate classification. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1253–1257.

265. Morchhale, S.; Pauca, V.P.; Plemmons, R.J.; Torgersen, T.C. Classification of pixel-level fused hyperspectral and LiDAR data using deep convolutional neural networks. In Proceedings of the 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Los Angeles, CA, USA, 21-24 August, 2016; pp. 1–5.

266. Wu, H.; Prasad, S. Convolutional recurrent neural networks for hyperspectral data classification. *Remote Sens.* **2017**, *9*, 298.

267. Venkatesan, R.; Prabu, S. Hyperspectral image features classification using deep learning recurrent neural networks. *J. Med. Syst.* **2019**, *43*.

268. Paoletti, M.E.; Haut, J.M.; Plaza, J.; Plaza, A. Scalable recurrent neural network for hyperspectral image classification. *J. Supercomput.* **2020**.

269. Zhang, X.; Sun, Y.; Jiang, K.; Li, C.; Jiao, L.; Zhou, H. Spatial sequential recurrent neural network for hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 1–15.

270. Hao, S.; Wang, W.; Salzmann, M. Geometry-Aware Deep Recurrent Neural Networks for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 2448–2460.

271. Vapnik, V. The support vector method of function estimation. In *Nonlinear Modeling*, Suykens, J.A.K., Ed.; Springer-Science-Business Media: Dordrecht, The Netherlands, 1998.

272. Camps-Valls, G.; Bruzzone, L. Kernel-based methods for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*.

273. Gualtieri, J.A.; Chettri, S.R.; Cromp, R.F.; Johnson, L.F. Support vector machines applied to AVIRIS data. In Proceedings of the Summaries of the Airborne Earth Science Workshop, 1999.

274. Gualtieri, J.A.; Cromp, R.F. Support vector machines for hyperspectral remote sensing classification. In Proceedings of the SPIE, 1998; pp. 221–232.

275. Melgani, F.; Lorenzo, B. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1778–1790.

276. Camps-Valls, G.; Serrano-López, A.J.; Gómez-Chova, L.; Martín-Guerrero, J.D.; Calpe-Maravilla, J.; Moreno, J. Regularized RBF networks for hyperspectral data classification. In Proceedings of the International Conference Image Analysis and Recognition, 2004; pp. 429–436.
277. Mercies, G.; Lennon, M. Support vector machines for hyperspectral image classification with spectral-based kernels. In Proceedings of the International Geoscience and Remote Sensing Symposium, Toulouse, France, 21-25 July, 2003.
278. Camps-Valls, G.; Gomez-Chova, L.; Munoz-Mari, J.; Vila-Frances, J.; Calpe-Maravilla, J. Composite kernels for hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 93–97.
279. Okwuashi, O. Deep support vector machine for hyperspectral image classification. *Pattern Recognit.* **2020**, *103*.
280. Genuer, R.; Poggi, J.M.; Tuleau-Malot, C. Variable selection using random forests. *Pattern Recognit. Lett.* **2010**, *31*, 2225–2236.
281. Svetnik, V.; Liaw, A.; Tong, C.; Culberson, J.C.; Sheridan, R.P.; Feuston, B.P. Random forest: A classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inform. Comput. Sci.* **2003**, *43*, 1947–1958.
282. Amini, S.; Homayouni, S.; Safari, A. Semi-supervised classification of hyperspectral image using random forest algorithm. In Proceedings of the International Geoscience and Remote Sensing Symposium, Quebec, Canada, 13-18 July, 2014.
283. Niemeyer, J.; Rottensteiner, F.; Soergel, U. Classification of urban LiDAR data using conditional random field and random forests. In Proceedings of the Joint Urban Remote Sensing Event, Sao Paulo, Brazil, 21-23 April, 2013; pp. 139–142.
284. Jackson, Q.; Landgrebe, D. Adaptive bayesian contextual classification based on markov random fields. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2454–2463.
285. Izquierdo, A.; Lopez-Guede, J.M. Active Learning for Road Lane Landmark Inventory with Random Forest in Highly Uncontrolled LiDAR Intensity Based Image. In Proceedings of the 15th International Conference on Soft Computing Models in Industrial and Environmental Applications, Burgos, Spain, 16-18 September, 2020.
286. Romero, A.; Gatta, C.; Camps-Valls, G. Unsupervised deep feature extraction for remote sensing image classification. *IEEE* **2015**.
287. Hinton, G.; Salakhutdinov, R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507.
288. Lee, H.; Ekanadham, C.; Ng, A. Sparse deep belief net model for visual area v2. *Adv. Neural Inf. Process. Syst.* **2008**, *20*, 873–880.
289. Masci, J.; Meier, U.; Ciresan, D.; Schmidhuber, J. Stacked convolutional auto-encoders for hierarchical feature extraction. In Proceedings of the International Conference on Artificial Neural Networks, Espoo, Finland, 14-17 June, 2011; pp. 52-59.
290. Kavukcuoglu, K.; Sermanet, P.; Boureau, Y.L.; Gregor, K.; Mathieu, M.; LeCun, Y. Learning convolutional feature hierarchies for visual recognition. *Adv. Neural Inf. Process. Syst.* **2010**.
291. Lecun, Y.B.Y.; Bottou, L.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, 86, 2278-2324.
292. Sermanet, P.; Eigen, D.; Zhang, X.; Mathieu, M.; Fergus, R.; Lecun, Y. OverFeat: integrated recognition, localization and detection using convolutional networks. In Proceedings of the 2nd International Conference on Learning Representations, Canada, Banff, April 14-16, 2014.
293. Chen, Y.; Jiang, H.; Li, C.; Jia, X.; Ghamisi, P. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*.
294. Prokhorov, D.V. Object recognition in 3D LiDAR data with recurrent neural network. In Proceedings of the Computer Vision and Pattern Recognition Workshop, Miami, FL, USA, 20-25 June, 2009.
295. Waske, B.; Benediktsson, J.A.; Arnason, K.; Sveinsson, J.R. Mapping of hyperspectral AVIRIS data using machine-learning algorithms. *Can. J. Remote Sens.* **2009**, *35*, 106–116.
296. Senchuri, R.; Kuras, A.; Burud, I. Machine Learning Methods for Road Edge Detection on Fused Airborne Hyperspectral and LIDAR Data. In Proceedings of the 11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), Amsterdam, The Netherlands, 24–26 March 2021.

Paper B:

Comparison of 2D and 3D semantic segmentation in urban areas using fused hyperspectral and lidar data

Comparison of 2D and 3D semantic segmentation in urban areas using fused hyperspectral and lidar data

Agnieszka Kuras[1], Anna Jenul[1], Maximilian Brell[2], Ingunn Burud[1]

[1]   *Faculty of Science and Technology, Norwegian University of Life Sciences, PB 5003, 1430 Aas, Norway*

[2]   *Helmholtz Center Potsdam, GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany*

Comparison of 2D and 3D semantic segmentation in urban areas using fused hyperspectral and lidar data

Agnieszka Kuras,[a] Anna Jenul,[a] Maximilan Brell[b] and Ingunn Burud[a]

[a]Faculty of Science and Technology, Norwegian University of Life Sciences, PB 5003, 1430 Aas, Norway

[b]Helmholtz Center Potsdam, GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany

Contact:

Agnieszka Kuras: agnieszka.kuras@nmbu.no
https://orcid.org/0000-0002-0541-5614
Anna Jenul: anna.jenul@nmbu.no
https://orcid.org/0000-0002-6919-3483

Maximilian Brell: maximilian.brell@gfz-potsdam.de
https://orcid.org/0000-0002-3759-7483
Ingunn Burud: ingunn.burud@nmbu.no
https://orcid.org/0000-0003-0637-4073

Multisensor data fusion has become a hot topic in the remote sensing research community. This is thanks to significant technological advances and the ability to extract information that would have been challenging with a single sensor. However, sensory enhancement requires advanced analysis that enables deep learning. A framework is designed to effectively fuse hyperspectral and lidar data for semantic segmentation in the urban environment. Our work proposes a method of reducing dimensions by exploring the most representative features from hyperspectral and lidar data and using them for supervised semantic segmentation. In addition, we chose to compare segmentation models based on 2D and 3D convolutional operations with two different model architectures, such as U-Net and ResU-Net. All algorithms have been tested with three loss functions: standard Categorical Cross-Entropy, Focal Loss and a combination of Focal Loss and Jaccard Distance—Focal–Jaccard Loss. Experimental results demonstrated that the 3D segmentation of U-Net and ResU-Net with Focal and Focal–Jaccard Loss functions had significantly improved performance compared to the standard Categorical Cross-Entropy models. The results show a high accuracy score and reflect reality by preserving the complex geometry of the objects.

Keywords: semantic segmentation, 3D segmentation, urban environment, hyperspectral imaging, lidar, hyperspectral lidar fusion, data fusion, deep learning, multisensor fusion, remote sensing

Introduction

An urban environment is a complex mosaic of diverse materials and surfaces constantly undergoing natural and anthropogenic processes resulting in rapid urbanisation.[1,2] This introduces environmental and ecological challenges, such as urban flooding, poor air quality, urban heat island issues and microclimate changes.[3–6] Due to these challenges and the complex and heterogeneous nature of urban areas, a proper advanced urban analysis

is a prerequisite to getting information about urban land surfaces.

Over the last decades, passive and active remote sensing has been widely used in diverse analyses of urban environments, including land cover/material abundance mapping,[7,8] vegetation monitoring,[9–12] urban and suburban use analysis,[13] and three-dimensional urban geometry investigation.[14–16] Such advanced analyses are often achievable considering spectral, spatial and structural properties.[7,17] One sensor is often insufficient to obtain all that information.[18] This can be overcome with multisensory applications such as the fusion of hyperspectral (HS) data with lidar (light detection and ranging) scanning data,[19–22] synthetic aperture radar (SAR)[23] or thermal sensors.[24]

Hyperspectral imaging (HSI) provides detailed spectral information about materials, classifying complex urban structures and effectively distinguishing pixel-based soil, water, vegetation and impervious surfaces.[25] However, the classification is only considered at the near-surface without any elevation information. A land cover classification based on HS data in the spectral domain has been proposed in many studies, applying shallow machine learning methods.[26,27] However, a pure spectral analysis is often insufficient in object-based classification since it does not consider the spatial context—producing salt–pepper noise on the resulting image.[28] An object is a mixture of materials and structures differently sensitive to various atmospheric factors and illumination. This results in distorted results, bias and incomplete analysis.

Lidar scanning in the analysis of urban areas focuses primarily on geometric and textural information of objects and surfaces.[29] Lidar can improve the urban land cover analysis, especially of objects of the same material but differing in geometry, height and structure, such as high and low vegetation.[19,30] Since lidar, as an active sensor, is less sensitive to atmospheric influence and illumination, which challenge HS-based semantic segmentation, it can recompensate with a physical shadow correction for HS data analysis.[31,32]

Therefore, recent technological advances make multisensory fusion possible to exploit the potential of HS and lidar properties for urban analysis (HL-Fusion).[33–38] Such an HL-Fusion can be conducted on different levels focusing on the fusion of various data, products and application purposes based on physical and empirical approaches.[39] An HL-Fusion involves a physical understanding of sensors and awareness of the increased complexity of the analysis process. Therefore, no general process chain for an HL-Fusion exists yet, although deep learning networks

have become promising for different classification and segmentation purposes in recent years. However, classifications based on deep learning models, such as convolutional neural networks (CNN),[40,41] deep residual U-Net (deep ResU-Net)[42] or deep residual net (deep Resnet)[33] are often carried out for HS or lidar data separately.

In urban land cover classification based on HS and lidar data, it is crucial to develop an algorithm that deals effectively with high-dimensional data, considers spectral, spatial and geometrical information simultaneously, and is stable and transferable to other areas with similar problem setups. Since HS data are high-dimensional (3D hyperspectral data cubes), they contain inherent spatial (x, y) and spectral information (λ). One of the main characteristics of an HS image is that it exhibits a strong correlation between the adjacent bands in the spectrum. Therefore, segmentation models that consider spatial features (2D convolutional operations) and the spectral dimension (3D convolutional operations) are believed to improve the performance since they allow for the finding of more patterns (Figure 1).
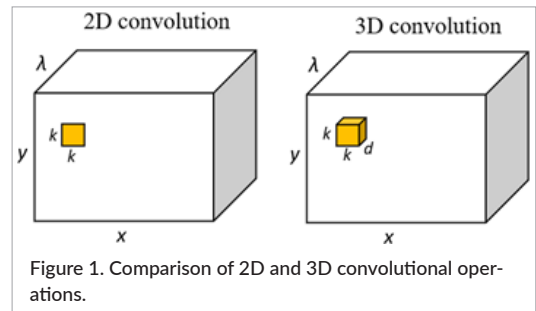

Figure 1. Comparison of 2D and 3D convolutional operations.

The 2D convolutional operation applied to hyperspectral and lidar data is powerful but has some limitations, especially in the spectral domain, since a 2D kernel slides only in the x and y directions. On the other hand, the 3D convolutional operation preserves the spectral information of the input HS data by moving in all three directions: x, y, λ. Furthermore, 3D convolutional operations are designed to exploit spatial continuity further and suppress noisy prediction.[43]

While many different 3D segmentation networks have been proposed in recent studies[43–46] for HL-Fusion, especially in urban environments, little research about 3D convolution-based networks for fused HS and lidar data exists.[47] Inspired by HL-Fusion and deep learning, we propose a method to analyse fused HS and lidar data for urban land cover classification. We compare two model architectures—U-Net and deep ResU-Net (residual deep

U-Net) at 2D and 3D convolutional operation levels applying different loss functions for multiclass segmentation. The two key contributions of this study are:

1) We suggest the design of an HL-Fusion on the feature extraction level for urban land cover classification.
2) We present the first study in HL-Fusion using 3D convolutional operations in U-Net and deep ResU-Net for urban land cover classification.

Theoretical background
Segmentation models
U-Net

The U-Net architecture, introduced in biomedical image segmentation by Ronneberger *et al.*,[48] is based on convolutional neural network layers. The architecture copies low-level features to the corresponding high-level features (Figure 2).

Due to the contracting path propagating between low-level features and high-level semantics in a facilitated way, U-Net takes context into account and hence helps to localise objects precisely.

ResU-Net

A residual neural network is built of stacked residual units.[49] The intuition behind the residual network is to introduce skip connections that prevent accuracy degradation in very deep networks with multiple hidden layers (Figure 3).[50]
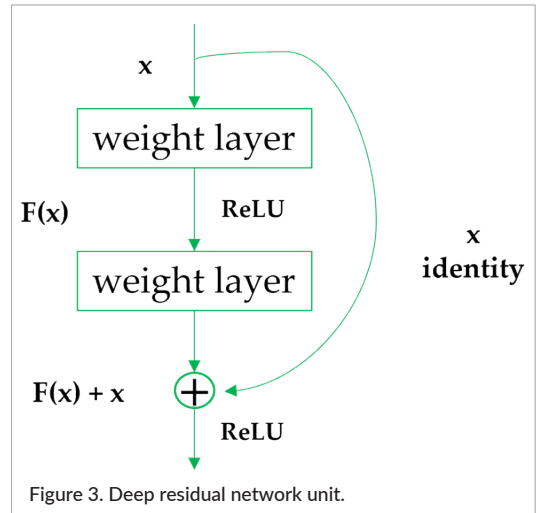

Figure 3. Deep residual network unit.

Skip connections learn identity functions that help follow higher layers perform at least as well as the previous lower layers. Therefore, information loss is reduced and the risk of vanishing gradients during training decreases. The residual unit facilitates the training and skip connections, making the segmentation stable and transferring the information from one layer to another without any information loss. This allows a calculation with fewer parameters.[42] The 3D ResU-Net consists of three parts: encoding, bridge and decoding (Figure 4). Bridge connects both encoding and decoding paths. The encoding part transforms the
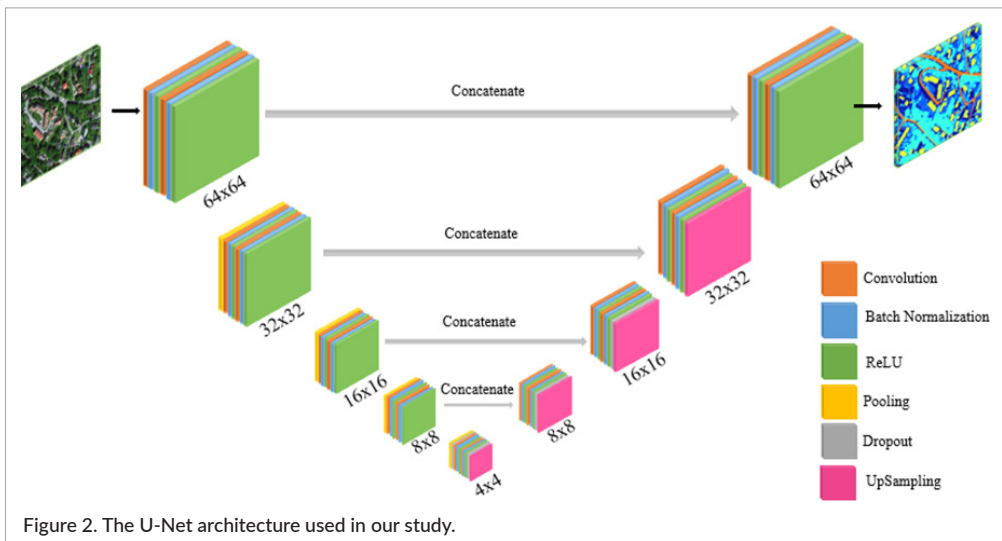

Figure 2. The U-Net architecture used in our study.

input into a compact representation. The decoding part puts the compact representations into a pixel-wise categorisation. All three parts include residual units built of a batch normalisation layer, a ReLU activation layer and a 3D convolutional layer. Both encoding and decoding paths do not contain a pooling operation that downsamples the output to maintain the spectral and spatial dependency in high-dimensional HS and lidar data.

Losses

A loss function must be specified to optimise the weights of the segmentation algorithms. The selection of the loss function is crucial and can significantly impact model optimisation. In Reference 51, Ma compares and categorises different loss functions; Duque *et al.*[52] divide them into two main groups: statistical-based and geometric-based loss functions. While statistical-based loss functions try to minimise the dissimilarity between distributions (derived from Kullback–Leibler divergence), geometric-based functions focus on overlapping regions between prediction and ground truth. Our work compares two statistical-based [Categorical Cross-Entropy (CE) Loss and Focal Loss] and one geometric-based representative (Jaccard Loss). Furthermore, to get the best out of both, we define the Focal–Jaccard Loss, combining Focal and Jaccard losses in an additive way.

Only the (ground truth) positive class contributes to the loss function as we use one-hot encoded target for

training the network. In our notation, p_t is the probability of assigning a pixel t to its true class resulting from a SoftMax activation. The SoftMax function normalises a multidimensional network output to a probability distribution over the model classes. The desired result is that the probability for the true class is significantly higher than for the other classes.

Widely used in classification problems, the CE Loss function is defined as

$$CE(p_t) = -\log(p_t). \tag{1}$$

Hence, the higher p_t is, the better the prediction and the lower the CE Loss is. Although CE is quite common and often the first choice when selecting loss functions, it shows weakness when applied to a model with imbalanced classes.[53] Suppose the dataset is heavily imbalanced, and the number of easy-to-classify samples is much higher than the number of difficult-to-classify samples. In that case, the CE Loss will learn more representations from easy-to-classify samples instead of representations from hard-to-classify samples.

Furthermore, the loss for pixels that are easy to classify (usually the background) is relatively high compared to hard-to-classify samples (such as houses or streets), as shown in Figure 5. As an example, we compare the CE Loss for $p_t = 0.6$ and $p_t = 0.3$. While for $p_t = 0.6$, the correct pixel classification is almost certain, and the CE Loss is still around 0.5 compared to $p_t = 0.3$, where a correct
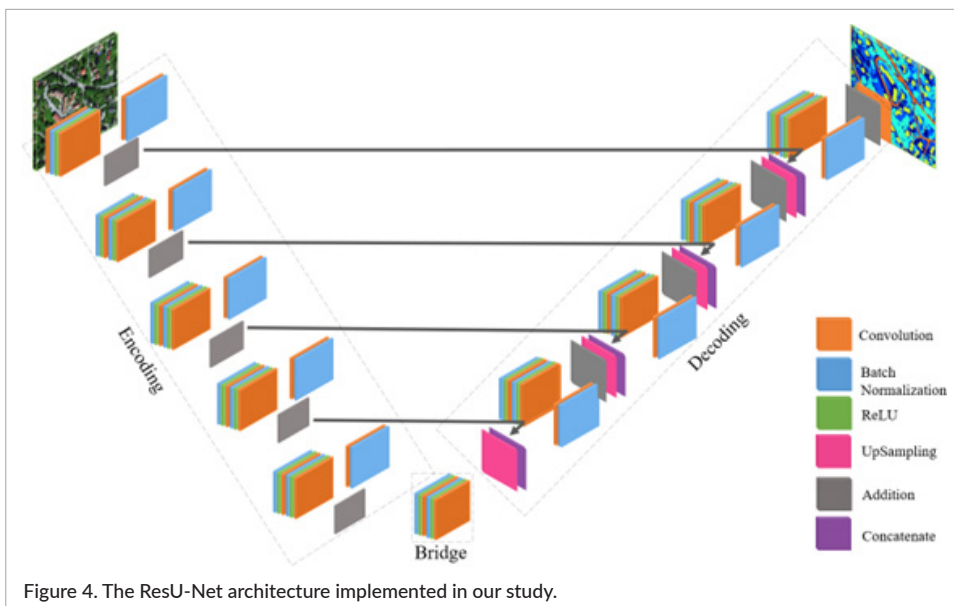

Figure 4. The ResU-Net architecture implemented in our study.

classification is unlikely, but where the loss between 1 and 1.5 is not much higher than 0.5. For this reason, Tsung-Yi Lin *et al.*[53] invented Focal Loss, which extends CE Loss by a multiplicative weight term α (assigning a higher weight to under-represented classes) and a modulating factor $(1 - p_t)^\gamma$, where $\gamma > 0$ is a shape parameter. Focal loss is defined as

$$FL(p_t) = -\alpha_t (1 - p_t)^\gamma \log(p_t). \tag{2}$$

If $\gamma = 0$ and $\alpha_t = 1$, the Focal Loss is equal to CE Loss. Both γ and α_t are hyperparameters that control the learning process. α_t helps weight classes and balance the importance of positive and negative labelled pixels.[53] The higher γ gets, the more Focal Loss focuses on difficult pixels, see Figure 5 where it does not penalise a lot when $p_t \gtrsim 0.6$, i.e. when the certainty in correct prediction increases. Nevertheless, if γ is too high, already weak probabilities would get a very low loss—a good hyperparameter selection is, therefore, essential. Different strategies can be found in the original paper.[53] In our work, we set $\gamma = 2$ and $\alpha_t = 0.25 \forall t$.

The Jaccard Loss relies on the Jaccard Index,[52] measuring the Intersection over the Union of two sets. For our one-hot encoded setup, we compute the Jaccard Loss for a single sample t as

$$1 - \frac{p_t + \varepsilon}{1 + \varepsilon} \tag{3}$$

which is a decreasing linear function with ε representing a small constant. In our case, the function's slope changes

slightly, but in the original definition, it intends to prevent division by zero. By default, we set epsilon, ε, to 1.

The fourth loss in our study is the additive combination of Focal Loss and Jaccard Loss which we refer to Focal–Jaccard Loss (FJL). The loss is given as

$$FJL(p_t) = -\alpha_t (1 - p_t)^\gamma \log(p_t) + 1 - \frac{p_t + \varepsilon}{1 + \varepsilon} \tag{4}$$

In Figure 5, we observe that FJL behaves like Focal Loss for small p_t, whereas it converges towards Jaccard Loss for increasing p_t.

Dimension reduction: spectral unmixing

Airborne-based optical data are characterised by high spectral resolution but correspondingly lower spatial resolution.[54] A single scene pixel usually contains various materials creating a mixed pixel.[55] In order to separate urban materials regardless of pixel size and mixture complexity, spectral unmixing is used to define the amount of a given material in a pixel (abundance map). Various scientific fields have applied spectral unmixing in HS data analysis for decades.[56–60]

Here, abundance maps for each endmember are generated with unsupervised endmember extraction and spectral unmixing, simplifying the analysis and effectively reducing dimensions. Hence, we avoid a curse of dimensionality,[61] which occurs easily when working with limited training data. In such scenarios, the deep learning model has to deal with many features to achieve reliable and accurate results.[19]
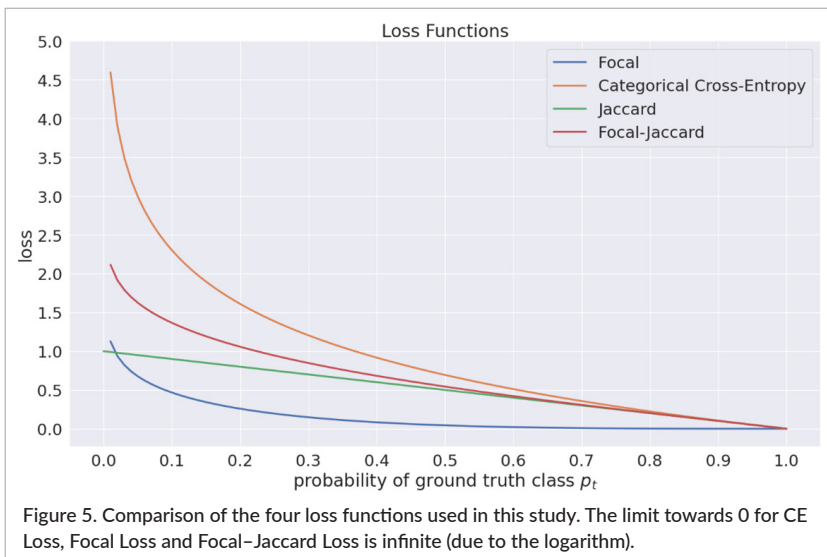

Figure 5. Comparison of the four loss functions used in this study. The limit towards 0 for CE Loss, Focal Loss and Focal–Jaccard Loss is infinite (due to the logarithm).

N-FINDR

The N-FINDR algorithm for endmember extraction was introduced by Winter.[62] This iterative approach aims to find endmembers corresponding to the purest spectra in the scene. N-FINDR assumes that the largest volume formed by an n-simplex with vertices can only be achieved by those purest pixels. The algorithm finds the final endmembers by randomly selecting the pixel set as initial endmembers, further iteratively investigating all spectra in the scene and searching for the largest volume of the simplex. The exact position of the defined endmember is then used to set the original spectrum before the dimension reduction as the final endmember signature. Each pixel in the scene is described as a linear mixture of the defined endmembers using non-negativity-constrained least squares fitting.[63]

Dataset

Terratec AS Company collected HS and lidar data in June 2021 over Baerum municipality in Norway (Figure 6).

The dataset represents a suburban environment with typical urban settlements, infrastructure and diverse vegetation. The airborne-based data were acquired under cloud-free conditions, including HS images and lidar-based 2.5-dimensional rasters. The HS data consist of images from HySpex sensors: VNIR-1800 (0.4–1.0 μm) and SWIR-384 (1–2.5 μm) with 0.3 m and 0.7 m spatial resolution, respectively. The lidar data were acquired using a Riegl VQ-1560i laser scanner, with five pulses per m² and intensity at 1.064 μm relevant for urban environments. The hyperspectral data were georeferenced and orthorectified by the Terratec AS using the PARGE software (Parametric Geocoding and Orthorectification for Airborne Optical Skanner Data).[64] The program corrects the geometry for each HS image pixel using a digital elevation model of the region, GPS position and attitude (ReSe Applications). The geocoded radiance data were converted to reflectance, adjusting illumination levels using ATCOR-4 (Atmospheric and Topographic Correction for airborne imagery). Absorption features associated with H_2O and OH close to bands at 1.4 μm

Figure 6. Study area representing a suburban environment in Høvik, near the capital Oslo in Norway.

and 1.9 µm were excluded from further analyses. The images from all flight lines with residual 176 channels for VNIR and 262 channels for SWIR were mosaicked separately using ENVI software (Environment for Visualizing Images). The noisy bands and outliers were set to bad bands and excluded from analyses resulting in 398 hyperspectral bands. The mosaic layers of VNIR and SWIR were stacked by applying the Savitzky–Golay filter[65] to all spectra. Our study area is located in Høvik with a coordinate extent of 588060, 6641500; 588878, 6641735 WGS 84/UTM zone 32N (Figure 6). The dataset contains the Joint Geospatial Database (FKB) that includes polygons of buildings, roads and other urban surfaces from 2010 to 2019. We carefully controlled FKB-polygons and adapted some modifications for labelling data. The residual classes (high and low vegetation) were extracted automatically.

Proposed method

Semi-automatic label preparation

In our analysis, we used five main classes extracted from the study area: low vegetation, high vegetation, buildings, roads and railway. Our segmentation is limited to these classes to facilitate the analysis and verify and deepen the understanding and optimisation of the method. In the semi-automatic label extraction, we identified high and low vegetation using the advantage of the high amount of relevant HS and lidar features. First, we masked out classes assigned to buildings to exclude potential errors. We applied the Normalised Difference Vegetation Index to HS data,[66] calculating the average for the red $\overline{x}_{RED}$ (660–670 nm) and near infrared band $\overline{x}_{NIR}$ (810–835 nm), respectively:

$$NDVI = \frac{\overline{x}_{NIR} - \overline{x}_{RED}}{\overline{x}_{NIR} + \overline{x}_{RED}} \qquad (5)$$

We concatenated the vegetation mask with raster-based lidar features such as normalised Digital Surface Model, surface roughness and multiple returns. Significant are multiple returns, which distinguish between low vegetation and high vegetation. The assumption is that high vegetation is represented by more returns than low vegetation, usually located at ground points or impervious surfaces through which the laser does not penetrate multiple times. The rest of the classes of interest, such as roofs, railways and roads, were visually validated and extracted from the FKB database. The ground truth map is displayed in Figure 7.

Endmember extraction and abundance maps

Our study separately implemented state-of-art iterative endmember extraction (EA) algorithm N-FINDR[62] for HS and lidar data. We applied the non-negativity-constrained least squares algorithm for abundance map generation. The preprocessed reflectance normalised image in the HS analysis was used to retrieve 27 endmembers. For lidar EA, we built a lidar feature space where the five most relevant raster-based features have been extracted, including slope, the intensity from the first return, multiple returns, normalised Digital Surface Model (nDSM) and point density. All the features have been then normalised separately before EA. The initially extracted endmembers for both HS and lidar data were used to generate abundance maps for each endmember, retrieving 32 abundance maps.

Semantic segmentation

In this study, we applied two semantic segmentation model architectures: a U-Net and a deep residual U-Net that takes full advantage of U-Net architecture and a residual neural network. Both architectures were tested on two- and three-dimensional convolutions. All four model types (2D U-Net, 2D ResU-Net, 3D U-Net and 3D ResU-Net) performed with different loss functions
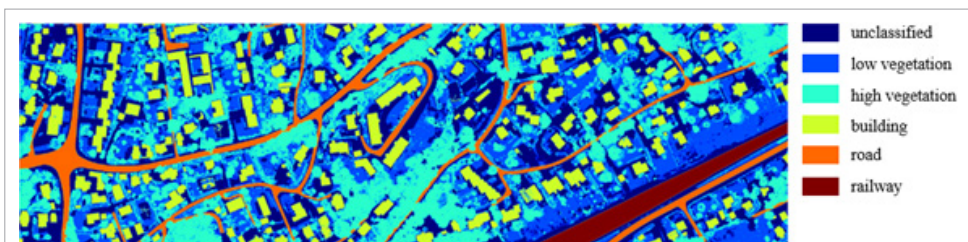

Figure 7. Ground truth map of the study area with five classes.

described above, resulting in 12 segmentation model combinations (Figure 8).

Implementation details

All 12 segmentation models were implemented in Python using Tensorflow, including GPU functionalities.[67] In all models, the maximum number of epochs was set to 500; however, we applied an early stopping function to save time, stop training when the model stopped improving and avoid model overfitting. For both U-Nets and ResU-Nets, we applied the Adam optimiser[68] with the learning rate of 1e–03, epsilon of 1e–07, $\beta_1 = 0.99$ and $\beta_2 = 0.999$. We implemented the Matthews Correlation Coefficient (MCC) to evaluate the algorithm's performance. MCC deals with unbalanced classes calculating the accuracy for each class separately.[69] Compared to F1-score, it takes all true positives, true negatives, false positives and false negatives into account, giving a more reliable performance result. MCC is bounded between –1 and 1, where –1 means that all predictions are incorrect, 1

indicates that all predictions are correct and a score of 0 denotes random predictions. Hence, the higher MCC gets, the better the results are. The input data dimensions for both 2D and 3D convolution-based models can be found in Table 1, splitting the study area into 80% training and 20% testing. The split between training and validation is carefully adapted so that all defined classes are involved equally to ensure that training data contain all known classes (Figure 9).

The input data for the 2D model consist of N number of patches with a size of 64×64 pixels and 32 abundance maps from both HS and lidar data. The ground truth for the 2D model contains five classes with the same N number and size of patches as the input. For the 3D model, the input data have to be expanded by one additional dimension compared to the 2D model. In the ground truth for the 3D model, we must stack the ground truth values 32 times for each abundance map, creating a matrix with the dimension $N \times 64 \times 64 \times 32 \times 5$ (Table 1).
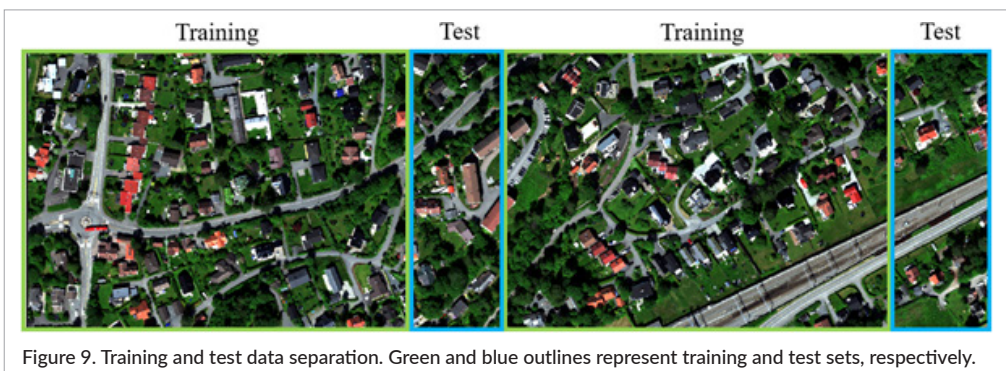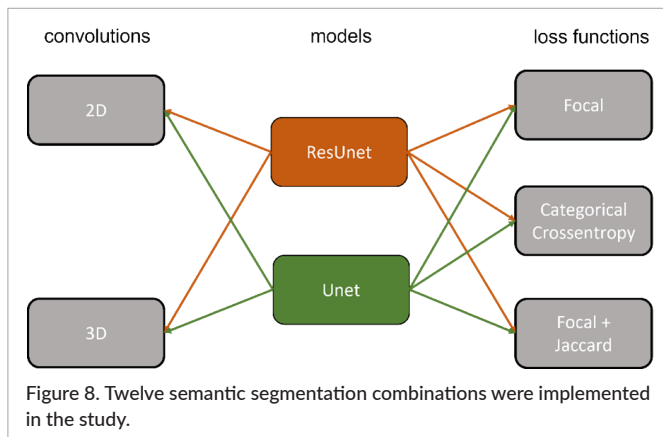

Figure 8. Twelve semantic segmentation combinations were implemented in the study.

Figure 9. Training and test data separation. Green and blue outlines represent training and test sets, respectively.

Table 1. The input data dimensions for both 2D and 3D convolution-based models.

	2D model	3D model
Input	$N \times 64 \times 64 \times 32$	$N \times 64 \times 64 \times 32 \times 1$
Ground truth	$N \times 64 \times 64 \times 5$	$N \times 64 \times 64 \times 32 \times 5$

Experimental results and discussion

Four segmentation methods, including 2D U-Net, 2D ResU-Net, 3D U-Net and 3D ResU-Net, with three different loss functions: CE, Focal Loss and combined Focal–Jaccard Loss, were compared and evaluated. Table 2 presents the experimental results of all segmentation models reporting the MCC metric for all five classes and the overall accuracy of the test set. In this section, the results are presented and discussed in detail.

In order to test and compare 2D and 3D convolutional operation-based segmentation models, we configured U-Net and ResU-Net models for 2D and 3D convolutions and kept the parameters consistent by comparing loss functions such as CE, Focal Loss and Focal–Jaccard Loss. All segmentation maps from the whole scene are shown in Figure 10.

Model evaluation

Based on the results presented in Table 2, it can be noted that the best results for individual classes (shaded fields in grey) were obtained mainly by models based on 3D convolutional operations. Comparing the number of classes best identified in the model, the 3D ResU-Net with combined Focal–Jaccard Loss outperforms, recognising 98 % of high vegetation and 99 % of buildings and railways.

Both 2D and 3D segmentation algorithms within the ResU-Net model architecture have shown that using either Focal Loss or combined Focal–Jaccard Loss significantly outperforms the overall accuracy score (based on MCC) of CE.

Looking at the overall accuracy of each model, we can derive the two best models, such as 2D U-Net with CE and 3D ResU-Net, using Focal Loss. Looking at the total segmentation result for each class, these two models achieved more than 80 % for each class. In the case of the simplest of the proposed models (2D U-Net), CE performed almost as well as the best of the models (3D ResU-Net), obtaining slightly lower accuracy than 3D ResU-Net. The 2D U-Net, despite the simplicity, is more stable and not sensitive to loss functions.

Single class accuracy

Starting the interpretation of the accuracy of a single class, we want to point out that the most challenging class
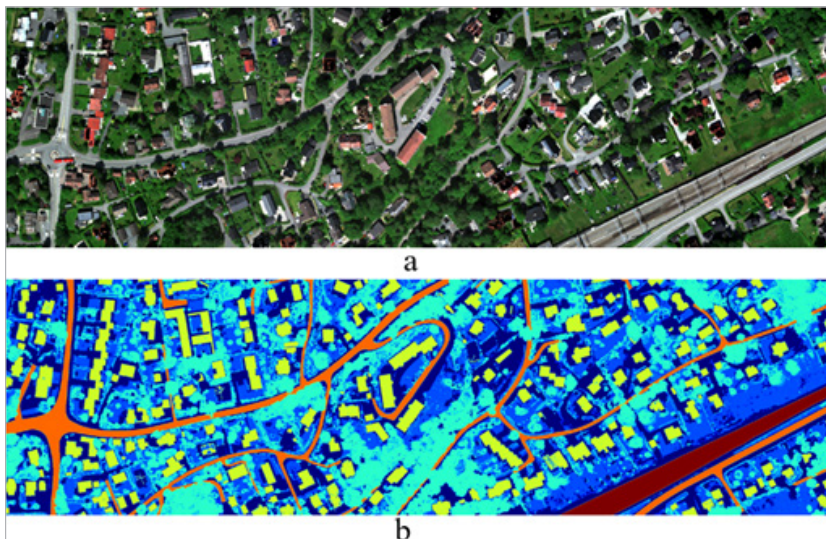

Figure 10. Qualitative comparison of 2D (c–h) and 3D (I–n) convolution-based model in urban land cover classification. a) HS image in RGB colour representation, b) ground truth with five classes such as low vegetation, high vegetation, building, road and railway.
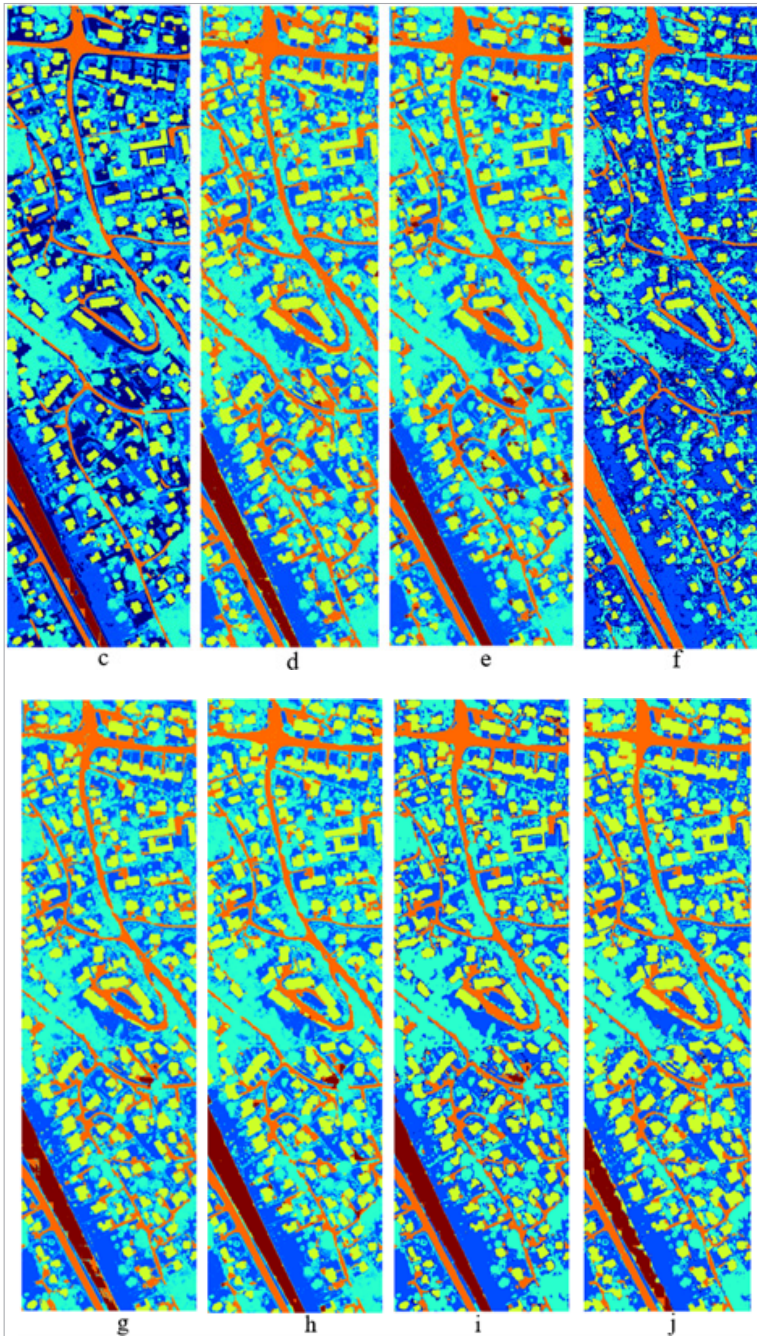
Figure 10. Qualitative comparison of 2D (c–h) and 3D (I–n) convolution-based model in urban land cover classification. c) 2D U-Net CE, d) 2D U-Net Focal Loss, e) 2D U-Net Focal–Jaccard loss, f) 2D ResU-Net CE, g) 2D ResU-Net Focal Loss, h) 2D ResU-Net Focal–Jaccard Loss, i) 3D U-Net CE, j) 3D U-Net Focal Loss.

Figure 10. Qualitative comparison of 2D (c–h) and 3D (I–n) convolution-based model in urban land cover classification. k) 3D U-Net Focal–Jaccard Loss, l) 3D ResU-Net CE, m) 3D ResU-Net Focal Loss, n) 3D ResU-Net Focal–Jaccard Loss.
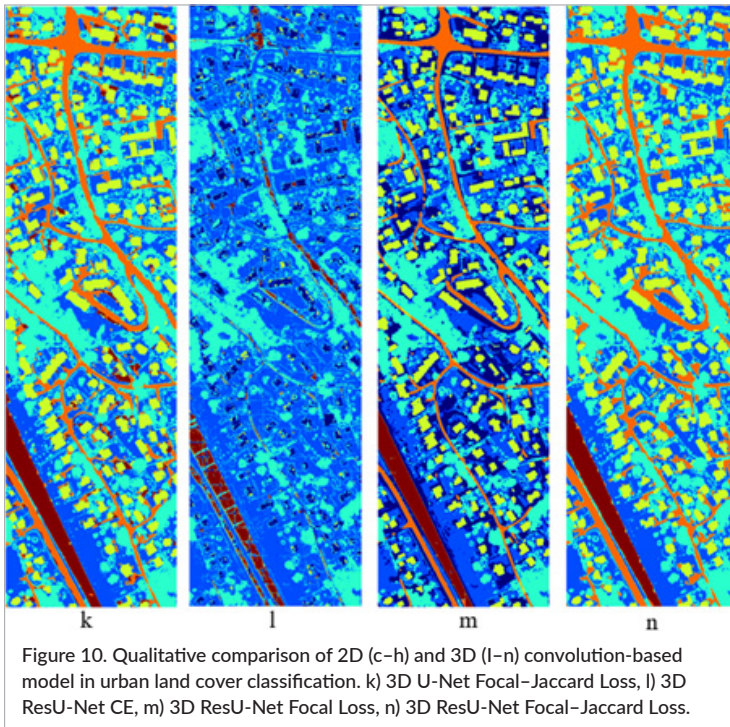
Table 2. Segmentation accuracies obtained by MCC. The grey shaded boxes indicate the highest accuracy score comparing all models for the five classes. Fields shaded in green represent the best accuracy result obtained by comparing loss functions within one model.

	2D U-Net			2D ResU-Net			3D U-Net			3D ResU-Net		
	CE	Focal	Focal–Jaccard	CE	Focal	Focal–Jaccard	CE	Focal	Focal–Jaccard	CE	Focal	Focal–Jaccard
Low vegetation	0.82	0.59	0.64	0.71	0.63	0.62	0.61	0.56	0.63	0.54	0.94	0.61
High vegetation	0.95	0.89	0.91	0.59	0.94	**0.98**	0.88	0.82	0.91	0.82	0.86	**0.98**
Building	0.94	**0.99**	**0.99**	0.89	0.97	**0.99**	0.97	**0.99**	0.98	0.16	0.95	**0.99**
Road	0.90	0.97	0.96	0.83	0.95	0.94	0.96	0.95	0.98	0.21	0.92	0.97
Railway	0.95	**0.99**	**0.99**	0.05	0.88	0.95	**0.99**	0.93	**0.99**	0.10	0.98	**0.99**
Overall accuracy (MCC)	0.84	0.80	0.82	0.62	0.80	0.81	0.78	0.73	0.80	0.31	0.88	0.84

to identify was the low vegetation class, reaching the accuracy of most models from 54% to 71%. However, in the 2D U-Net with CE (Figure 10c) and 3D ResU-Net with Focal Loss (Figure 10m), the low vegetation class was correctly classified, reaching 82% and 94%, respectively. Any supervised classification requires ground truth preparation. In our case, dealing with an airborne-based dataset, our ground truths are extracted directly from the

image (see above), where low and high vegetation are extracted from the laser features. This way speeds up the labelling process but can cause, in particular, sparse low vegetation pixel labelling as "not present".

All other defined classes belonging to impervious objects were classified correctly, even in 2D convolution-based models. We assume that separating these three classes, such as buildings, roads and railways, was

supported by a preprocessing method for HS and lidar data and a proper selection of lidar features based mainly on the review article from Kuras *et al.*[19]

Loss comparison

Considering the choice of loss functions in each model, Focal and combination of Focal–Jaccard Loss outperform because Focal Loss converges to a lower loss faster than CE, focusing more on classes that are difficult to predict (Figure 5). Measurements on Jaccard distance take more time than Focal Loss and need to be close to the truth before the loss converges. These losses fill each other out and take advantage of both functions: handling extremely imbalanced data and recognising objects.

2D and 3D models

Several important aspects can be noted by visually comparing 2D and 3D models (Figure 10). First, sometimes in 2D convolution-based segmentation models, despite their comparable accuracy score to 3D models, so-called edge effects appear which are the ends of the 64×64 patches into which the whole scene has been divided before feeding into the segmentation algorithm (Figure 11).
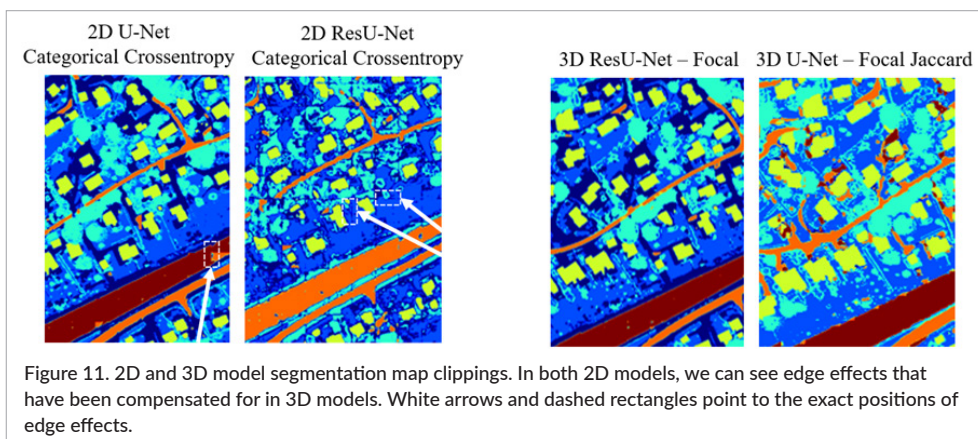
This problem is well-known in 2D patch-based image segmentation processes and has been reported in several studies.[70,71] In order to avoid additional post-processing steps and the potential generation of other artefacts to the final segmentation map, we compared 3D segmentation models in which the problem of edge effects has been mitigated. The 3D ResU-Net with CE (Figure 10l) is an exception, where the parameters were probably chosen incorrectly. However, since we compare the same 2D and 3D model architectures, we necessarily wanted to present all the results.

Another aspect that needs to be discussed is geometric accuracy. In Figure 10 and Figure 11, we can see (especially in the building class) that 2D and 3D models do not perform well with objects with irregular shapes. Often, we can see inaccurate blurred contours and, in some places, complete disruption of the geometry. However, in the 2D U-Net with CE and the 3D ResU-Net with Focal Loss, the geometric of identified objects/surfaces are adequate or close to reality and ground truth.

It must be noted that an important issue appeared in the attempt to classify the "unknown" class. Only in the case of the 2D U-Net with CE and the 3D ResU-Net with Focal Loss undefined pixels (no ground truth) were correctly identified as "unknown", while the rest of the segmentation models assigned a class to each pixel, creating many false positives. Identification of unknown pixels is crucial, especially in a heterogeneous complex urban environment where our selected scene consists of more than the five defined classes. This "unknown" class has been assigned to those pixels where no defined class has reached 50% in the prediction.

Furthermore, the proper input preparation and the reduction of dimensions through endmember extraction allowed us to skip statistical transformation methods, such as Principal Component Analysis (PCA). Due to its purely statistical nature, PCA loses spectral information since the spectral positions are not considered in the calculations.[72] Hence the extracted (spectral) endmembers can be used for further analysis or to generate local spectral libraries.

Figure 11. 2D and 3D model segmentation map clippings. In both 2D models, we can see edge effects that have been compensated for in 3D models. White arrows and dashed rectangles point to the exact positions of edge effects.

The last but not least aspect relates to time consumption comparing 2D and 3D segmentation models. Due to the requirements of advanced calculations in 3D convolutional operations, 3D models took about 20 times longer than 2D models. For this reason, with a simple segmentation of high-dimensional HS and lidar data, simple 2D models seem to be sufficient. However, we are convinced that analysis based on much more advanced segmentation, intraclass segmentation, as well as identification of materials and their properties in an urban environment will require the choice of 3D convolutional operations despite the time consumption.

Conclusion and future perspectives

This paper explores diverse deep learning models with different parameters and convolutional dimensions for urban land cover classification using fused HS and lidar data. In particular, four deep learning models have been compared, 2D-U-Net (Figure 10c–e), 2D-ResU-Net (Figure 10f–h), 3D-U-Net (Figure 10i–k) and 3D-ResU-Net (Figure 10l–n). All models have been trained with three different loss functions: state-of-art CE, Focal Loss, and a combination of Focal Loss and Jaccard Loss (Focal–Jaccard Loss).

As a whole, we can conclude that segmentation of both U-Net and ResU-Net performed very well, especially implementing Focal and Focal–Jaccard Loss functions. However, an important factor to consider is the choice of the model architecture and the proper selection of parameters. Despite the out-performance of the 3D U-Net and Res-U-Net models, the training often requires more training data than analysing shallow machine learning or deep 2D convolution-based algorithms. Alternatively, to save time, reduce computational resources and limit the need for a large amount of training data, we will explore the potential of hybrid 2D–3D models and carry out analyses based on ensemble learning where several models are combined. Thus, it is possible to reduce the number of parameters in the deep learning model and, at the same time, use the information from the 3D convolution-based algorithm.

Another important factor to consider when working with image segmentation on HS and lidar data is the correct labelling of the ground truth. Due to the fuzzy class assignment of some pixels responding from semi-automatic label preparation, the network may try to learn incorrect patterns. Further, a single pixel can contain more than one class, but the model assumes that the pixel contains only a single class, which also hinders the learning process. Hence, in future research, we will focus on unsupervised segmentation to avoid other errors in creating ground truths and develop a method to identify unknown classes, such as testing open set recognition for deep learning algorithms, e.g. for change detection based on HL-Fusion data.

Funding

References

1. X.-P. Song, J.O. Sexton, C. Huang, S. Channan and J.R. Townshend, "Characterizing the magnitude, timing and duration of urban growth from time series of Landsat-based estimates of impervious cover", *Remote Sens. Environ.* **175,** 1–13 (2016). https://doi.org/10.1016/j.rse.2015.12.027

2. J.A. Hepinstall, M. Alberti and J.M. Marzluff, "Predicting land cover change and avian community responses in rapidly urbanizing environments", *Landsc. Ecol.* **23,** 1257–1276 (2008). https://doi.org/10.1007/s10980-008-9296-6

3. Y.B. Chen, H.L. Zhou, H. Zhang, G.M. Du and J.H. Zhou, "Urban flood risk warning under rapid urbanization", *Environ. Res.* **139,** 3–10 (2015). https://doi.org/10.1016/j.envres.2015.02.028

4. F. Chen, H. Kusaka, R. Bornstein, J. Ching, C.S.B. Grimmond, S. Grossman-Clarke, T. Loridan, K.W. Manning, A. Martilli, S. Miao, D. Sailor, F.P. Salamanca, H. Taha, M. Tewari, X. Wang, A.A. Wyszogrodzki and C. Zhang, "The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems", *Int. J. Climatol.* **31(2),** 273–288 (2011). https://doi.org/10.1002/joc.2158

5. J.H. Lee and K.B. Woong, "Characterization of urban stormwater runoff", *Water Res.* **34(6),** 1773–1780 (2000). https://doi.org/10.1016/S0043-1354(99)00325-5

6. Y. Qian, W. Zhou, W. Yu and S.T.A. Pickett, "Quantifying spatiotemporal pattern of urban

greenspace: new insights from high resolution data", *Landsc. Ecol.* **30,** 1165–1173 (2015). https://doi.org/10.1007/s10980-015-0195-3

7. M. Herold, M.E. Gardner and D.A. Roberts, "Spectral resolution requirements for mapping urban areas", *IEEE Trans. Geosci. Remote Sens.* **41(9),** 1907–1919 (2003). https://doi.org/10.1109/TGRS.2003.815238

8. S. Roessner, K. Segl, U. Heiden and H. Kaufmann, "Automated differentiation of urban surfaces based on airborne hyperspectral imagery", *IEEE Trans. Geosci. Remote Sens.* **39(7),** 1525–1532 (2001). https://doi.org/10.1109/36.934082

9. Y.-C. Lin, C. Lin, M.-D. Tsai and C.-L. Lin, "Object-based analysis of LiDAR geometric features for vegetation detection in shaded area", *XXIII ISPRS Congress*. Prague, Czech Republic (2016). https://doi.org/10.5194/isprsarchives-XLI-B1-43-2016

10. M. Rutzinger, B. Höfle and N. Pfeifer, "Detection of high urban vegetation with airborne laser scanning data", *Forestsast*. Montpellier, France (2007).

11. M.K. Ridd, "Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities", *Int. J. Remote Sens.* **16,** 2165–2185 (1995). https://doi.org/10.1080/01431169508954549

12. M. Alonzo, B. Bookhagen, J.P. McFadden, A. Sun and D.A. Roberts, "Mapping urban forest leaf area index with airborne lidar using penetration metrics and allometry", *Remote Sens. Environ.* **162,** 141–153 (2015). https://doi.org/10.1016/j.rse.2015.02.025

13. C. Jürgens, "Urban and suburban growth assessment with remote sensing", *OICC 7th International Seminar on GIS Applications in Planning and Sustainable Development*. 13–15 February, Cairo, Egypt (2001).

14. M. Batty and P. Longley, *Fractal Cities: A Geometry of Form and Function*. Academic Press, London and San Diego (1994).

15. M.A. Stretton, W. Morrison, R.J. Hogan and S. Grimmond, "Evaluation of the SPARTACUS-urban radiation model for vertically resolved shortwave radiation in urban areas", *Bound.-Lay. Meteorol.* **184,** 301–331 (2022). https://doi.org/10.1007/s10546-022-00706-9

16. V.S. Saravanarajan, R.C. Chen and L.S. Chen, "Geometric feature learning network for detecting the objects in urban streets", *SSRN*, preprint (2022). https://doi.org/10.2139/ssrn.4218483

17. E. Ben-Dor, N. Levin and H. Saaroni, "A spectral based recognition of the urban environment using the visible and near-infrared spectral region (0.4–1.1 μm). A case study over Tel-Aviv, Israel", *Int. J. Remote Sens.* **22(11),** 2193–2218 (2001). https://doi.org/10.1080/01431160117759

18. P. Gamba, F. Dell'Acqua and B.V. Dasarathy, "Urban remote sensing using multiple data sets: Past, present, and future", *Inf. Fusion* **6,** 319–326 (2005). https://doi.org/10.1016/j.inffus.2005.02.007

19. A. Kuras, M. Brell, J. Rizzi and I. Burud, "Hyperspectral and lidar data applied to the urban land cover machine learning and neural-network-based classification: a review", *Remote Sens.* **13(17),** 3393 (2021). https://doi.org/10.3390/rs13173393

20. R. Senchuri, A. Kuras and I. Burud, "Machine learning methods for road edge detection on fused airborne hyperspectral and lidar data", *2021 11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2021*. pp. 1–5 (2021). https://doi.org/10.1109/WHISPERS52202.2021.9484007

21. F. Priem and F. Canters, "Synergistic use of LiDAR and APEX hyperspectral data for high-resolution urban land cover mapping", *Remote Sens.* **8(10),** 787 (2016). https://doi.org/10.3390/rs8100787

22. M. Alonzo, B. Bookhagen and D.A. Roberts, "Urban tree species mapping using hyperspectral and lidar data fusion", *Remote Sens. Environ.* **148,** 70–83 (2014). https://doi.org/10.1016/j.rse.2014.03.018

23. D. Borghys, M. Shimoni and C. Perneel, "Change detection in urban scenes by fusion of SAR and hyperspectral data", *Proc. SPIE 6749, Remote Sensing for Environmental Monitoring, GIS Applications, and Geology VII* 67490R (2007). https://doi.org/10.1117/12.738767

24. K. Segl, S. Roessner, U. Heiden and H. Kaufmann, "Fusion of spectral and shape features for identification of urban surface cover types using reflective and thermal hyperspectral data", *ISPRS J. Photogramm. Remote Sens.* **58(1–2),** 99–112 (2003). https://doi.org/10.1016/S0924-2716(03)00020-0

25. J.M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader and J. Chanussot, "Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches", *IEEE J. Select. Topics Appl. Earth Observ. Remote Sens.* **5(2),** 354–379 (2012). https://doi.org/10.1109/JSTARS.2012.2194696

26. G. Hegde, J.M. Ahamed, R. Hebbar and U. Raj, "Urban land cover classification using hyperspectral data", *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **XL-8,** 751–754 (2014). https://doi.org/10.5194/isprsarchives-XL-8-751-2014

27. S. Amini, S. Homayouni and A. Safari, "Semi-supervised classification of hyperspectral image using random forest algorithm", *2014 IEEE Geoscience and Remote Sensing Symposium.* pp. 2866–2869 (2014). https://doi.org/10.1109/IGARSS.2014.6947074

28. X. Huang and L. Zhang, "An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery", *IEEE Trans. Geosci. Remote Sens.* **51(1),** 257–272 (2013). https://doi.org/10.1109/TGRS.2012.2202912

29. A. Serna and B. Marcotegui, "Urban accessibility diagnosis from mobile laser scanning data", *ISPRS J. Photogramm. Remote Sens.* **84,** 23–32 (2013). https://doi.org/10.1016/j.isprsjprs.2013.07.001

30. M. Dalponte, L. Bruzzone and D. Gianelle, "Fusion of hyperspectral and LiDAR remote sensing data for classification of complex forest areas", *IEEE Trans. Geosci. Remote Sens.* **46(5),** 1416–1427 (2008). https://doi.org/10.1109/TGRS.2008.916480

31. M. Brell, K. Segl, L. Guanter and B. Bookhagen, "Hyperspectral and Lidar intensity data fusion: a framework for the rigorous correction of illumination, anisotropic effects, and cross calibration", *IEEE Trans. Geosci. Remote Sens.* **55(5),** 2799–2810 (2017). https://doi.org/10.1109/TGRS.2017.2654516

32. M. Brell, C. Rogass, K. Segl, B. Bookhagen and L. Guanter, "Improving sensor fusion: a parametric method for the geometric coalignment of airborne hyperspectral and LiDAR data", *IEEE Trans. Geosci. Remote Sens.* **54(6),** 3460–3474 (2016). https://doi.org/10.1109/TGRS.2016.2518930

33. C. Ge, Q. Du, W. Sun, K. Wang, J. Li and Y. Li, "Deep residual network-based fusion framework for hyperspectral and LiDAR data", *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **14,** 2458–2472 (2021). https://doi.org/10.1109/JSTARS.2021.3054392

34. D. Hong, L. Gao, R. Hang, B. Zhang and J. Chanussot, "Deep encoder-decoder networks for classification of hyperspectral and LiDAR data", *IEEE Geosci. Remote Sens. Lett.* **99,** 1–5 (2020). https://doi.org/10.1109/LGRS.2020.3017414

35. R. Hang, Y. Li, P. Ghamisi, D. Hong, G. Xia and Q. Liu, "Classification of hyperspectral and LiDAR data using coupled CNNs", *IEEE Trans. Geosci. Remote Sens.* **68(7),** 4939–4950 (2020). https://doi.org/10.1109/TGRS.2020.2969024

36. S.Y. Sadjadi and S. Parsian, "Combining hyperspectral and LiDAR data for building extraction using machine learning technique", *Int. J. Comput.* **2,** 88–93 (2017).

37. Q. Feng, Q. Zhu, J. Yang and B. Li, "Multisource hyperspectral and LiDAR data fusion for urban land-use mapping based on a modified two-branch convolutional neural network", *ISPRS Int. J. Geoinf.* **8(1),** 28 (2019). https://doi.org/10.3390/ijgi8010028

38. H. Li, P. Ghamisi, U. Soergel and X.X. Zhu, "Hyperspectral and LiDAR fusion using deep three-stream convolutional neural networks", *Remote Sens.* **10(10),** 1649 (2018). https://doi.org/10.3390/rs10101649

39. H. Torabzadeh, F. Morsdorf and M.E. Schaepman, "Fusion of imaging spectroscopy and airborne laser scanning data for characterization of forest ecosystems", *ISPRS J. Photogramm. Remote Sens.* **97,** 25–35 (2014). https://doi.org/10.1016/j.isprsjprs.2014.08.001

40. J. Yin, C. Qi, Q. Chen and J. Qu, "Spatial-spectral network for hyperspectral image classification: a 3-D CNN and Bi-LSTM framework", *Remote Sens.* **13(12),** 2353 (2021). https://doi.org/10.3390/rs13122353

41. H. Gong, Q. Li, C. Li, H. Dai, Z. He, W. Wang, H. Li, F. Han, A. Tuniyazi and T. Mu, "Multiscale information fusion for hyperspectral image classification based on hybrid 2D-3D CNN", *Remote Sens.* **13(12),** 2268 (2021). https://doi.org/10.3390/rs13122268

42. Z. Zhang, Q. Liu and Y. Wang, "Road extraction by deep residual U-net", *IEEE Geosci. Remote Sens. Lett.* **15(5),** 749–753 (2018). https://doi.org/10.1109/LGRS.2018.2802944

43. C. Shi and C.M. Pun, "Superpixel-based 3D deep neural networks for hyperspectral image classification", *Pattern Recognit.* **74,** 600–616 (2018). https://doi.org/10.1016/j.patcog.2017.09.007

44. C. Yu, R. Han, M. Song, C. Liu and C.I. Chang, "A simplified 2D-3D CNN architecture for hyperspectral image classification based on spatial-spectral fusion", *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **13,** 2485–2501 (2020). https://doi.org/10.1109/JSTARS.2020.2983224

45. M. Rao, P. Tang and Z. Zhang, "A developed Siamese CNN with 3D adaptive spatial-spectral pyramid pooling for hyperspectral image classification", *Remote Sens.* **12(12),** 1964 (2020). https://doi.org/10.3390/rs12121964

46. Q. Li, Q. Wang and X. Li, "Mixed 2D/3D convolutional network for hyperspectral image super-resolution", *Remote Sens.* **12(10),** 1660 (2020). https://doi.org/10.3390/rs12101660

47. M.K.K. Singh, S. Mohan and B. Kumar, "Fusion of hyperspectral and LiDAR data using sparse stacked autoencoder for land cover classification with 3D-2D convolutional neural network", *J. Appl. Remote Sens.* **16(3),** 034523 (2022). https://doi.org/10.1117/1.JRS.16.034523

48. O. Ronneberger, P. Fischer and T. Brox, "U-Net: convolutional networks for biomedical image segmentation", in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Ed by N. Navab, J. Hornegger, W. Wells and A. Frangi. Lecture Notes in Computer Science. Springer, Vol. 9351, pp. 234–241 (2015). https://doi.org/10.1007/978-3-319-24574-4_28

49. K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition", *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* pp. 770–778 (2016). https://doi.org/10.1109/CVPR.2016.90

50. K. He and J. Sun, "Convolutional neural networks at constrained time cost", *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* pp. 5353–5360 (2015). https://doi.org/10.1109/CVPR.2015.7299173

51. J. Ma, "Segmentation loss odyssey", *arxiv:2005* 13449v1 (2020). https://doi.org/10.48550/arXiv.2005.13449

52. D. Duque-Arias, S. Valesco-Forero, J.E. Deschaud, F. Goulette, A. Serna, E. Decenciere and B. Marcotegui, "On power Jaccard losses for semantic segmentation", *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications.* Vol. 5, pp. 561–568 (2021). https://doi.org/10.5220/0010304005610568

53. T.Y. Lin, P. Goyal, R. Girshick, K. He and P. Dollar, "Focal loss for dense object detection", *2017 IEEE International Conference on Computer Vision (ICCV).* pp. 2999–3007 (2017). https://doi.org/10.1109/ICCV.2017.324

54. M. Hasanlou and S.T. Seydi, "Hyperspectral change detection: an experimental comparative study", *Int. J. Remote Sens.* **39(20),** 7029–7083 (2018). https://doi.org/10.1080/01431161.2018.1466079

55. A. Ertürk, M.K. Güllü, D. Çeşmeci, D. Gerçek and S. Ertürk, "Spatial resolution enhancement of hyperspectral images using unmixing and binary particle swarm optimization", *IEEE Geosci. Remote Sens. Lett.* **11(12),** 2100–2104 (2014). https://doi.org/10.1109/LGRS.2014.2320135

56. A. Ertürk, M.-D. Iordache and A. Plaza, "Sparse unmixing with dictionary pruning for hyperspectral change detection", *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **20(1),** 321–330 (2017). https://doi.org/10.1109/JSTARS.2016.2606514

57. S. Liu, L. Bruzzone, F. Bovolo and D. Peijun, "Unsupervised multitemporal spectral unmixing for detecting multiple changes in hyperspectral images", *IEEE Trans. Geosci. Remote Sens.* **54(5),** 2733–2748 (2016). https://doi.org/10.1109/TGRS.2015.2505183

58. Y. Fang, Y. Wang, L. Xu, R. Zhuo, A. Wong and D.A. Clausi, "BCUN: Bayesian fully convolutional neural network for hyperspectral spectral unmixing", *IEEE Trans. Geosci. Remote Sens.* **60,** 5523714 (2022). https://doi.org/10.1109/TGRS.2022.3151004

59. C.V.S.S.M. Kumar, S.S. Jha, R.R. Nidamanuri and B.K. Dadhwal, "Benchmark studies on pixel-level spectral unmixing of multi-resolution hyperspectral imagery", *Int. J. Remote Sens.* **43(4),** 1451–1484 (2022). https://doi.org/10.1080/01431161.2022.2040755

60. A. Gomez-Sanchez, M. Marro,M. Marsal, S. Zacchetti, R.R. de Oliveira, P. Loza-Alvarez and A. de Juan, "Linear unmixing protocol for hyperspectral image fusion analysis applied to a case study of vegetal tissues", *Sci. Rep.* **11,** 18665 (2021). https://doi.org/10.1038/s41598-021-98000-0

61. G.F. Hughes, "On the mean accuracy of statistical pattern recognizers", *IEEE Trans. Inf. Theory* **14,** 55–63 (1968). https://doi.org/10.1109/TIT.1968.1054102

62. M.E. Winter, "N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data", *Proc. SPIE 3753, Imaging Spectrometry V* **3753,** 266–275 (1999). https://doi.org/10.1117/12.366289

63. R. Bro and S. De Jong, "A fast non-negativity-constrained least squares algorithm", *J. Chemometr.* **11(5),** 393–401 (1997). https://doi.org/10.1002/

(SICI)1099-128X(199709/10)11:5<393::AID-CEM483>3.0.CO;2-L

64. D. Schläpfer and R. Richter, "Geo-atmospheric processing of airborne imaging spectrometry data. Part 1: Parametric orthorectification", *Int. J. Remote Sens.* **23(13),** 2609–2630 (2002). https://doi.org/10.1080/01431160110115825

65. C. Ruffin and R.L. King, "The analysis of hyperspectral data using Savitzky-Golay filtering-Theoretical basis (part 1)", *IEEE 1999 International Geoscience and Remote Sensing Symposium IGARSS'99* **2,** 756–758 (1999). https://doi.org/10.1109/IGARSS.1999.774430

66. P.S. Thenkabail, R.B. Smith and E. De Pauw, "Hyperspectral vegetation indices and their relationships with agricultural crop characteristics", *Remote Sens. Environ.* **71(2),** 158–182 (2000). https://doi.org/10.1016/S0034-4257(99)00067-X

67. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu and X. Zheng, "Tensorflow: large-scale machine learning on heterogeneous distributed systems", arXiv:160304467 (2016). https://doi.org/10.48550/arXiv.1603.04467

68. E. Adam, O. Mutanga and D. Rugege, "Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review", *Wetl. Ecol. Manag.* **18,** 281–296 (2010). https://doi.org/10.1007/s11273-009-9169-z

69. B.W. Matthews, "Comparison of the predicted and observed secondary structure of t4 phage lysozyme", *Biochim. Biophys. Acta-Protein Structure* **405(2),** 442–451 (1975). https://doi.org/10.1016/0005-2795(75)90109-9

70. N. Pielawski and C. Wählby, "Introducing Hann windows for reducing edge-effects in patch-based image segmentation", *PLoS One* **15(3),** e0229839 (2020). https://doi.org/10.1371/journal.pone.0229839

71. N.G. Kurale and M. Vaidya, "Retinal hemorrhage detection using splat segmentation of retinal fundus images", *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, pp. 364–375 (2017). https://doi.org/10.1109/ICCUBEA.2017.8463939

72. H.A. Khan, M.M. Khan, K. Khurshid and J. Chanussot, "Saliency based visualization of hyperspectral images", *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1096–1099 (2015). https://doi.org/10.1109/IGARSS.2015.7325961

Paper C:

Multitemporal Feature-Level Fusion on Hyperspectral and LiDAR Data in the Urban Environment

Multitemporal Feature-Level Fusion on Hyperspectral and LiDAR Data in the Urban Environment

Agnieszka Kuras [1], Maximilian Brell [2], Kristian Hovde Liland [1], Ingunn Burud [1]

[1]   *Faculty of Science and Technology, Norwegian University of Life Sciences, PB 5003, 1430 Aas, Norway*

[2]   *Helmholtz Center Potsdam, GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany*

Article 1

Multitemporal Feature-Level Fusion on Hyperspectral and LiDAR Data in the Urban Environment

Agnieszka Kuras [1,*], Maximilian Brell [2], Kristian Hovde Liland [1], and Ingunn Burud [1] 4

[1] Faculty of Science and Technology, Norwegian University of Life Sciences, PB 5003, 1430 Aas, Norway; 5
ingunn.burud@nmbu.no 6
[2] Helmholtz Center Potsdam, GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, 7
Germany; maximilian.brell@gfz-potsdam.de 8
* Correspondence: agnieszka.kuras@nmbu.no 9

Abstract: Technological innovations and advanced multidisciplinary research increase the demand 10
for multisensor data fusion in Earth observations. Such fusion has great potential, especially in the 11
remote sensing field. One sensor is often insufficient in analyzing urban environments to obtain 12
comprehensive results. Inspired by the capabilities of hyperspectral and LiDAR data in multisensor 13
data fusion at the feature level, we present a novel approach to the multitemporal analysis of urban 14
land cover in a case study in Høvik, Norway. Our generic workflow is based on bitemporal datasets; 15
however, it is designed to include datasets from other years. Our framework extracts representative 16
endmembers in an unsupervised way, retrieves abundance maps fed into segmentation algorithms, 17
and detects the main urban land cover classes by implementing 2D ResU-Net for segmentation 18
without parameter regularizations and with effective optimization. Such segmentation optimization 19
is based on updating initial features and providing them for a second iteration of segmentation. We 20
compared segmentation optimization models with and without data augmentation, achieving up 21
to 11 % better accuracy after segmentation optimization. In addition, a stable spectral library is au- 22
tomatically generated for each land cover class, allowing local database extension. The main product 23
of the multitemporal analysis is a map update, effectively detecting detailed changes in land cover 24
classes. 25

Keywords: multisensory data fusion, feature-level fusion, hyperspectral imaging lidar, urban envi- 26
ronment, urban remote sensing 27
28

1. Introduction 29

Urban surface types are a mix of complex materials and surfaces, such as low, mid- 30
dle, and high vegetation, non-vegetated pervious surfaces, and partially and fully imper- 31
vious surfaces, including asphalt, concrete, and various roofing systems [1]. These mate- 32
rials and surfaces undergo natural and anthropogenic processes, constantly increasing 33
urban heterogeneity [2]. This diversity of urban land cover is additionally characterized 34
by high-frequent changes and complex transitions [3,4] due to the growing urban popu- 35
lation contributing to a concomitant increase in environmental problems [5]. In order to 36
effectively monitor the highly dynamic urban environment, appropriate technologies and 37
methods are needed to cope with such change analysis within the urban environment. 38
Active and passive remote sensing has been widely used in urban land cover map- 39
ping and monitoring in recent decades [6-9]. Hyperspectral (HS) data at the airborne scale 40
have gained particular attention, identifying materials effectively based on their physical 41
and chemical properties [10,11]. HS imaging has increasingly become a valuable tool for 42
multitemporal analysis, such as change detection in urban areas. 43
Multitemporal analysis of HS data compares materials, material conditions, stability, 44
degradation, pollution, alteration, and anthropological and atmospheric changes. HS- 45

based change detection uses rich spectral information distinguishing materials and fine spectral changes [12]. The high amount of spectral features enables effective real-time detection of changing areas. However, the information about the changes is often contained in different bands simultaneously, complicating the HS analysis. In airborne-based HS data, the spatial resolution is significantly higher than in satellite-based images. This results in high spectral complexity of objects of similar materials showing similar spectral responses [13]. Any change detection technique must deal with high dimensionality, computational cost, and limited data, including ground truth data [14].

One of the limitations in high-resolution airborne-based HS change detection is pixel-based classification. Such a classification requires assumptions that neighbor pixels are independent of each other and that radiometric properties of multitemporal images are identical. However, these assumptions are not valid in the urban environment due to the heterogeneity of the urban surfaces, different atmospheric conditions during data acquisition, and sensor geometry [15]. Due to the potentially miscellaneous spectral behavior of urban surfaces and adjacent pixel dependency, semantic meaning and spatial context analysis are critical. Such spatial-context information is included when extracting textures, calculating morphological filters [16-19], using adaptive pixel neighborhood [20], applying contextual Support Vector Machines [21], Markov Random Fields (MRF) [22,23], Convolutional Neural Networks (CNN) combined with MRF [24], 3D CNN extracting spectral and spatial information simultaneously [25]. However, another possibility to compensate for this problem and to complement multitemporal HS analysis is multisensor data fusion with LiDAR (HL-Fusion) [8,26-30].

The use of LiDAR in multitemporal analysis focuses mainly on structural and textural changes, e.g., canopy gaps [31] and single-tree levels in forestry applications [32,33], and mining subsidence [34]. Applications of LiDAR data have been mainly limited to analysis based on data acquisition from a single time (single-data analysis). This is mainly due to the lack of a multitemporal database and technical limitations such as widely varying intensity values and the irregular distribution of cloud points between multitemporal data.

Airborne campaigns are being launched increasingly in which data from different sensors are acquired simultaneously, such as RGB cameras, LiDAR scanners, and multispectral and HS sensors [35-41]. This opens up the possibility to fuse data from different sensors at different levels ranging from raw data fusion, feature-level fusion, or application-level fusion [42]. Of particular interest in HL-Fusion is the ability to operate in a common feature vector using the potential of each sensor and performing the fusion on the feature level. The analysis based on feature-level HL-Fusion enables a complete spectral-local analysis and diversifies the results and products obtained from the fusion of these two sensors [43]. In addition, the analysis based on multitemporal HS and LiDAR data allows the detection and evaluation of complex changes in the urban environment [44]. Man et al. [45] proposed a method for urban land cover classification by extracting normalized Digital Surface Model (nDSM) and backscattered intensity features from LiDAR. The authors first applied Principal Component Analysis (PCA) to an HS dataset, using the first PC to generate texture features based on the grey level co-occurrence matrix (GLCM) [46,47] and additionally retrieved the Normalized Difference Vegetation Index (NDVI) [48]. All extracted features were fed into pixel-based supervised classification algorithms, including SVM and Maximum Likelihood Classification (MLC). Hasani et al. [49] generated a hybrid feature space including spectral and structural features from HS and LiDAR data and an optimized classification system applying cuckoo search. Khodadadzadeh et al. [50] proposed a feature-level fusion method integrating multiple types of HS and LiDAR-based features without model parameter regularization. Kuras et al. [35] extracted endmembers from an HS dataset in an unsupervised way by applying N-FINDR [51] and retrieving raster-based LiDAR features for segmentation purposes.

Simultaneous feature-level fusion of multitemporal data is among the analyses that not only require an understanding of the physical functions of each sensor but are also

very complex compared to analyses relying on single sensors or lower dimensional data. Deep learning has proven to be a critical basis and remarkable breakthrough for handling such issues in image processing in recent years [52]. Continuous improvements and innovations in deep learning models show that no single generic and transferable classification model can correctly analyze the selected target of interest. Very often, the selection of deep learning algorithms depends on the complexity of the classification task, the type of data, their dimensionality, training data availability, and the final classification purpose.

For high-dimensional HS and LiDAR data, the algorithm for multiclass classification of urban land cover should include a wide range of information. Such fundamental information is the spectral context using multidimensional convolutional operations. Also important is the location of the class and the occurrence environment.

Inspired by HL-Fusion and deep learning for multitemporal analysis, we present: (1) a novel land cover multitemporal analysis based on fused HS and LiDAR data at the feature level that integrates abundance representation from HS and LiDAR datasets applying 2D ResU-Net (2D Residual U-Net) [53,54]; (2) we automatically generate spectral libraries as a by-product for a local database expansion creating a spectral library for each defined class based on endmember extraction and forced class assignment based on a synthetic mixture including intraclass variability; (3) we propose a generic method for stable updating of local maps using a case study bitemporal HL-Fusion dataset.

The article is structured as follows: Section 2 describes the dataset used in our study. Section 3 introduces the framework of our proposed method for multitemporal analysis of HL-Fusion data. Section 4 presents the results of segmentation optimization, spectral library generation, and the change detection approach. The results are further discussed in Section 5. Finally, Section 6 points out concluding remarks on our method and suggestions for future directions in this research field.

2. Dataset

The company Terratec AS collected airborne-based HS and LiDAR data in August 2019 and June 2021 over Bærum municipality near Oslo, Norway (Figure 1).
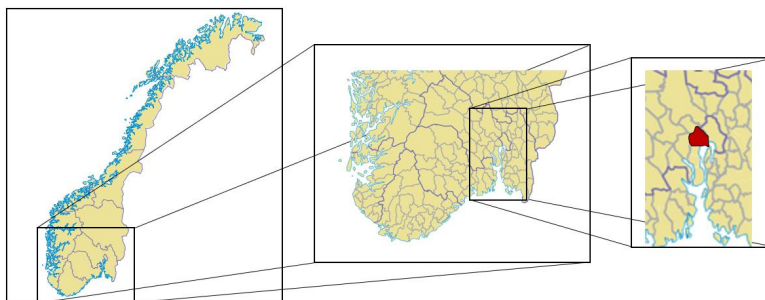

Figure 1. The area of interest located in Norway, in Baerum (red polygon on the figure).

Our study area is located in Høvik with a coordinate extent of 588060, 6641500; 588878, 6641735 WGS 84 / UTM zone 32N (Figure 2 a, b). The datasets contained bitemporal cloud-free airborne-based HS images and LiDAR scans. The HS data were acquired by two HySpex sensors: VNIR-1800 (0.4 – 1.0 μm) and SWIR-384 (1 – 2.5 μm) with 0.3 and 0.7 m spatial resolution, respectively. The HS data were preprocessed by conducting georeferencing and orthorectification using the PARGE software (Parametric Geocoding and Orthorectification for Airborne Optical Scanner Data) [55]. The geocoded radiance data were converted to reflectance, conducting atmospheric correction using ATCOR-4 (Atmospheric and Topographic Correction for airborne imagery). Absorption features

associated with H_2O and OH close to bands at 1.4 µm and 1.9 µm and noisy bands were excluded from further analyses.

The LiDAR data were acquired using Riegl VQ-1560i, with five pulses per m² and intensity at 1.064 µm. From the LiDAR-based point cloud (Figure 2d), noise and outliers were removed. Five different raster-based features were extracted based on other studies [8], including intensity from the first return, height derivatives such as slope, normalized Digital Surface Model (nDSM), multiple returns, and point density. All features were aligned to the spatial resolution of 0.3 m of the HS VNIR scene.

2.1. Ground truth

The ground truth consists of a local database in Norway (FKB database) that includes polygons of artificial objects such as buildings, railways, and roads manually updated from 2011 to 2019. Ground truth data were unavailable for low and high vegetation due to high dynamic and seasonal differences. Therefore these classes were extracted semi-automatically, calculating the Normalised Difference Vegetation Index (NDVI) for the HS scene [48] and distinguished high and low vegetation based on raster-based LiDAR features, relying on the method from Kuras et al. [35] (Figure 2).

a) HS scene from dataset from 2019

b) HS scene from dataset from 2021

c) Ground truth data

d) LiDAR point cloud

unclassified   low vegetation   high vegetation   building   road   railway
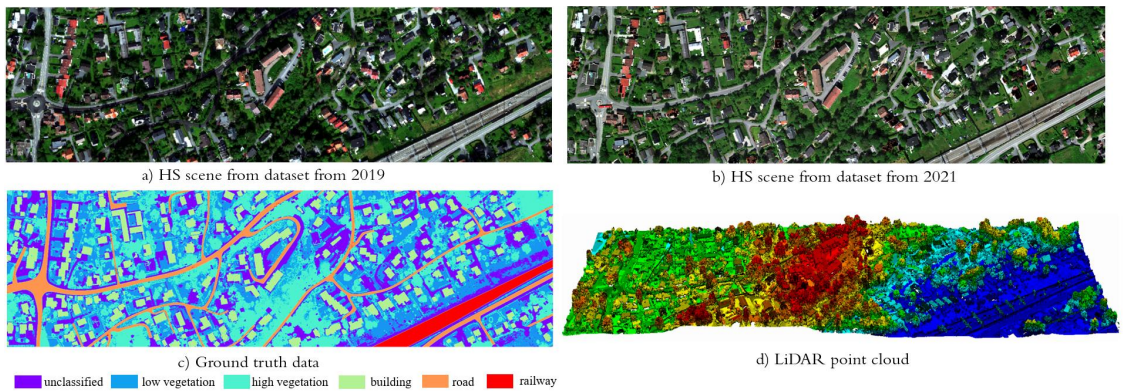
Figure 2. Our study area representing a) the HS scene from 2019, b) the HS scene from 2021, c) the ground truth data, and d) the LiDAR point cloud.

These ground truth data were used for the dataset from 2019. For the dataset from 2021, we used segmentation results from 2019 as reference data.

2.2. Data simulation

In order to diversify the analysis and prove the proposed method's stability and correctness, we created simulated changes, adding a building in place of low and high vegetation. We assumed that adding a building in this location is feasible and typical for urban/suburban areas where vegetation is removed to build new residential neighborhoods. This building was not annotated in the ground truth in training.

3. Proposed method

The following section describes an approach for multitemporal HL-Fusion at the feature level (Figure 3). The analysis begins with unsupervised endmember extraction separately for HS (Figure 3, box 1.1) and LiDAR data (Figure 3, box 1.2) for the first time point (Figure 3, Dataset A). From the created endmembers, abundance maps are generated for HS (Figure 3, box 2.1) and LiDAR data (Figure 3, box 2.2), retrieving the percentage of each endmember per pixel in the scene. These abundance maps and the prepared ground truth data (Figure 3, box 3) are fed into a segmentation algorithm (Figure 3, box 4). Then, analogous to the first dataset, the second time point (Figure 3, Dataset B) is analyzed,

starting from endmember extraction to generation of abundance maps for HS (Figure 3, box 5A and 6A) and LiDAR data (Figure 3, box 5B and 6B).
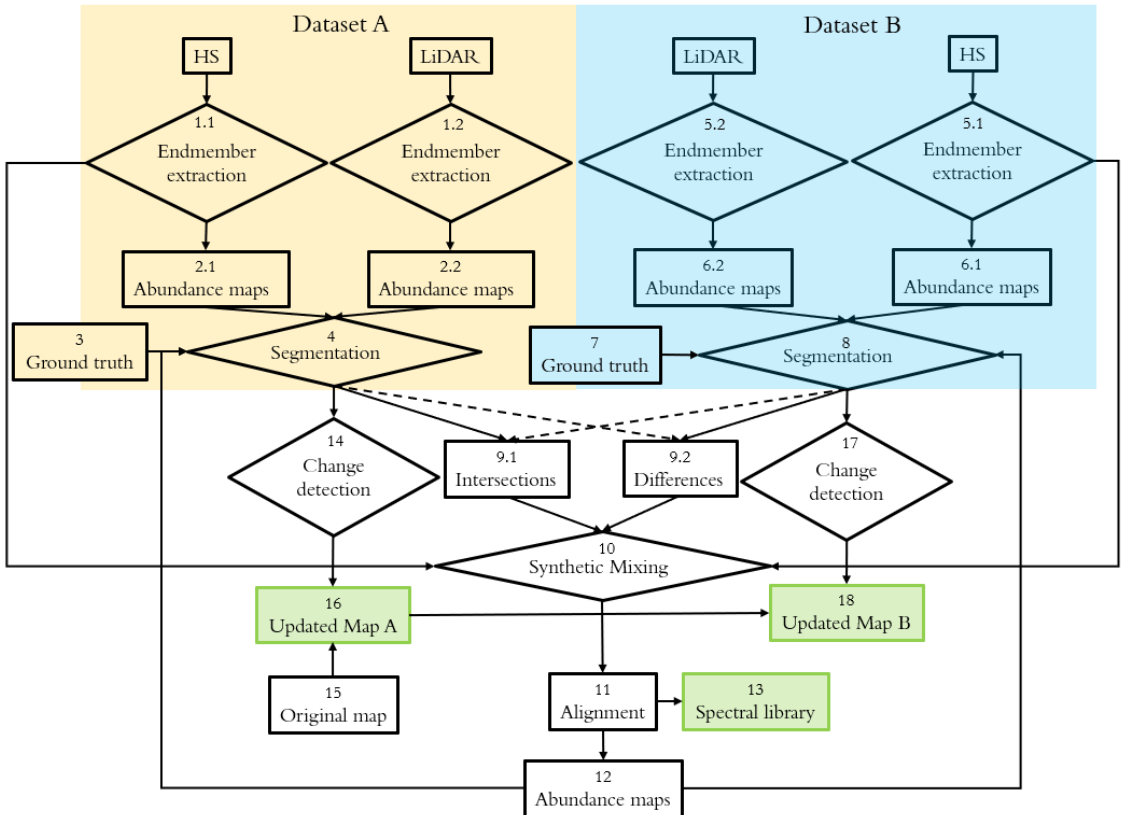
Figure 3. Schematic workflow for multitemporal HL-Fusion at the feature level.

Then, ground truth (Figure 3, box 7) is added to segmentation (Figure 3, box 8) along with retrieved abundance maps. From the results of the first segmentation from Dataset A and B, segments are extracted for each of the defined classes showing class intersections (Figure 3, box 9.1) and segmentation differences between Dataset A and B (Figure 3, box 9.2). The most representative endmembers are extracted from each intersection and difference segments for the classes, such as low and high vegetation, building, road, and railway. In case endmembers from the difference group belong to one of the five predefined classes, synthetic mixing is applied (Figure 3, box 10) to effectively and automatically align endmembers from the already existing difference group to the corresponding class (Figure 3, box 11) and to update initial endmembers extracted before the first segmentation (Figure 3, box 1.1, 1.2). From the updated endmembers, abundance maps are generated (Figure 3, box 12) and fed into the next iteration as optimized segmentation without parameter optimization (Figure 3, box 3, 4, 7, 8). For spectral library generation, Intersections, Differences, synthetic mixing, and alignment are applied (Figure 3, box 9.1, 9.2, 10, 11). After another endmember alignment (Figure 3, box 11), a stable local spectral library with intraclass variability is generated as a by-product (Figure 3, box 13). The final step of the multitemporal analysis is change detection (Figure 3, box 14), where the original map to be updated (Figure 3, box 15) is compared with the segmentation results from Dataset

A, creating an updated map (Figure 3, box 16), which also shows the changes that have occurred in each of the five predefined classes. Analogously, the change detection can be applied on Dataset B (Figure 3, box 17), comparing the map from Dataset A (Figure 3, box 16) to the segmentation results from Dataset B, generating an updated map (Figure 3, box 18).

3.1. Endmember extraction and abundance maps

In our study, we implemented the state-of-the-art iterative endmember extraction (EA) algorithm N-FINDR [51] for HS and LiDAR data, respectively (Figure 3, box 1.1, 1.2). We generated normalized abundance maps based on the extracted endmembers (Figure 3, box 2.1, 2.2) by applying the non-negativity-constrained least squares algorithm [56]. For HS EA, the preprocessed reflectance image was used to retrieve the most representative endmembers. For EA of LiDAR data, we built a LiDAR feature space where the five most relevant raster-based features have been extracted, including slope, the intensity from the first return, multiple returns, normalized digital surface model (nDSM), and point density. All the features were normalized separately before EA due to significant differences in the value scale. The initially extracted endmembers for LiDAR data were used to generate abundance maps for each endmember, analogously to HS data.

3.2. Semantic segmentation

The final input to semantic segmentation algorithms (Figure 3, box 4) consists of abundance maps from HS and LiDAR data. We considered the 2D ResU-Net model architectures in this study [53,54], comparing the segmentation process with and without training data augmentation for 2019 (Figure 3, box 4) and 2021 (Figure 3, box 8) datasets without model parameter regularizations. The original U-net consists of an encoder part with multiple blocks of convolutions and max pools for feature extraction and a corresponding decoder with transposed convolutions for upscaling after each convolution block [57]. Skip connections between corresponding convolution blocks in the encoder and decoder are used for improved class location and signal propagation. The Residual U-net extends the original U-net with local skip connections in the convolution blocks, further enhancing signal pathways and granularity of predictions. In the ResU-Net model, we implemented 2D convolutional operations, which are sufficient and not time-consuming in this type of land cover analysis [35].

The 2D ResU-Net models were implemented in Python using the module Tensorflow with GPU functionalities [58]. To evaluate the model performance, we implemented the Matthews Correlation Coefficient (MCC) that effectively handles unbalanced classes, computing accuracy for each class independently [59] and F1 score [60].

3.2.1. Implementation details

For the segmentation, the study dataset was divided into 64x64 pixel patches, of which 70 % is training and 30 % is the test dataset. 20 % of the training data is used for validation. Training, validation, and testing were selected, considering all classes equally in training, validation, and testing. In addition, data augmentation was applied to training data by applying a 50 % overlap of each patch (Figure 4), of which no patch in the training dataset was part of the test dataset.
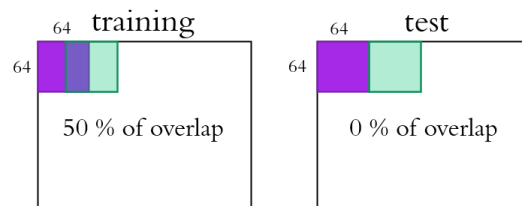

Figure 4. Schematic illustration of data augmentation created for training dataset for 64x64 pixel patches with 50 % overlap. Image patches in the test dataset have an overlap of 0 %.

244

3.2.2. Multitemporal analysis – intersection and differences

Given that the semantic segmentation is conducted for all available datasets X at time $X_1, \ldots, X_n$ for classes $C_1, \ldots, C_i$, each estimated segment of a single class from $C_1, \ldots, C_i$ is handled individually.

For each class $C_1, \ldots, C_i$, common areas occurring in all datasets from different times $X_1, \ldots, X_n$ are grouped into «intersection» (Figure 3, box 9.1). It means that a pixel classified as class C_1 (e.g., high vegetation) in all datasets $X_1, \ldots, X_n$ (e.g., 2019 and 2021) is assigned to the intersection group. For each intersection of a class $C_1, \ldots, C_i$, we extracted representative endmembers from all datasets $X_1, \ldots, X_n$.

In contrast to an intersection, difference means that a pixel in dataset X_1 was assigned to another class than in dataset(s) $X_2, \ldots, X_n$ (Figure 3, box 9.2). For each such difference of each class $C_1, \ldots, C_i$, we found representative endmembers and collected them in one difference group (Figure 5).

245
246
247
248
249
250
251
252
253
254
255
256
257

258
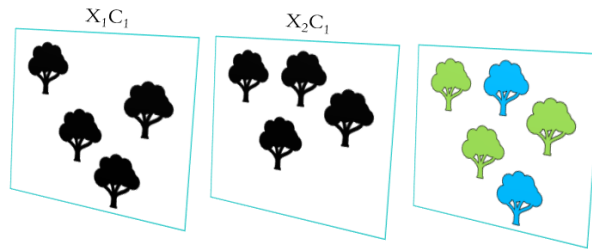
Figure 5. The clustering principle of intersections (green) and differences (blue) for two datasets, X_1 and X_2, for class C_1.

259

3.3. Synthetic mixing for spectral library generation

We assume some endmembers from the difference group can be assigned to any defined classes $C_1, \ldots, C_i$. Therefore, we synthetically mixed all intersection endmembers $EM_{i0}, \ldots, EM_{in}$ with all difference endmembers $EM_{d0}, \ldots, EM_{dm}$ and initial endmembers generated for the first segmentation in proportion 50:50 percent (Figure 3, box 10). After the synthetic mixing, we unmixed the new synthetic matrix spectrally (Figure 6).
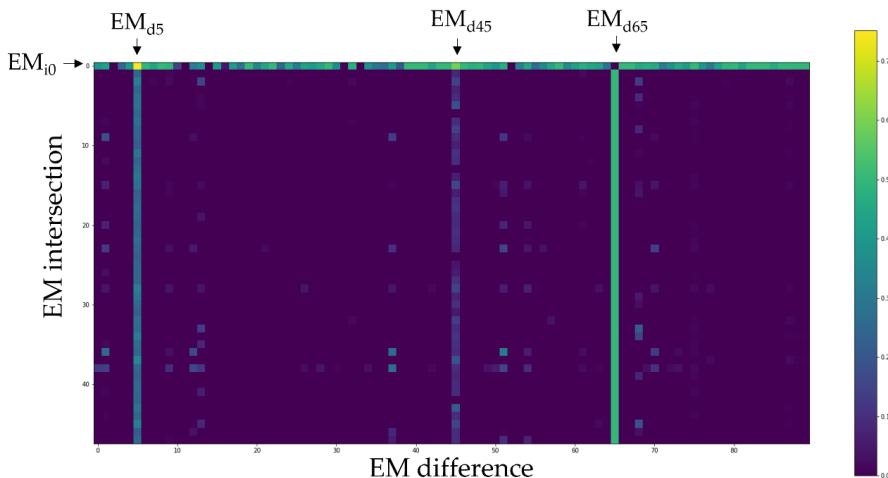
260
261
262
263
264
265
266

267

Figure 6. Spectral unmixing example based on synthetic mixing result.

268

The intuition behind the spectral unmixing in this study is to align difference and initial endmembers to any of the defined classes $C_1, \ldots C_i$, comparing all difference and initial endmembers to all intersection endmembers (Figure 3, box 11). Figure 6 presents a spectral unmixing example, where each class $C_1, \ldots, C_i$ consists of ten intersection endmembers (Table 1).

Table 1. Intersection endmember defined classes.

Defined class	Intersection EM
road	EM_{i0}-EM_{i9}
building	EM_{i10}-EM_{i19}
low vegetation	EM_{i20}-EM_{i29}
high vegetation	EM_{i30}-EM_{i39}
railway	EM_{i40}-EM_{i49}

Given that the intersection endmember EM_{i0} belongs to the road class, we are searching for difference and initial endmembers with a similar unmixing value to EM_{i0}. In this example, two difference endmembers, EM_{d5} and EM_{d65}, are similar to EM_{i0} and were aligned automatically to the road class as the EM_{i0}. For the updated intersection endmembers, abundance maps have been retrieved and summed up for each class separately. Such optimized intersection endmembers are the basis for a stable local spectral library (Figure 3, box 13).

The difference endmembers that were not aligned automatically to any defined classes were used to retrieve abundance maps for the second segmentation iteration. However, to avoid noise and endmembers that are not substantial, we calculated the average of each difference endmember occurring in our study area and limited the amount of the endmembers contributing with a value above 0.1 %. The updated intersection abundance maps, difference abundance maps, and LiDAR features (Figure 3, box 12) extracted for the first segmentation were merged and fed as input for the second segmentation iteration (Figure 3, box 4).

3.4. Change detection

In order to update the local map that served as the ground truth (FKB) for segmentation for the 2019 dataset, we subtracted each segment separately from the 2019 dataset from the FKB reference data (Figure 3, box 15) highlighting land cover changes (Figure 3, box 14). The resulting update map (Figure 3, box 16) indicates the changes in objects/surfaces added or removed in 2019. This procedure was analogously repeated for change detection (Figure 3, box 17) for the 2019 (Figure 4, box 16) to 2021 dataset retrieving an updated map (Figure 4, box 18). Since there were no significant changes in the defined classes in artificial objects such as buildings and railways from 2019 to 2021, we simulated a change and added a random building in place of low/high vegetation in the dataset from 2021.

4. Experimental results

This section provides results for the initial (first iteration) and optimized the second iteration after the abundance map update) segmentation task for 2D ResU-Net with and without data augmentation. The results of the spectral library generated from the best segmentation results of 2019 and 2021 are presented. Then, the results of the change simulation are shown, as well as the results of the change detection for each of the defined classes taking into account the changes from the FKB reference data for 2019 and from 2019 to 2021.

4.1. Segmentation results

311

Table 2 presents the segmentation results for 2019 of the ResU-Net without and with data augmentation for initial segmentation (segmentation I) and optimized segmentation (segmentation II). The results are based on MCC for each class and the F1 score metric for overall segmentation. Corresponding segmentation maps are shown in Figure 7.

312
313
314
315

Table 2. Segmentation results for the dataset from 2019 for the ResU-Net model based on MCC calculated for each class and overall F1 score.

316
317

Color	Dataset	2019 without augmentation		2019 with augmentation	
	Segmentation	I	II	I	II
	Low vegetation	0.79	**0.81**	0.73	**0.75**
	High vegetation	0.92	0.92	0.94	**0.97**
	Building	0.88	**0.94**	0.97	**0.99**
	Road	0.78	**0.89**	0.92	**0.95**
	Railway	0.85	**1**	1	1
	F1	0.818	**0.831**	0.814	**0.843**

318

a) HS scene from dataset from 2019

b) FKB Ground truth data

c) I segmentation without augmentation

d) II segmentation without augmentation

e) I segmentation with augmentation
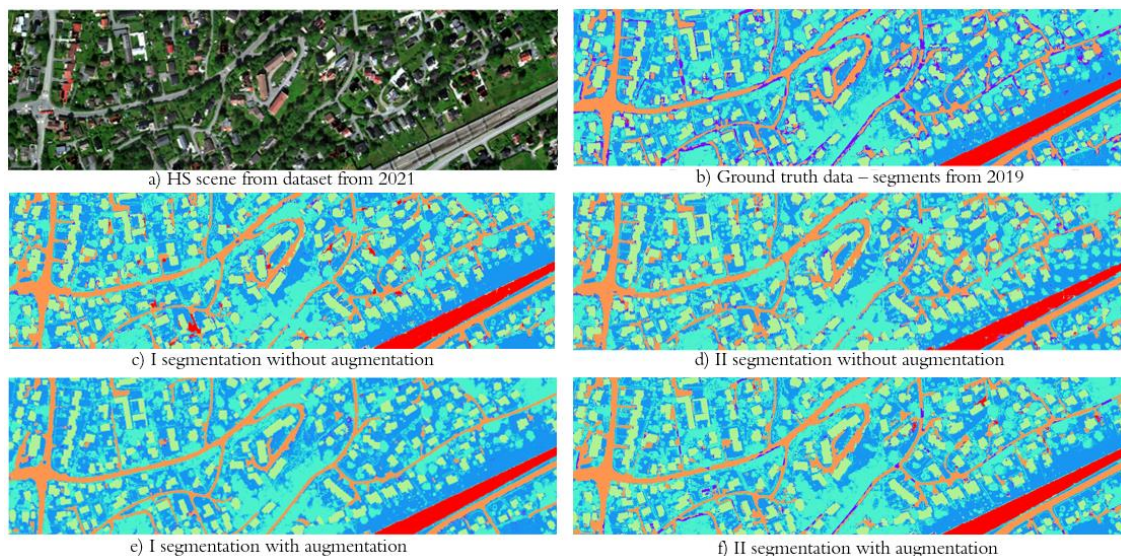
f) II segmentation with augmentation

319

Figure 7. Segmentation results for the dataset from 2019 for the ResU-Net showing a) the HS scene in RGB color representation, b) the FKB reference map to be updated, c) the first segmentation iteration without data augmentation, d) the second segmentation iteration without data augmentation, e) the first segmentation iteration with data augmentation, and f) the second segmentation iteration with data augmentation.

320
321
322
323

Table 2 shows that, regardless of the data augmentation process, the accuracy increases for both individual classes and the overall F1 score. Comparing the results of segmentation without and with data augmentation, the segmentation with data augmentation achieves higher results in both initial (I) and optimized (II) segmentation.

324
325
326
327

Similarly, the 2021 results are reported in Table 3, where, as in 2019, the initial (I) and optimized (II) segmentation results for segmentation without and with data augmentation are compared. The second iteration of segmentation in the main improves the accuracy results in both cases without and with data augmentation. Comparable to 2019, in 2021 also, segmentation with data augmentation outperformed.

328
329
330
331
332

Table 3. Segmentation results for the dataset from 2021 for the ResU-Net model based on MCC calculated for each class and overall F1 score.

Color	Dataset	2021 without augmentation		2021 with augmentation	
	Segmentation	I	II	I	II
	Low vegetation	0.7	**0.81**	0.8	**0.81**
	High vegetation	0.76	**0.91**	0.95	**0.97**
	Building	0.92	**0.98**	0.99	0.99
	Road	0.82	**0.91**	0.99	0.99
	Railway	**0.99**	0.98	1	1
	F1	0.752	**0.776**	0.886	**0.892**

a) HS scene from dataset from 2021

b) Ground truth data − segments from 2019

c) I segmentation without augmentation

d) II segmentation without augmentation

e) I segmentation with augmentation

f) II segmentation with augmentation

Figure 8. Segmentation results for the dataset from 2021 for the ResU-Net showing a) the HS scene in RGB color representation, b) the reference map from 2019 to be updated, c) the first segmentation iteration without data augmentation, d) the second segmentation iteration without data augmentation, e) the first segmentation iteration with data augmentation, and f) the second segmentation iteration with data augmentation.

4.2. Spectral library

The spectral library was generated based on the results of the initial (I) segmentation of 2019 and 2021 and then updated after the results of the optimized (II) segmentation. The spectral library, shown in Figure 9, demonstrated the final spectra for each class, including low and high vegetation, buildings, roads, and railways. Each spectrum covers the 0.4-2.35 µm spectral range. Noisy bands from the preprocessing are not included either in the analysis or in the built spectral library. Each class contains the most representative spectra within its definition, i.e., the building class consists of different roof materials depending on their complexity and heterogeneity in the selected study of interest.
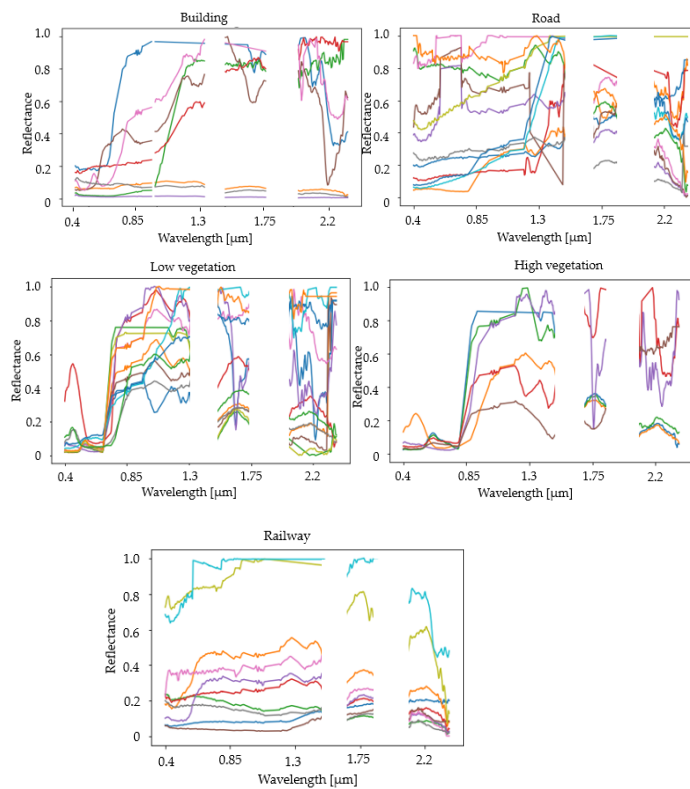
Figure 9. Spectral library of low and high vegetation, building, road, and railway generated after alignment after the II segmentation iteration.

4.3. Change detection

Figure 10 demonstrates the change detection results for the changes from the FKB reference data to 2019 and from 2019 to 2021 for classes such as buildings, roads, and railways. Figure 11 depicts the changes in low and high vegetation from 2019 to 2021 only due to the lack of ground truth data available in the local database before 2019.
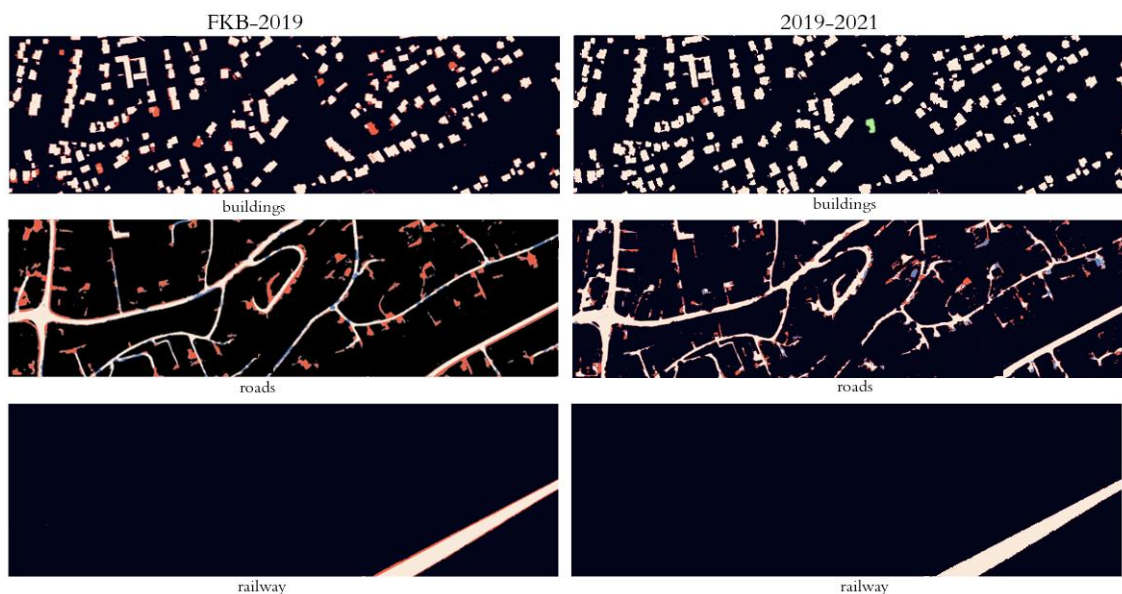
Figure 10. Change detection for buildings, roads, and railways from FKB reference data to 2019 (left) and from 2019 to 2021 (right). Different colors are used to highlight changes, i.e., orange for changed pixels (class present in the updated map), blue for changed pixels (class present in the map to be updated), white for unchanged pixels, and green for simulation.
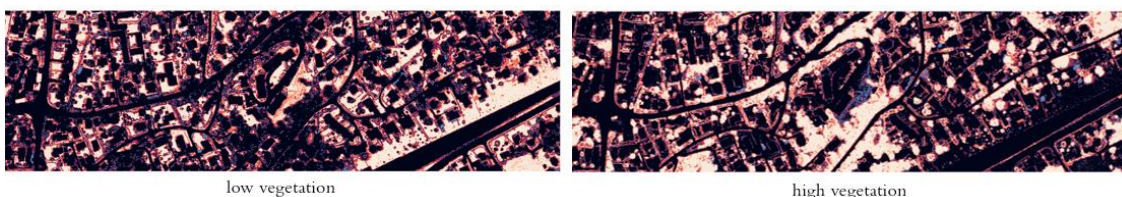

Figure 11. Change detection of low and high vegetation from 2019 to 2021. Different colors are used to highlight changes, i.e., orange for changed pixels (class present in the updated map), blue for changed pixels (class present in the map to be updated), and white for unchanged pixels.

Table 4 presents the MCC accuracy results for segmentation on a dataset where a building was added in the 2021 dataset. Figure 12 highlights the simulated building addition with the building found in the segmentation process.

Table 4. The segmentation results of a simulated dataset from 2021.

Color	Dataset	2021
	Low vegetation	0.79
	High vegetation	0.92
	Building	0.97
	Road	0.99
	Railway	1
	F1	0.859

Figure 12. Simulated change in the dataset from 2021. The red dashed rectangle represents an added building in place of low/high vegetation.

5. Discussion

5.1. Segmentation process

In the segmentation of high-dimensional data, dimension reduction is crucial. Unsupervised endmember extraction and retrieval of abundance maps provide a stable and reliable method for obtaining the most representative features of a scene that are not calculated based on statistics. Based on the segmentation results from 2019 and 2021, it has to be noted that only in single classes, such as railway in 2021, the second iteration of segmentation after optimization deteriorated accuracy from 99 to 98 %. Comparing all classes in general, low and high vegetation achieved significantly lower accuracy than the buildings, roads, and railways. This is because ground truth data for these classes were created semiautomatically, which is insufficiently accurate. In addition, in many places in our study area, vegetation partially covers some objects, such as roads and buildings, depending on the season in which the data acquisition campaign was carried out. For this reason, some road pixels were not found in the segmentation. The main aspect is that HS data do not penetrate the surface, and some extracted features from LiDAR, such as intensity, include information from the surface only, i.e., from the first laser return. Figure 13 shows an example of high vegetation covering a road.
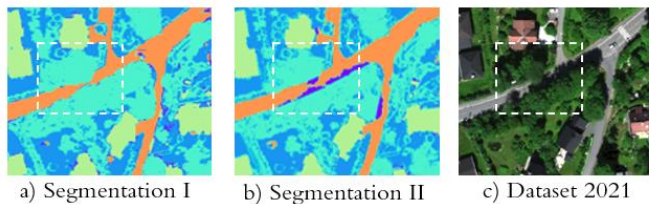

a) Segmentation I b) Segmentation II c) Dataset 2021

Figure 13. An example of vegetation covering the road marked with a white dashed rectangle.

In the I segmentation (Figure 13a), vegetation was identified on the road, degrading the road accuracy results. In the II segmentation, vegetation covering the road was reclassified to the unknown class. When a classification of new objects or surfaces is required, one of the new classes in this study of interest could be "vegetation on the road". Such information about vegetation covering main roads can be an indicator for municipalities to remove or secure high vegetation that threatens vehicular traffic.

In addition to the improved accuracy of the results in the second iteration of segmentation, after segmentation optimization, in most objects, the edges have been sharpened, and the geometry in the 2D plane is approximated to reality. This is especially noticeable in objects not marked in ground truth data but present in the analyzed dataset, such as building detection in Figure 14.
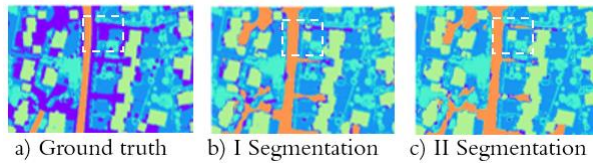
a) Ground truth   b) I Segmentation   c) II Segmentation

Figure 14. Improvement of the shape of an object on example building detection marked with a dashed white rectangle.

A frequent challenge in multiclass segmentation using 2D convolutional operations is patch edge effects. These effects relate to the generation of "contours" of each 64x64 pixel patch in the classification results, meaning that contextual information from neighboring pixels is not included at the pixel patch edges (Figure 7c) [61]. This unintentional effect has been reported in previous studies, where Kuras et al. [35] implemented 3D instead of 2D convolutional operations in the model, mitigating edge effects in the final segmentation map. However, in this study, 3D convolutions require increased computation time, especially when applying data augmentation.

In our study, the second iteration of the segmentation eliminates patch edge effects, thereby improving accuracy results (Figure 7d, f). On the other hand, data augmentation through a patch overlap of 50 % in the training dataset not only levels edge effects in the I segmentation already (Figure 7e) but also allows for more stable learning of the algorithm of localization patterns helping in object identification.

Another important aspect worth discussing is achieving 100 % accuracy in some classes. This is a sign of overfitting, which can be compensated for by reducing the complexity of the segmentation model. The splitting of training may have caused overfitting, and test data was carried out patch-wise rather than strictly object-wise, causing the same object/surface to occur both in training and test data. Also, the weight of each class may have been disproportionate, with the rest getting more focus in the segmentation.

5.2. Spectral library

Since the defined classes in the segmentation in the urban environment are complex and heterogeneous, the automatically created spectral library for each class contains spectra belonging to different materials and surfaces within that class. One example is the road class which consists of not only the primary road material – asphalt or concrete, but also road surface markings which, due to the spatial resolution of 0.3 m, are mixed with the asphalt or the vehicle on the street. In Figure 9, we can observe that some of the spectra, especially in low and high vegetation, experience saturation caused by technical problems in data acquisition or atmospheric correction. Errors and artifacts in the spectral library can also be caused by the level adjustment between VNIR and SWIR sensors and the fact that these two cameras do not point to the same spot from the airplane, which is particularly problematic when dealing with a dynamic environment where, for example, vehicles are constantly moving.

5.3. Change detection

Change detection for all defined classes is shown in Figures 10 and 11. The most significant changes could be experienced when the updated map represented 2019. In 2019, several new buildings were detected and updated. Sometimes in building maps, we notice detected changes with already known buildings. This is because boats and cars on the properties were assigned to the building class. In the case of road change detection, as roads were defined, property entrances or alleys were not marked in the original reference data. In addition, the algorithm also detected changes when ground truth was prepared manually, and the object from the airborne-based image perspective slightly shifted or changed object edges.

Furthermore, it is crucial to point out that our novel framework has effectively identified actual or simulated objects and when ground truth data are not aligned with the current position, as is the case with local map updates.

Our framework allows the addition of new time point datasets, thanks to which the focus of the analysis can be on high-frequent and low-frequent changes or mobile and static object recognition in an urban scene. Moreover, adding another dataset allows for building a stable spectral library and features that can be transferred to other study areas. Objects and surfaces not identified during segmentation can be manually added, allowing for dynamic class extension of urban land cover.

6. Conclusions

This study presents a novel approach to feature-level-based multisensor data fusion of HS and LiDAR data proposing a method for an effective segmentation optimization based on the unsupervised endmember extraction and abundance map retrieval of HS and LiDAR data without parameter regularizations. Objects that have not been identified can be added manually, with the possibility of dynamic expansion and a variety of land cover classes. All the models achieved increased segmentation results after segmentation optimization. The ResU-Net with implemented data augmentation outperformed compared to models without data augmentation. In addition, a local spectral library has been generated automatically as a by-product that can be used to expand the local urban database and serve as a basis for further updates of this region. Based on the segmentation and generated spectral library, we created a change map of each defined class, creating a local map update.

7. Future work

Our novel approach serves as a promising basis for developing a change detection framework based on unsupervised segmentation. Such unsupervised segmentation would limit issues related to the preparation of the ground truth data, which are not always available and updated for the algorithm to learn correctly segment complex objects and surfaces. The proposed framework of change detection applying fused HS and LiDAR data at the feature level can be expanded with more datasets of the study area, allowing segmentation and spectral signatures of individual objects or surfaces to be more stable and reliable.

Author Contributions: Conceptualization, AK and MB; analysis, AK; investigation, AK; writing—original draft preparation, AK; writing—review and editing, MB, KHL, IB; visualization, AK; supervision, IB. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

8. Bibliography

1.	Heiden, U.; Segl, K.; Roessner, S.; Kaufmann, H. Determination of robust spectral features for identification of urban surface materials in hyperspectral remote sensing data. *Remote Sens. Environ.* **2007**, *111*, 537-552.

2.   Cadenasso, M.L.; Pickett, S.T.A.; Schwarz, K. Spatial heterogeneity in urban ecosystems: reconceptualizing
      land cover and a framework for classification. *Frontiers in Ecology and the Environment* **2007**, *5*, 80-88.

3.   Jing, C.; Zhou, W.; Qian, Y.; Yu, W.; Zheng, Z. A novel approach for quantifying high-frequency urban land
      cover changes at the block level with scarce clear-sky Landsat observations. *Remote Sens. Environ.* **2021**, *255*.

4.   Banzhaf, E.; Kabisch, S.; Knapp, S.; Rink, D.; Wolff, M.; Kindler, A. Integrated research on land-use changes in
      the face of urban transformations - An analytic framework for further studies. *Land Use Policy* **2017**, *60*, 403-
      407.

5.   Hegazy, I.R.; Kaloop, M.R. Monitoring urban growth and land use change detection with GIS and remote
      sensing techniques in Daqahlia governorate Egypt. *International Journal of Sustainable Built Environment* **2015**, *4*,
      117-124.

6.   Wellmann, T.; Lausch, A.; Andersson, E.; Knapp, S.; Cortinovis, C.; Jache, J.; Scheuer, S.; Kremer, P.;
      Mascarenhas, A.; Kraemer, R.; et al. Remote sensing in urban planning: Contributions towards ecologically
      sound policies? *Landsc. Urban Plan.* **2020**, *204*.

7.   Yin, J.; Dong, J.; Hamm, N.A.S.; Li, Z.; Wang, J.; Xing, H.; Fu, P. Integrating remote sensing and geospatial big
      data for urban land use mapping: A review. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *103*.

8.   Kuras, A.; Brell, M.; Rizzi, J.; Burud, I. Hyperspectral and Lidar Data Applied to the Urban Land Cover
      Machine Learning and Neural-Network-Based Classification: A Review. *Remote Sens.* **2021**, *13*,
      doi:https://doi.org/10.3390/rs13173393.

9.   Shahtahmassebi, A.R.; Li, C.; Fan, Y.; Wu, Y.; Iin, Y.; Gan, M.; Wang, K.; Malik, A.; Blackburn, G.A. Remote
      sensing of urban green spaces: A review. *Urban Forestry & Urban Greening* **2021**, *57*.

10.   Roessner, S.; Segl, K.; Heiden, U.; Kaufmann, H. Automated differentiation of urban surfaces based on
      airborne hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*.

11.   Tan, K.; Wang, H.; Chen, L.; Du, Q.; Du, P.; Pan, C. Estimation of the spatial distribution of heavy metal in
      agricultural soils using airborne hyperspectral imaging and random forest. *Journal of Hazardous Materials* **2020**,
      382.

12.   Qu, J.; Hou, S.; Dong, W.; Li, Y.; Xie, W. A Multi-Level Encoder-Decoder Attention Network for Change
      Detection in Hyperspectral Images. *IEEE Trans. Geosci. Remote Sens.* **2021**, doi:10.1109/TGRS.2021.3130122.

13.   Campbell, J.B. *Introduction to Remote Sensing*, Guilford Press ed.; New York, 2010.

14.   Song, A.; Choi, J.; Han, Y.; Kim, Y. Change Detection in Hyperspectral Images Using Recurrent 3D Fully
      Convolutional Networks. *Remote Sens.* **2018**, *10*.

15.   Bruzzone, L.; Bovolo, F. A Novel Framework for the Design of Change-Detection Systems for Very-High-
      Resolution Remote Sensing Images. *Proceedings of the IEEE* **2012**, *101*, 609-630,
      doi:10.1109/JPROC.2012.2197169.

16.   Aksoy, S. Spatial techniques for image classification. In *Signal and Image Processing for Remote Sensing*; 2008; pp.
      491-513.

17.   Benediktsson, J.A.; Palmason, J.A.; Sveinsson, J.R. Classification of hyperspectral data from urban areas based
      on extended morphological profiles. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 480-491.

18.   Jouni, M.; Mura, M.D.; Comon, P. Hyperspectral Image Classification Based on Mathematical Morphology
      and Tensor Decomposition. *Math. Morphol. Theory Appl.* **2020**, *4*, 1-30.

19.   Mura, M.D.; Benediktsson, J.A.; Waske, B.; Bruzzone, L. Extended profiles with morphological attribute filters
      for the analysis of hyperspectral data. *Int. J. Remote Sens.* **2010**, *31*, 5975-5991.

20. Bovolo, F. A Multilevel Parcel-Based Approach to Change Detection in Very High Resolution Multitemporal Images. *IEEE Geosci. Remote Sens. Lett.* **2008**, *6*, 33-37.

21. Plaza, A.; Benediktsson, J.A.; Boardman, J.W.; Brazile, J.; Bruzzone, L.; Camps-Valls, G.; Chanussot, J.; Fauvel, M.; Gamba, P.; Gualtieri, A.; et al. Recent advances in techniques for hyperspectral image processing. *Remote Sens. Environ.* **2009**, *13*, 110-122.

22. Sun, L.; Wu, Z.; Liu, J.; Xiao, L.; Wei, Z. Supervised spectral-spatial hyperspectral image classification with weighted Markov Random Fields. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1490-1503.

23. Li, J.; Bioucas-Dias, J.M.; Plaza, A. Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and markov random fields. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*.

24. Cao, X.; Zhou, F.; Xu, L.; Meng, D.; Xu, Z.; Paisley, J. Hyperspectral image classification with markov random fields and a convolutional neural network. *IEEE Trans. Image Process.* **2017**, *27*, 2354-2367.

25. Li, Y.; Zhang, H.; Shen, Q. Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sens.* **2017**, *9*.

26. Alonzo, M.; Bookhagen, B.; Roberts, D.A. Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sens. Environ.* **2014**, *148*, 70-83.

27. Zhao, X.; Tao, R.; Li, W.; Li, H.C.; Du, Q.; Liao, W.; Philips, W. Joint Classification of Hyperspectral and LiDAR Data Using Hierarchical Random Walk and Deep CNN Architecture. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 7355-7370.

28. Hong, D.; Gao, L.; Hang, R.; Zhang, B.; Chanussot, J. Deep encoder-decoder networks for classification of hyperspectral and LiDAR data. *IEEE Geosci. Remote Sens. Lett.* **2020**, *99*, 1-5.

29. Feng, Q.; Zhu, D.; Yang, J.; Li, B. Multisource Hyperspectral and LiDAR Data Fusion for Urban Land-Use Mapping based on a Modified Two-Branch Convolutional Neural Network. *ISPRS Int. J. Geoinf.* **2019**, *8*.

30. Fang, L.; Zhu, D.; Yue, J.; Zhang, B.; He, M. Geometric-Spectral Reconstruction Learning for Multi-Source Open-Set Classification With Hyperspectral and LiDAR Data. *IEEE/CAA Journal of Automatica Sinica* **2022**, *9*, 1892-1895.

31. Gaulton, R.; T.Malthus. LiDAR mapping of canopy gaps in continuous cover forests: A comparison of canopy height model and point cloud based techniques. *Int. J. Remote Sens.* **2008**, *31*, 17-19.

32. Morsdorf F., M.E., Allgöwer B., Nüesch D. Clustering in airborne laser scanning raw data for segmentation of single trees. **2003**.

33. Marinelli, D.; Paris, C.; Bruzzone, L. An Approach to Tree Detection Based on the Fusion of Multitemporal LiDAR Data. *IEEE Geosci. Remote Sens. Lett.* **2019**, *99*, 1-5.

34. Yu, H.; Lu, X.; Cheng, G.; Ge, X. Detection and volume estimation of mining subsidence based on multi-temporal LiDAR data. *19th International Conference on Geoinformatics* **2011**.

35. Kuras, A.; Jenul, A.; Brell, M.; Burud, I. Comparison of 2D and 3D semantic segmentation in urban areas using fused hyperspectral and lidar data. *Journal of Spectral Imaging* **2022**, *11*.

36. Senchuri, R.; Kuras, A.; Burud, I. Machine Learning Methods for Road Edge Detection on Fused Airborne Hyperspectral and LIDAR Data. In Proceedings of the 11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), Amsterdam, Netherlands, 2021.

37. Singh, M.K.K.; Mohan, S.; Kumar, B. Fusion of hyperspectral and LiDAR data using sparse stacked autoencoder for land cover classification with 3D-2D convolutional neural network. *J. Appl. Remote Sens.* **2022**, *16*, 034523.

38. Degerickx, J.; Roberts, D.A.; McFadden, J.P.; Hermy, M.; Somers, B. Urban tree health assessment using airborne hyperspectral and LiDAR imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 26-38.

39. Hänsch, R.; Hellwich, O. Fusion of Multispectral LiDAR, Hyperspectral and RGB Data for Urban Land Cover Classification. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 366-370.

40. Brell, M.; Segl, K.; Guanter, L.; Bookhagen, B. Hyperspectral and Lidar Intensity Data Fusion: A Framework for the Rigorous Correction of Illumination, Anisotropic Effects, and Cross Calibration. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*.

41. Brell, M.; Segl, K.; Guanter, L.; Bookhagen, B. 3D hyperspectral point cloud generation: fusing airborne laser scanning and hyperspectral imaging sensors for improved object-based information extraction. *ISPRS J. Photogramm. Remote Sens.* **2019**, *149*, 200-214.

42. Khaleghi, B.; Khamis, A.; Karray, F.; Razavi, S.N. Multisensor Data Fusion: A Review of the State-of-the-art. *Inf. Fusion* **2013**, *14*.

43. Kahramann, S.; Bacher, R. A comprehensive review of hyperspectral data fusion with lidar and sar data. *Annual Reviews in Contol* **2021**, *51*, 236-253.

44. Voss, M.; Sugumaran, R. Seasonal Effect on Tree Species Classification in an Urban Environment Using Hyperspectral Data, LiDAR, and an Object- Oriented Approach. *Sensors* **2008**, *8*, 3020-3036.

45. Man, Q.; Dong, P.; Guo, H. Pixel- and feature-level fusion of hyperspectral and lidar data for urban land-use classification. *Int. J. Remote Sens.* **2015**, *36*, 1618-1644.

46. Ojala, T.; Pietikainen, M.; Maenpaa, T.T. Multi resolution gray scale and rotation invariant texture classification with local binary pattern. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971-987.

47. Shirowzhan, S.; Trinder, J. Building classification from LiDAR data for spatial-temporal assessment of 3D urban developments. *Procedia Eng.* **2017**, *180*, 1453-1461.

48. Thenkabail, P.S.; Smith, R.B.; Pauw, E.D. Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics. *Remote Sens. Environ.* **2000**, *71*, 158-182.

49. Hasani, H.; Samadzadegan, F.; Reinartz, P. A metaheuristic feautre-level fusion strategy in classification of urban area using hyperspectral imagery and LiDAR data. *Eur. J. Remote Sens.* **2017**, *50*, 222-236.

50. Khodadadzadeh, M.; Li, J.; Prasad, S.; Plaza, A. Fusion of Hyperspectral and LiDAR Remote Sensing Data Using Multiple Feature Learning. *IEEE Journal of Selected Topics in Applied Earth Obersvations and Remote Sensing* **2015**, *8*, 2971-2983.

51. Winter, M.E. N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data. *Imaging Spectroscopy V* **1999**, *3753*, 266-275.

52. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436-444.

53. Yang, X.; Li, X.; Ye, Y.; Zhang, X.; Zhang, H.; Huang, X.; Zhang, B. Road detection via deep residual dense u-net. In Proceedings of the International Joint Conference on Neural Networks, Budapest, Hungary, 2019.

54. Zhang, Z.; Liu, Q.; Wang, Y. Road extraction by deep residual u-net. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*.

55. Schläpfer, D.; Richter, R. Geo-atmospheric processing of airborne imaging spectrometry data. Part1: Parametric orthorectification. *Int. J. Remote Sens.* **2002**, *23*, 2609-2630.

56. Bro, R.; Jong, S.d. A fast non-negativity-constrained least squares algorithm. *Journal of Chemometrics* **1997**, *11*, 393-401.

57. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, October 2015, 2015; pp. 234-241.

58. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467* **2016**.

59. Matthews, B.W. Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure* **1975**, *405*, 442-451.

60. Sasaki, Y. The truth of the F-measure. *Teach tutor mater* **2007**, *1*, 1-5.

61. Pielawski, N.; Wählby, C. Introducing Hann windows for reducing edge-effects in patch-based image segmentation. *PLoS One* **2020**, *15*, e0229839.

Paper D:

Integration of Hyperspectral and Magnetic Data for Geological Characterization of the Niaqornarssuit Ultramafic Complex in West-Greenland

Integration of Hyperspectral and Magnetic Data for Geological Characterization of the Niaqornarssuit Ultramafic Complex in West-Greenland

Agnieszka Kuras [1,*] , Björn H. Heincke [2] , Sara Salehi [2] , Christian Mielke [3,] Nicole Köllner [4,5],

Christian Rogass [6], Uwe Altenberger [5] and Ingunn Burud[1]

[1]   *Faculty of Science and Technology, Norwegian University of Life Sciences, PB 5003, 1430 Aas, Norway*

[2]   *Geological Survey of Denmark and Greenland, 1350 Copenhagen, Denmark*

[3]   *Rad.Data, 14482 Potsdam, Germany*

[4]   *German Research Centre of Geosciences, Telegrafenberg, 14473 Potsdam, Germany*

[5]   *Institute of Geosciences, University of Potsdam, 14476 Potsdam, Germany*

[6]   *Helmholtz Centre for Environmental Research, 04318 Leipzig, Germany*

MDPI

Article

Integration of Hyperspectral and Magnetic Data for Geological Characterization of the Niaqornarssuit Ultramafic Complex in West-Greenland

Agnieszka Kuras [1,*], Björn H. Heincke [2], Sara Salehi [2], Christian Mielke [3], Nicole Köllner [4,5], Christian Rogass [6], Uwe Altenberger [5] and Ingunn Burud [1]

[1] Faculty of Science and Technology, Norwegian University of Life Sciences, PB 5003, 1430 Aas, Norway
[2] Geological Survey of Denmark and Greenland, 1350 Copenhagen, Denmark
[3] Rad.Data, 14482 Potsdam, Germany
[4] German Research Centre of Geosciences, Telegrafenberg, 14473 Potsdam, Germany
[5] Institute of Geosciences, University of Potsdam, 14476 Potsdam, Germany
[6] Helmholtz Centre for Environmental Research, 04318 Leipzig, Germany
[*] Correspondence: agnieszka.kuras@nmbu.no

Abstract: The integration of imaging spectroscopy and aeromagnetics provides a cost-effective and promising way to extend the initial analysis of a mineral deposit. While imaging spectroscopy retrieves surface spectral information, magnetic responses are used to determine magnetization at both shallower and greater depths using 2D and 3D modeling. Integration of imaging spectroscopy and magnetics improves upon knowledge concerning lithology with magnetic properties, enhances understanding of the geological origin of magnetic anomalies, and is a promising approach for analyzing a prospective area for minerals having a high iron-bearing content. To combine iron diagnostic information from airborne hyperspectral and magnetic data, we (a) used an iron absorption feature ratio to model pseudo-magnetic responses and compare them with the measured magnetic data and (b) estimated the apparent susceptibility along the surface by some equivalent source modeling, and compared them with iron ratios along the surface. For this analysis, a Modified Iron Feature Depth index was developed and compared to the surface geochemistry of the rock samples in order to validate the spectral information of iron. The comparison revealed a linear increase in iron absorption feature depths with iron content. The analysis was performed by empirically modeling the statistical relationship between the diagnostic absorption features of hyperspectral (HS) image spectra of selected rock samples and their corresponding geochemistry. Our results clearly show a link between the spectral absorption features and the magnetic response from iron-bearing ultra/-mafic rocks. The iron absorption feature ratio of Fe^{3+}/Fe^{2+} integrated with aeromagnetic data (residual magnetic anomaly) allowed us to distinguish main rock types based on physical properties. This separation matches the lithology of the Niaqornarssuit complex, our study area in West Greenland.

Keywords: geological remote sensing; magnetics; hyperspectral; hyperspectral-magnetic integration; ultramafic complex; Greenland; iron; susceptibility; imaging spectroscopy; data fusion

1. Introduction

In harsh environments such as in the Arctic, conventional mineral exploration techniques are challenging due to inaccessibility and remoteness. Optical remote sensing and airborne geophysics offer cost- and time-efficient tools to analyze mineral deposits based on physical properties. Datasets from such methods typically cover large areas and provide physical parameter information which enables the spatial description of the structural and lithological conditions. This makes these methods particularly attractive in an early exploration phase.

However, ore-formation typically involves several processes, and deposits are often hosted in complex background geology and may be overprinted by subsequent tectonic

or alteration processes. Therefore, it is rarely the case that a single physical parameter is a diagnostic criterion for the mineralization and that a single method sufficiently describes a deposit. Optical remote sensing and geophysical methods should be combined to address these limitations since consideration of several physical parameters can significantly impact a more successful mineral exploration [1,2].

The key component of the integration of magnetic and optical remote sensing with lithological information is iron. Iron is the most relevant element since (1) a few iron minerals are almost exclusively responsible for the overall magnetic properties of rocks and, hence, anomalies measured in magnetic data, and (2) iron alteration minerals have strong absorption features in multi-/hyperspectral data.

Magnetizations in rocks are mainly associated with ferromagnetic magnetite, other iron-titanium oxide minerals, and the monoclinic sulfide pyrrhotite [3]. This makes magnetic measurements attractive for targeting several mineral deposits containing Fe-oxides and sulfides [4,5]. However, most other minerals, including the alteration minerals of iron (goethite, hematite), have no significant magnetic properties, so a comprehensive lithological characterization is hardly possible based on magnetic data alone.

Magnetite has no characteristic spectral absorption features in the solar-reflective spectral range. However, its presence can be estimated from Fe-alteration products, which show more distinct spectral signatures [5,6]. The transformation of goethite and hematite to magnetite was observed during laboratory heating experiments [7–9]. Another limiting factor in an accurate geological analysis is that in many regions either the bedrock is not fully exposed (e.g., partly covered by vegetation or soil layer) or the exposed bedrock surface is primarily covered by lichens [10] as is common in the Arctic. Such lichen coverage can lead to misidentification in the spectra and strongly hampers the mineral and rock composition mapping from the spectral analysis [11,12].

Magnetic data might help to overcome the optical analysis limitations at the rock surface in such cases. It enables tracing geological trends in areas where the information from HS sensors is vague due to full or partial coverage of the rock surface by other materials. Integration of the two methods would also assist in the assignment of magnetic properties to the mineralogy and rock types by geologically contextualizing the magnetic data.

Despite many integration approaches combining optical remote sensing methods (e.g., HS sensors with different spectral ranges [13–17], HS with LiDAR [18,19], or HS with photogrammetry [20,21]), only a little research has been done to combine datasets from the two groups (optical remote sensing and airborne/UAV geophysics) at the interpretation stage for geological applications [22,23]. However, especially for ultramafic/mafic complexes exhibiting distinct magnetic properties, the data integration of airborne magnetic and hyperspectral (HS) imagery (HSM integration) appears to be particularly meaningful in improving the knowledge about lithology.

Most applications combining such methods are seen in areas where bedrock is well-exposed and only sparsely covered by vegetation, such as desert, high-alpine, or arctic environments, and where both magnetic surveys and HS data of high resolution and good quality exist at the same time. Geological characterization with HS data is possible in most of these areas [12,24–30] because the bedrock is often well-exposed and is only slightly covered by vegetation. Moreover, HS data allows a general mineralogical characterization of the rock surface and can contribute to closing the gap of missing lithological information. Many mineral exploration areas have been uniformly covered with HS airborne surveys and local helicopter-borne magnetic surveys with dense line spacing (<200 m).

Although an HSM integration can provide largely complementary information about the mineralogical composition and appears promising for several mineral exploration target types in many regions, few papers have been published on where data from these methods are combined [2,23,31,32].

This is likely due to the fact that the proper integration of HS and magnetic data is not straightforward for several reasons. First, the measured magnetic responses are generated by magnetization in the whole sub-surface and not only by its near-surface

contributions. Furthermore, the magnetic method is known to have a low resolution, such that the exact position and distribution of the magnetization in the ground cannot be determined without additional information, and an infinite number of magnetization distributions in the subsurface can describe the measured data response. This means that without general assumptions about how the magnetization is distributed in the ground, it is challenging to compare magnetics with HS properties at the surface reliably. This situation is further complicated by the fact that magnetization is a vector quantity. This vector quantity consists of an induced component parallel to the current Earth's magnetic field and a remanent component that can point in a different direction and reflects the Earth's magnetic field direction when rocks were formed and/or overprinted.

Therefore, it is expected that a simple comparison of data responses presented as magnetic maps (typically given as residual magnetic anomaly) with image products from the HS can only coarsely describe the relationships of magnetic and HS properties on the ground. This is particularly true if magnetic surveys are gathered at greater flight heights and wider line spacings, or when the geology is complex and varies over short distances, or the remanent magnetization component is strong and points in a largely different direction than the induced magnetization.

A plausible future solution for this problem could be to develop magnetic modeling or inversion schemes, where the HS information is incorporated as constraining information along the surface. Such approach would improve the resolution of the magnetics at the surface and, hence, make the magnetization directly comparable to the hyperspectral image results at each pixel. This would allow for a reliable presentation of the relationships between hyperspectral image products and magnetizations in scatterplots and the use of these as inputs in statistical analysis. However, to develop such schemes, it must first be investigated how spectral and magnetic properties are related to each other, such that a geologically reasonable constraint can be found that links the two datasets with each other.

This contribution aims to investigate if it is generally possible to gain more geological information from mineral exploration sites and iron-bearing ultramafic complexes in particular by combining airborne magnetic and HS data. In this contribution, we make the first step in developing an approach to integrating airborne HS and magnetic data properly. We apply an HSM integration using simple forward modeling and inversion tests—fully aware of their limitations. In particular, we compare the resulting apparent magnetic susceptibility estimates along the surface with iron ratios determined from the HS data to make a geological characterization in terms of Fe-mineral variations. The study is supplemented with hyperspectral laboratory and geochemical measurements on rock samples which helps to validate the observed relationships of the different physical parameters in the airborne data and link them to mineralogical knowledge.

2. Study Area—The Niaqornarssuit Complex

As a test area, we selected the intrusive Niaqornarssuit Complex located within the southernmost part of the Paleoproterozoic Nagssugtoqidian Orogen in West Greenland (66.83°N, −52.02°E) which is explored for disseminated Ni and Cu sulfide mineralizations [33]. The area is fully covered both with a regional airborne hyperspectral survey (HyMAP [29]) and a local high-resolution helicopter-borne magnetic survey [34]. The rocks in the complex are largely exposed at the surface, such that it satisfies the conditions to test the integration of HS and magnetic data properly.

The Niaqornarssuit Complex formed in the Palaeoproterozoic from 2050 to 1750 Ma [35] is a layered ultramafic intrusion hosted in Archean gneisses. It is elliptically shaped with a long axis of 1.8 km striking east-west, and a short axis measuring 0.9 km [36]. The complex is a peridotite lens consisting mainly of two homogeneous yellowish-green weathered dunite bodies separated by a thin olivine-poorer peridotite zone with lherzoliteic and harzburgitic composition surrounded by older basement orthogneisses (Figure 1). Some limited observations of intrusive contacts and a significant chilled margin in the southeastern part of the Niaqornarssuit Complex imply that the intrusion was embedded

in older basement gneisses of the South Nagssugtoqidian Orogen [33]. The mafic and ultramafic rocks of the study area experienced a metamorphic and tectonic overprint forming the present shape of the complex.
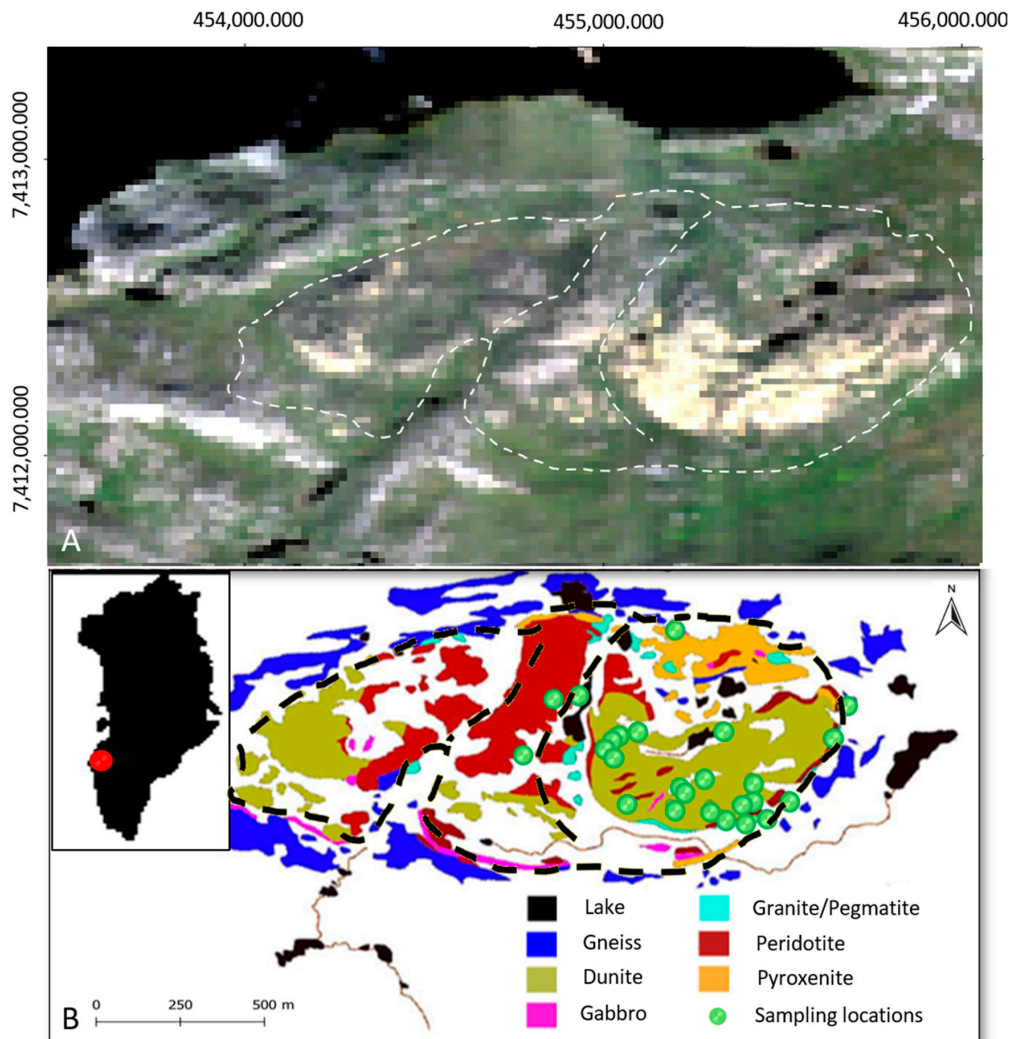

Figure 1. (**A**) True color HyMap image of the Niaqornarssuit Complex and (**B**) generalized geological map (1:10,000) of the Niaqornarssuit Complex (modified from [37]) with 23 rock sample locations (green marks) collected by the mineral exploration company 21st North Exploration and geochemically analyzed by the University of Potsdam. The white (**A**) and black (**B**) dashed polygon lines outline the ultramafic complex with a highlighted intrusion.

The gently undulating rolling terrain with moderate relief mostly lacks distinct vegetation except in low-lying south-facing slopes and depressions with 1–2 m high shrubs. However, up to 90% of the exposed bedrock surfaces are widely covered by lichens (Figure 2) [12,28,37].
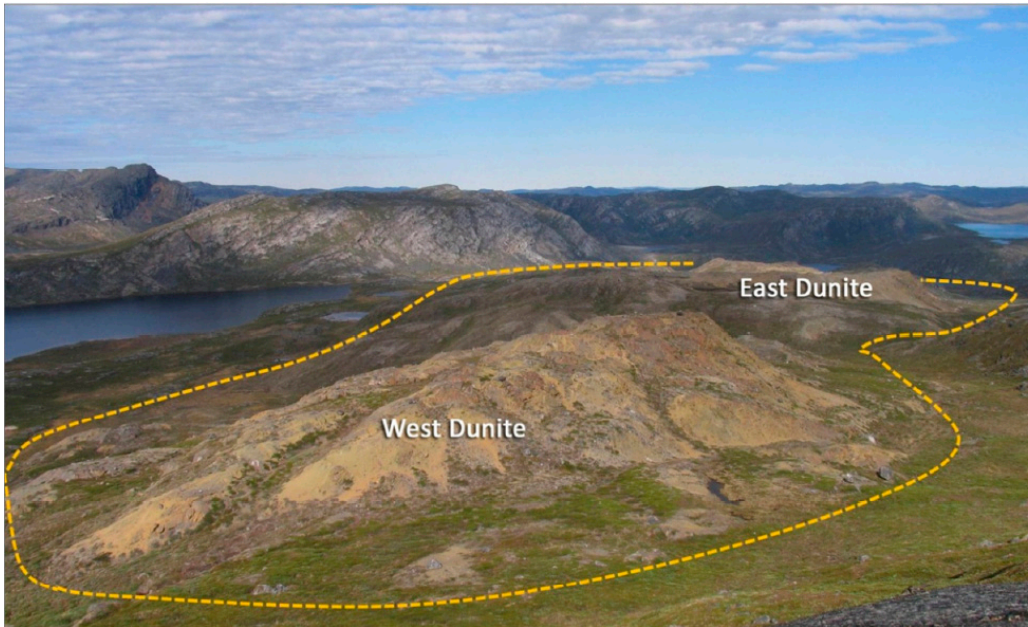
Figure 2. The Niaqornarssuit complex from the southwest. The dashed lines outline the complex of 2 × 1 km size [37].

The exploration company 21st North collected more than 160 samples from the complex and defined the stratigraphy of the complex with the following rock units [33]:

(I)   A chilled margin with a black aphanitic-fine-grained peridotite rock composition is located at the contact zone to basement gneisses (navy blue color in Figure 1) and contact-metamorphic granites (light blue color in Figure 1). This formation is 5–30 m thick, sheared, and contains a variable amount of olivine, pyroxene, and oxides;

(II)   A unit with magnetite-chromite-rich homogeneous medium-grained dunite that contains common peridotite-pyroxenite layers and intrusive dikes. The unit is mainly present in two dunite bodies (green beige color in Figure 1);

(III)   A unit that comprises medium- to coarse-grained peridotite olivine-rich at the bottom and pyroxene-rich at the upper level (maroon color in Figure 1);

(IV)   A homogenous unit of coarse-grained to pegmatitic pyroxenite that forms a massive block in the northeastern part of the complex (orange color in Figure 1);

(V)   A discontinuous layer of medium-grained and banded metagabbro (magenta color in Figure 1) interleaved with hornblende-gneiss rocks.

The complex is characterized by strongly weathered gossans and small rusty beds hosted by peridotite and pyroxenite layers at the surface. The weathered zone consists of strongly oxidized malachite-stained gossan boulders. Fresh sulfides such as pentlandite, pyrrhotite, and chalcopyrite are rarely found due to their low resistance to weathering [33,37]. These mineralizations at the surface are almost entirely restricted to the rusty beds of the eastern dunite body.

In this study, the focus of our analysis lies on II (dunite), III (peridotite), and IV (pyroxenite) rock units of the mafic-ultramafic intrusion, and we did not further consider the surrounding host rock (see Section 2). The intrusive rocks mainly consist of mafic minerals (olivine, pyroxene, and amphibole) and minor amounts of felsic minerals and other accessories.

All rock material for our laboratory measurements was collected exclusively from the eastern part of the complex since we used samples from the 21st North that concentrated their investigations on this region showing most ore concentrations.

3. Mafic and Ultramafic Rocks

Ultramafic rocks are the main component of the Earth's mantle which initially formed when the Earth differentiated into an iron-rich core and a silicate mantle [38]. Peridotites containing more than 40% olivine can be distinguished from pyroxenites characterized by an olivine content of less than 40% [39,40]. Dunites are peridotites with a very high olivine content of more than 90%. Since in this study we are analyzing both peridotites and dunites, for simplicity, peridotites are referred to as peridotites with between 40 and 90% olivine content.

Iron is the common element to characterize ultramafic rocks with optical remote sensing and the chemical element retaining magnetic properties [3]. It is the most common element in the Earth's crust and is exposed at the Earth's surface in nearly all rocks, especially in the ultramafic bodies [41] and, hence, well-suited to generally distinguish different lithologies for such ultramafic intrusions. Ferrous iron (Fe^{2+}) mainly occurs in the original mantle minerals such as olivines and pyroxenes [42]. Minerals reacting with water or atmosphere undergo a process of oxidation and alteration, resulting in the formation of iron oxide/hydroxide minerals (ferric iron Fe^{3+}) as hematite ($Fe^{3+}_2O_3$), goethite ($Fe^{3+}O(OH)$), limonite ($Fe^{3+}O(OH)\cdot nH_2O$), and magnetite ($Fe^{3+}_2Fe^{2+}O_4$) [43]. Olivine, the main component of dunite, weathers very quickly and undergoes a serpentinization process under the right atmospheric conditions producing secondary minerals such as serpentines [44].

3.1. Spectral Signatures of Ultramafic Rocks

Ferrous (Fe^{2+}) and ferric (Fe^{3+}) iron have distinct absorption features close to 1000 nm and 650 nm, respectively (Figure 3).
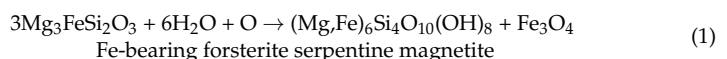
However, not all minerals containing iron show distinct absorption features within the spectral range of 400 to 2500 nm (Figure 4), such as magnetite which can be easily confused with, e.g., the Fe-bearing sulfide mineral chalcopyrite. Furthermore, magnetite is an accessory mineral in intrusive rocks and is typically not identifiable in airborne HS data.

The ferrous iron content in the ultramafic rocks is dominant and complies with the higher ferrous absorption bands (close to 1000 nm). For dunite, containing almost all olivine (having iron included as Fe^{2+}), only absorption feature at 1000 nm is exhibited [45]. Altered minerals containing ferric ions in Fe-O by oxidation exhibit a significant fall-off of the reflectance intensity in the UV-blue region (100–400 nm) of the electromagnetic spectrum [45]. Some colored Fe minerals display absorption features in visible light (400–700 nm) called color centers. These spectral characteristics are not caused by the material's chemistry but by electronic processes [5].

In this study, iron absorption features close to 850 nm related to Fe^{2+}–Fe^{3+} intervalence charge transfer are not included in the calculations. This band is not indicative of ferrous or ferric iron and overlaps with the absorption feature at 1000 nm.

3.2. Magnetism in Ultramafic Rocks

Ultramafic rocks such as pyroxenites, peridotites, and serpentinized dunites have typically high concentrations of magnetic minerals [4]. Primary ferromagnetic magnetite is the most significant magnetic mineral here, although its content in the rock is relatively low (<1 volume percent) and classified as an accessory. In addition, secondary magnetite forms during the serpentinization process of ultramafic rocks:

$$3Mg_3FeSi_2O_3 + 6H_2O + O \rightarrow (Mg,Fe)_6Si_4O_{10}(OH)_8 + Fe_3O_4$$
Fe-bearing forsterite   serpentine   magnetite

(1)
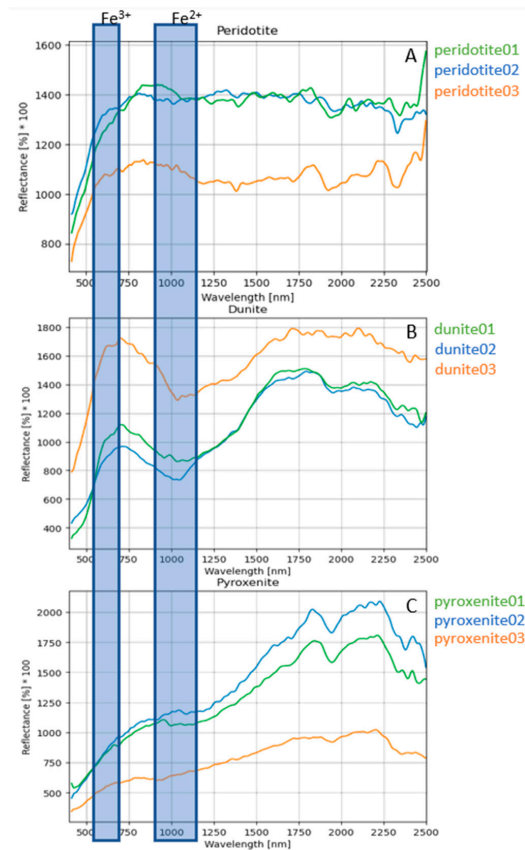
Figure 3. Laboratory examples of spectral signatures in the spectral range of 400 to 2500 nm of (**A**) peridotite (olivine 40–90%), (**B**) dunite (olivine > 90%), and (**C**) pyroxenite (olivine < 40%) of collected rock samples from the Niaqornarssuit Complex. The measurements were conducted on the weathered surfaces of the rock samples from the study site (Figure 1). Blue rectangles represent the spectral range of absorption features from ferrous (Fe^{2+}) and ferric (Fe^{3+}) iron.

Figure 4. Comparison of chalcopyrite, magnetite, and olivine spectra extracted from the splib07 USGS spectral library [46].

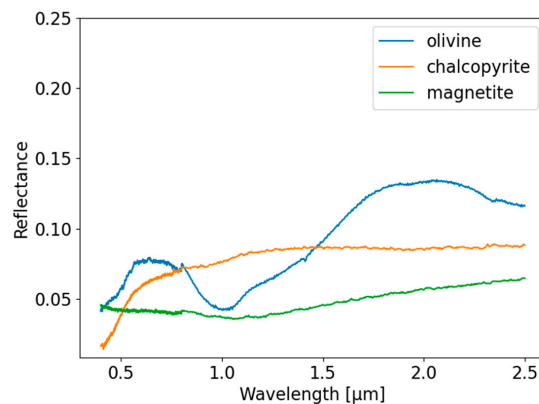Ultra-basic rocks with Fe-bearing forsterite and orthopyroxene react with water and are altered to serpentine and secondary magnetite during serpentinization [47]. The secondary magnetite amount is small and non-linearly correlated with the serpentinization degree [48,49]. However, it is relevant in the comprehension of the mechanism and evolution of the upper mantle and lower crust [50]. Magnetism increases with serpentinization until the metamorphic grade of amphibolite facies is reached [51]. At higher grades, ultramafic rocks become deserpentinized, and magnetite disappears due to the replacement of iron by other elements [3].

Since olivine and pyroxene are relatively weakly magnetic (paramagnetic: μ = 12–540 and 12–330, respectively [4]), minerals responsible for features in the spectra are not considered to be mainly responsible for stronger magnetic responses despite the significant content of olivine and pyroxene in mafic rocks [52].

4. Data Acquisition

The dataset consists of several data types analyzed in the HSM-integration approach including airborne and laboratory measurements. The airborne data comprises hyperspectral (HyMAP) and magnetic data (a local helicopterborne survey). The laboratory scale analysis includes HS scans of rock samples from the ultramafic complex using an imaging spectrometer—HySpex, and Laser-Induced Breakdown Spectrometer (LIBS) measurements to obtain the surface geochemistry.

4.1. Airborne Surveys
4.1.1. Magnetic Survey

The magnetic data were acquired as part of a combined time-domain electromagnetic and magnetic helicopter-borne survey in 2012 by Geotech Ltd. from Aurora, Canada, [34], using a horizontal magnetic gradiometer with two cesium magnetometers (Table 1).

Table 1. Magnetic gradiometer properties [34].

Magnetic Gradiometer	Horizontally Separated
Mean Altitude [m]	87
Average speed [km/h]	80
Sampling interval [s]	0.1
Sensitivity [nT]	0.001
Traverse line spacing [m]	100 and 200
Tie line spacing [m]	2000

The magnetometers were placed on a loop hanging ~24 m below the helicopter and had a horizontal distance of ~12.5 m from each other. A GPS navigation system and a radar altimeter were built-in, allowing to generate a local Digital Elevation Model (DEM) of the terrain. The helicopter flew with an average speed of 80 km/h at a mean altitude of 87 m, such that the magnetic sensors were positioned on average 63 m above the ground. Magnetic data were recorded with a sampling interval of 0.1 s. The split-beam cesium vapor magnetometer had a sensitivity of 0.001 nT. General in-line and tie-line spacings were 200 m and 2000 m, respectively, but the denser in-line spacing of 100 m was used across the mafic intrusion (Figure 5). A combined magnetometer and GPS base station was located at the Kangerlussuaq airport, about 60 km away [33].
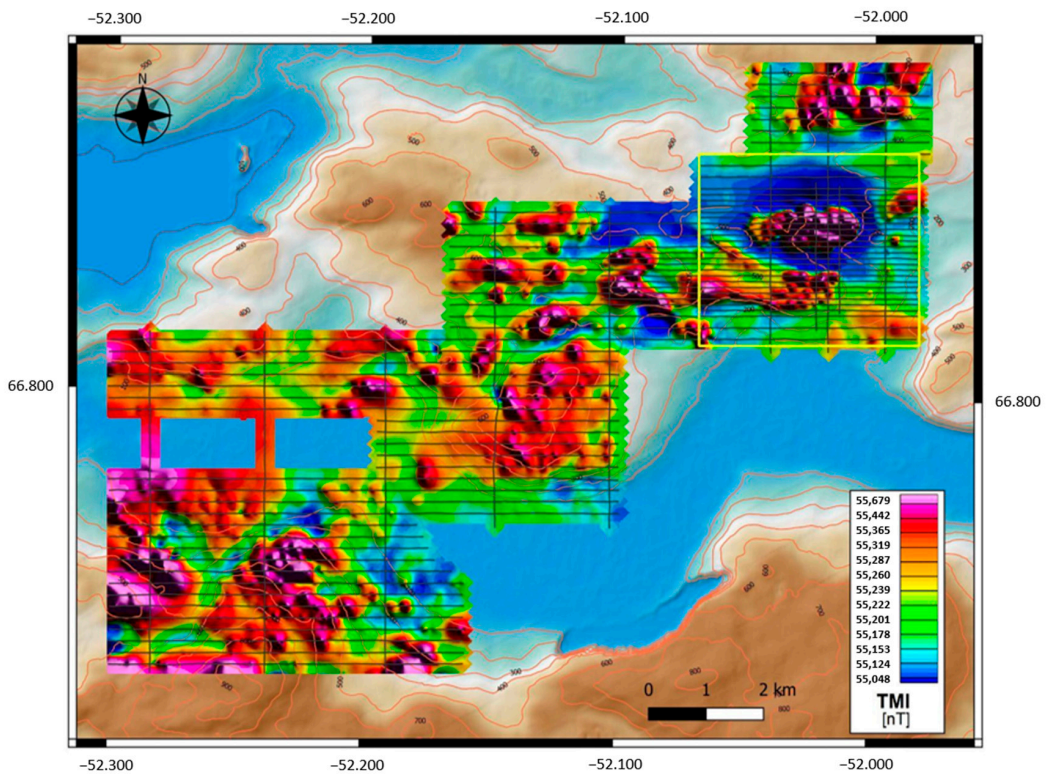
Figure 5. Total magnetic intensity map of the aeromagnetic survey. The black lines indicate the flight lines. The yellow rectangle outlines the area around a mafic intrusion considered in magnetic and hyperspectral data integration.

The total magnetic intensity map obtained through preprocessing (of the data from the two magnetometers) that included, among others, diurnal corrections and micro-leveling [34], is shown in Figure 5.

Since we focus on near-surface anomalies, we removed the impact of the core field by subtracting the IGRF (International Geomagnetic Reference Field) (box 14 in Figure 6) from these preprocessed data. Finally, we applied a 2-D Butterworth high-pass filter with a cut-off wavelength of 3000 m (filter order of 8) to a grid (cell sizes: 20 × 20 m), which was created from the residual magnetic data points by minimum curvature gridding (box 15 in Figure 6). Afterwards, the filtered values were determined at the measurement locations by means of sampling the grid.

An additional filter to prepare the magnetic data for the HSM integration could be a Reduce-To-Pole transformation that removes the effect of the Earth's magnetic field direction on the shape of magnetic anomalies from the induced magnetization. We did not apply the Reduce-To-pole correction in this study since the study area is close to the magnetic south pole; hence, the differences are minor, but the correction might be considered for surveys conducted at lower latitudes.
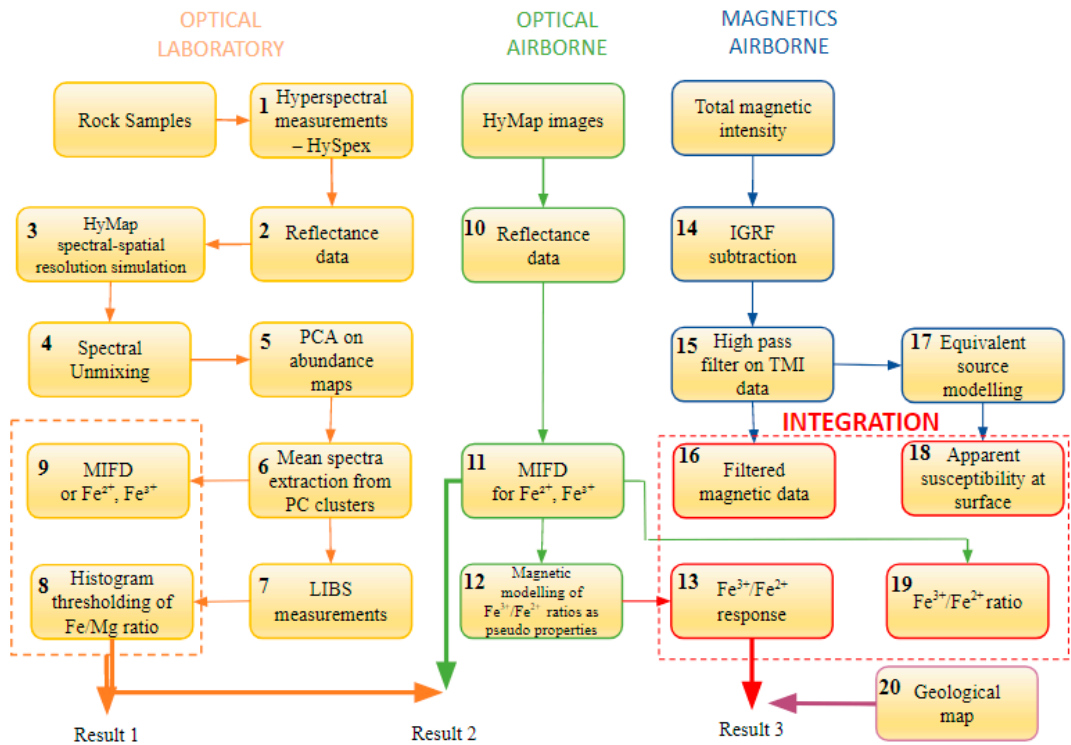
Figure 6. Generalized process chain of the proposed data integration approach.

4.1.2. Hyperspectral Survey

The HS HyMAP data were acquired in South West Greenland in 2002 [29] (Table 2).

Table 2. Key characteristics of the HyMap hyperspectral sensor [53] and HySpex cameras (AS).

Sensor	HyMap		HySpex	
			VNIR-1600	SWIR-320-e
Sensor type	hyperspectral		hyperspectral	
Altitude [m]	2500			
Setting	airborne		laboratory	
Wavelength [nm]	450–890	1950–2480	400–1000	1000–2500
	890–1350			
	1400–1800			
Bandwidth [nm]	15–16	18–20	3.7	6.25
Spatial resolution	3–10 m		24 μm	53 μm
Detector	HyMap MK 1 512 pixels		Si CCD 1600 × 1200	HdCdTe 320 × 256
FOV across track [°]	61.3		17	14
Pixel FOV across track [mrad]	2.0		0.18	0.75
Pixel FOV along track [mrad]	2.5		0.36	0.75

The whisk-broom sensor was mounted on a Piper Navajo Chieftain aircraft flown ~2500 m above the mean sea level, giving the scanner's swath width of 3000 m [53,54]. The data comprise 53 flight lines covering 7500 km². For this study, three flight lines were selected. The geocoded radiance data were converted to at-surface reflectance (box 10 in Figure 6), adjusting illumination levels using ATCOR4 software with rugged terrain optimization [11]. The features associated with atmospheric H_2O and OH close to 1400 and 1900 nm bands were excluded for further analysis.

The images with residual 106 channels were mosaicked using ENVI® version 5.6. Moreover, clouds, shadows, water, snow, ice, and steep terrain with low illumination were masked from hyperspectral images. The albedo differences in the atmospherically corrected data are caused by the bidirectional reflectance distribution effects (BRDF) [55], and the atmospheric correction causes the remaining artifacts. The final hyperspectral images had 20 × 20 m pixel sizes later used to discretize magnetic modeling and inversion (see Section 5.2).

4.2. Laboratory Measurements

We used 23 rock samples from the eastern part of the Niaqornarssuit Complex for laboratory measurements (boxes 1 to 9 in Figure 6). The samples represent the main lithological rock types in the study area, including peridotites and pyroxenites. The exact sampling locations for all rocks on a 1:10,000 geological map are shown in Figure 1. All 23 samples were scanned in the hyperspectral laboratory on the weathered sample side to imitate the real acquisition conditions of the airborne HyMap sensor that mainly measured weathered rock surfaces in the study area. In order to obtain the samples' lithogeochemistry, additional whole-rock analysis was performed.

Spectral Data

In the laboratory, the rock specimens were scanned using two HS imaging spectrometers, HySpex VNIR-1600 and SWIR-320-e sensors (Table 2). Laser-Induced Breakdown Spectrometer (LIBS) was then used to identify and quantify the first few micrometers of the sample surface's chemical components. LIBS is an analytical technique used to determine the elemental composition of materials. A focused, pulsed laser beam is directed at a sample surface, where laser energy absorption and material ablation produce high-temperature microplasma. Small amounts of the measured material are dissociated and ionized at the laser focal point, and during cooling, atomic/ionic emissions in the plasma are generated. An integrated detector was used to spectrally/temporally detect the plasma signals and record the emission lines of all elements present in the material. The resulting LIBS spectrum represents the complete chemical composition of the analyzed material [56]. LIBS data using the SciAps Z-300 was collected within the 190–950 nm wavelength range. The 50 Hz laser emits 5–6 mJ per pulse and analyzes every element in the periodic table. The LIBS measurements were taken pointwise with a fourfold shot at each homogeneous region where HySpex mean spectra were determined.

5. Proposed Method

The presented integration approach is set up to provide a first analysis suited to characterize exposed iron-rich geologies by a combination of optical and magnetic properties. The whole process chain is depicted in Figure 6. The analysis starts with comparing modified iron feature depths (MIFDs) in the laboratory and lithogeochemical information of rock samples—Result 1. In the next step, the MIFDs calculated from airborne data are compared with the MIFDs of the rock samples on a laboratory scale—Result 2. Then, airborne-based MIFDs image information is combined with airborne magnetic results by applying magnetic modeling and inversion tests. Results from the integrated analysis are finally compared with the lithology of the ultramafic complex—Result 3. It has to be noted that the laboratory results are based on measurements of only 23 rock samples that were

distributed within and nearby the eastern dunite body. Therefore, all considerations and discussions are related to this part of the ultramafic complex.

5.1. Preparation of Laboratory Optical Remote Sensing Data

The original laboratory HS raw data in 16-bit integer digital numbers (DN) from the HySpex spectrometer were transformed to radiance, followed by conversion to reflectance data (box 2 in Figure 6) using an in-house algorithm [57]. The approach is an iterative log-polar phase correlation technique based on least square regressions from Averbuch and Keller that reduces spatial non-uniformities and detects the reflection of white [58]. The reflectance retrieval also included a geometric alignment and a leveling adjustment between the HySpex system sensors—VNIR (400–1000 nm) and SWIR (1000–2500 nm) for the spectral overlap region of both sensors (VNIR and SWIR) between 940 and 1000 nm.

The absolute reflectance of laboratory measurements was calculated using the white reference's whole reflectance spectrum (VNIR and SWIR) [57]. Depending on the target's albedo, the white reference of 5, 20, 50, 90, or 95% can be chosen to resemble the reflection factor of the sample. In this way, the optimal saturation of the measurement is achieved. The extrapolated irradiance of radiance spectra in all pixels was normalized and built up the reflectance data as output. The calculated reflectance data in the laboratory were adapted to the spectral resolution of the airborne images from HyMap, simulating the HyMap properties in the laboratory (box 3 in Figure 6). The simulated laboratory data were further analyzed applying supervised spectral unmixing (box 4 in Figure 6) to identify the surface component and retrieve the fractional abundance of the pure elements at every pixel of the HS scene in each rock sample.

For spectral unmixing, the Bounded Variable Least-Squares algorithm (BVLS) was applied to find the percentage of each endmember in each pixel, assuming that measured spectra at each pixel are mixtures of different endmembers [59]. The BVLS results in abundance maps. For this approach, a user-defined spectral library was considered that consists of (1) reference spectra that were measured on pure minerals and lichen using the HySpex spectrometer (Table 2) and (2) spectra extracted from the USGS Digital Spectral Libraries: splib06 [60] and splib07a [46].

Next, the Principal Component Analysis (PCA) [61] (box 5 in Figure 6) was applied to the abundance maps to find abundance clusters in each sample. The simulated BVLS results were individually considered in the PCA using the related covariance matrix. The PCA was used to remove the redundancy in the spectral data, extract spectrally homogeneous regions within a sample, and retrieve a mean reflectance spectrum for each abundance cluster (box 6 in Figure 6). The calculated mean spectrum could be more straightforwardly used than the lower-resolution airborne-based data. It was assumed that an airborne pixel spectrum could be considered a spatio-spectral, quasi-linear mixture of multiple millimeter spectra. A PCA-based dimensional reduction of regional unmixing abundances might reflect the fractal scaling property between meter and millimeter pixel spectra, which was required in this study.

After retrieving the mean spectra for each abundance cluster, we used the LIBS device to determine the surface geochemistry of the rock sample's homogeneous regions (abundance clusters) (box 7 in Figure 6). For this purpose, LIBS was used without a beforehand prepared iron calibration curve. It has to be noted that LIBS does not provide information about ferrous and ferric iron separately. Therefore, we calculated a ratio of the primary mantle elements iron (Fe) to magnesium (Mg) for every spectrum and used it for further analysis (box 8 in Figure 6) to get a semiquantitative result. It has to be noted that for some rock samples, we extracted more than one homogeneous region resulting in multiple LIBS scans on one sample.

A new iron index was developed to relate the iron content from the LIBS measurements with the spectral signatures. Such an index was calculated for each mean spectrum for ferrous ($MIFD_{1000}$) and ferric iron ($MIFD_{650}$), respectively (box 9 in Figure 6). We assumed

that the depth of absorption feature of a given element increases with the chemical content of that element in the sample.

The new iron index—Modified Iron Feature Depth index (MIFD)—is based on the Iron Feature Depth (IFD) classification tool for hyperspectral data [62], providing the spatial distribution of iron-bearing minerals due to iron absorption features. However, our MIFD uses division instead of subtraction to avoid negative values in ferrous and ferric iron ratios (Equations (2) and (3)) and to suppress remaining micro shadow effects that impact the at-surface reflectance retrieval. The MIFD algorithm is based on the left λ_{left} and the right shoulder wavelength λ_{right} of a feature, its absorption center wavelength λ_{center}, and corresponding reflectance values $r_{left}, r_{right}, r_{center}$, respectively (Figure 7).
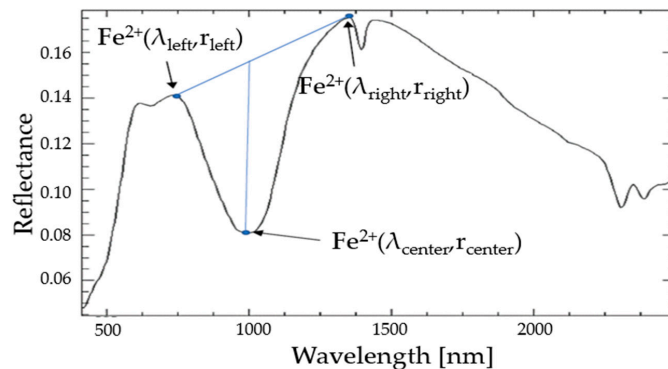

Figure 7. Schematic illustration of the MIFD$_{1000}$ parameter estimation on a spectrum example of a peridotite sample (analogous to MIFD$_{650}$).

The depth is estimated by setting an interpolated continuum between both shoulders of the absorption feature and calculating the quotient between the interpolated line $line_{int}$ and given absorption center r_{center} by the following formula:

$$line_{int} = r_{left} + \left(r_{right} - r_{left}\right) \frac{\lambda_{center} - \lambda_{left}}{\lambda_{right} - \lambda_{left}} \tag{2}$$

$$MIFD = \frac{line_{int}}{r_{center}} \tag{3}$$

For the *MIFD* retrieval in this study, the iron spectral features close to 1000 nm (related to ferrous iron) and 650 nm (related to ferric iron) are set up as absorption centers (Figure 7).

MIFD$_{650}$ and MIFD$_{1000}$ absorption depth estimates were calculated on all 23 rock samples (approximately on two regions of each sample) measured by HySpex and from the airborne HS HyMap image at the ultramafic complex (box 11 in Figure 6). Depending on the optical sensor, the closest bands to 650 nm and 1000 nm are selected as the centers for absorption features r_{center}. The MIFD at 1000 nm is considered as more reliable for airborne-based HS data than the 650 nm band since the 650 nm is positioned close to the end of the spectral region, which may distort the results. In contrast, the MIFD at 650 nm iron feature is more reliable than the one at 1000 nm in the HS data obtained by HySpex under laboratory conditions. This is due to the detector jump effect mainly caused by a misalignment between the spectra of VNIR and SWIR detectors. To achieve semiquantitative information about the correlation between MIFD$_{650}$ and MIFD$_{1000}$ and Fe/Mg-ratio, the Fe/Mg-ratio values determined for rock samples in the laboratory were used as validation points since the locations of the analyzed samples are known (Figure 1).

5.2. Magnetic Forward Modeling and Inversion for the Integration of Hyperspectral and Magnetic Data

To combine the magnetic and hyperspectral information from the air-/helicopterborne surveys with each other, we choose two strategies—a magnetic forward modeling approach and a magnetic inversion approach. For both strategies, we limited our investigation to a rectangle area of 3820 × 3300 m with the Niaqornarssuit Complex in its center (see yellow polygon in Figure 5), and we determined the ratio of MIFD ($MIFD_{650}/MIFD_{1000}$) at each pixel of the HS image having cell sizes of 20 × 20 m. In both approaches, we used the high-resolution DEM determined from the magnetic helicopter survey to describe the topography in our models (Figure 8C), and we considered the same magnetic data (residual magnetic anomalies after IGRF correction and high pass-filtering; see Section 4.1.1. and Figure 8A) with their exact positions and heights for the analyses.
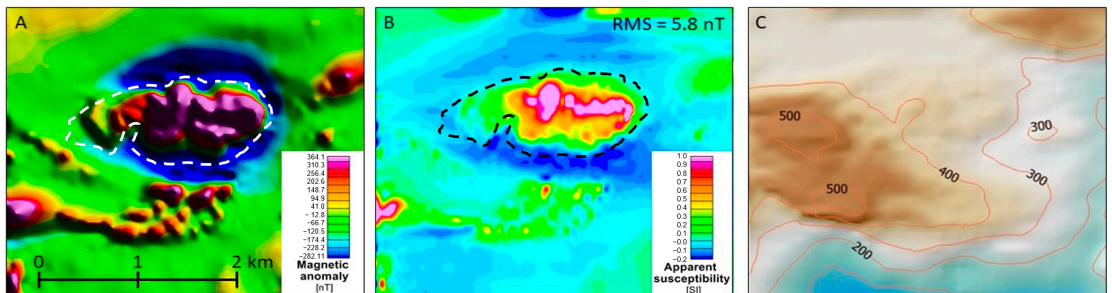

Figure 8. (**A**) Residual magnetic anomaly map after high pass filtering with a cut-off of 3000 m. (**B**) Inversion results from an equivalent dipole layer modeling. The dipoles are arranged in a uniformly spaced grid along the surface topography and are located in the center of the HS pixels. The precise DEM derived from the airborne survey (**C**) is used as the topography estimate. Dashed lines outline the shape of the ultramafic complex.

We used the software Oasis Montaj [63] for the magnetic forward approach. The model consisted of a single layer of cuboidal cells (number of cells: 192 × 166) that were arranged along the topography and had precisely the same horizontal extent and locations as the HS pixels and a small vertical dimension of Δz = 5 m.

We filled modeling cells along the surface with the MIFD values, conducted magnetic modeling, and calculated the responses at the data points. We assumed that the iron ratios represent some pseudo-magnetization (see box 12 in Figure 6). Since the magnetization direction in all cells was aligned parallel to the Earth Magnetic Field, it was considered that the magnetizations had only an induced (susceptibility) component. The amount of lichen coverage was considered in the modeling by weighting the MIFDs estimates by its rock surface percentage. For the $MIFD_{650}$ parameter, the lichen cover of the study area was increased by 30% to reduce the lichen influence on the spectral information of rocks.

We compared these modeled pseudo-magnetic responses with the processed and filtered residual magnetic anomaly data in cross-plots (see boxes 13 and 16 in Figure 6) and assigned a color code that represents the mapped lithology immediately underneath the data point locations.

Since two data responses and, hence, equivalent entities are compared in the modeling, this approach is preferable to a simple comparison of the actual MFID values at the ground and the residual magnetic anomaly data, where the former is directly associated with physical properties. However, the latter is a data response (measured at some flight height), not a physical parameter at the ground.

A drawback of this easily accomplishable modeling approach is that total responses at the data points are derived from a linear combination of the individual responses from the magnetizations of the different cells. Accordingly, the response at a data point is not only affected by the magnetization immediately underneath but also from other

locations. However, since the magnetic field rapidly decays with distance (between $1/r^2$ and $1/r^3$, where r is the distance between the source location and data point) and since flight altitudes were shallow (see Table 1), it is appropriate to assume that the data response is predominantly associated with the magnetization of rocks located underneath.

We applied a magnetic inversion approach to account for this limitation and directly relate magnetic properties with the hyperspectral iron features along the surface topography (box 17 in Figure 6). In this approach, we estimated apparent magnetic susceptibilities for a layer of magnetic dipole sources (e.g., [63,64]) that were arranged in a regularly spaced grid along the surface topography, where the dipole locations coincided with the pixel centers of the HS images (192 × 166 dipole sources with a grid spacing of 20 m both in x- and y-direction). The dipoles were oriented parallel to the Earth's magnetic field, such that it was assumed that the magnetization had no remanent component whose direction deviated from one of the induced magnetizations. We used the same filtered residual magnetic anomaly data as input data for the forward modeling approach.

The apparent susceptibilities were determined using a self-developed Python script with an iterative deterministic Gauss–Newton inversion scheme. A regularization term in the inversion ("smoothing"; e.g., [65]) stabilized the inversion results and ensured that apparent susceptibility values only gradually changed in areas with little data coverage and, hence, model resolution. The weight of the regularization term λ was carefully selected and decreased with the number of iterations (λ ranges from 1.0 to 0.4) to ensure that the resulting data misfit was reasonably low and little impacted by the regularization.

The final inversion result (Figure 8B) after 50 iterations had a low RMS misfit of 5.8 nT. The magnetization estimate at the ground allows a direct comparison with the MFID values in cross-plots (boxes 18 and 19 in Figure 6) and an exact spatial assignment of the lithological units.

Although an (apparent) susceptibility distribution was assigned directly to the surface, these estimates still suffer from inaccuracies due to the limited resolution of the magnetic method. The resolution in horizontal directions is estimated to be in the range of 50–150 m by considering a dense flight line pattern and shallow flight heights (see Section 4.1.1). A more precise estimate of the resolution could be determined for each pixel by analyzing the resolution matrix or the model covariance matrix [66], which is beyond the scope of this conceptual study. Further inaccuracies are introduced by the fact that some contributions of the dipole source values are still associated with magnetic sources that are not located immediately at the surface despite the high-pass filter that generally reduces the impact of deeper-seated magnetic sources. Finally, the remanent part of the magnetization that points in a different direction than the Earth Magnetic Field is not considered in the modeling such that the response of this part is assigned to the induced magnetization component. This means that the determined apparent susceptibilities, which were finally compared with the Fe ratios in the cross-plots, are a mixture of reduced and remanent magnetization contributions.

6. Results

6.1. Fe/Mg Ratio and Modified Iron Feature Depth Index

Result 1—laboratory scale: The first preliminary result presents the Fe/Mg ratio plotted against the estimated sum of ferrous ($MIFD_{650}$) and ferric ($MIFD_{1000}$) modified iron feature depths for dunites (red) and peridotites (blue) (Figure 9).

The Fe/Mg content ratio was calculated on the basis of the LIBS measurements on hand specimens to display the dependency of the iron content on its absorption feature depths. The absorption features were extracted from the hand specimens measured in the laboratory in the simulated HyMap resolution (coarser spatial and spectral resolution of spectra). Figure 9 shows an overall linear trend indicating that the absorption feature depth tends to increase with increasing Fe-content or decreasing Mg-content for dunites and peridotites.
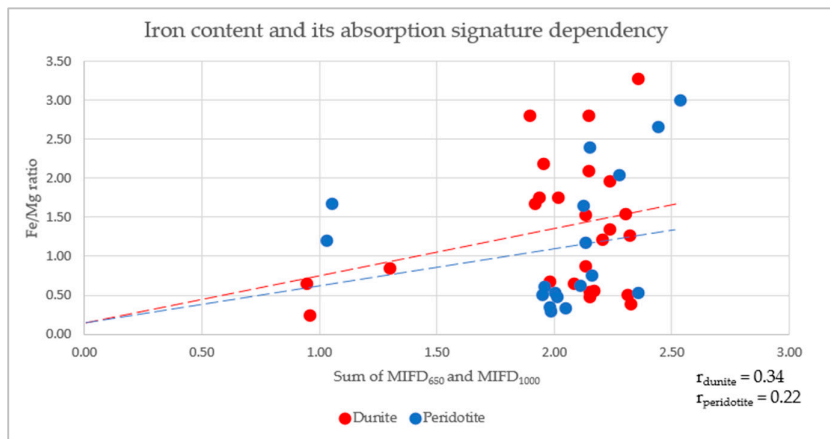
Figure 9. Fe/Mg ratio versus MIFD of summed ferrous and ferric iron absorption depths for dunites (red) and peridotites (blue) calculated for HyMap spectral resolution in the laboratory. r values represent the correlation coefficients for each lithology. Dashed lines show the fitted linear trendlines for each lithology with the corresponding color.

Result 2—airborne scale: Analogously, ferrous and ferric MIFDs were calculated for airborne HyMap scenes of the ultramafic complex. The Fe/Mg ratio values from rock samples were plotted against the ferric $MIFD_{650}$ (Figure 10) and ferrous $MIFD_{1000}$ (Figure 11) from HyMap data. Both figures show a robust, increasing linear trend for peridotite with a large correlation coefficient of 0.85 and 0.68) and a more vague decreasing linear trend for dunite (with lower correlation coefficients of -0.23 and 0.14).
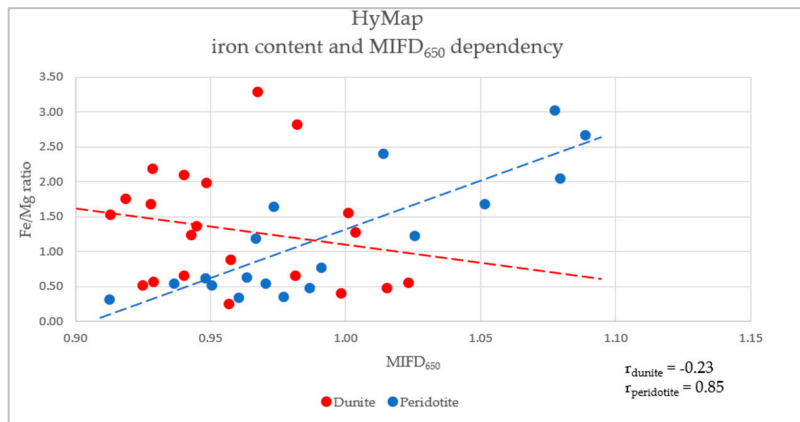

Figure 10. In-situ Fe/Mg ratio of the validation points versus $MIFD_{650}$ for HyMap data. Dashed lines show the fitted linear trendline for each lithology with the corresponding color. r values represent the correlation coefficient for each lithology.

The dunite body in the eastern part of the ultramafic complex shows a high abundance of large $MIFD_{1000}$ values (dark red color), whereby $MIFD_{650}$ values are lower abundant (light green color) in this area compared to its surroundings (Figure 12).
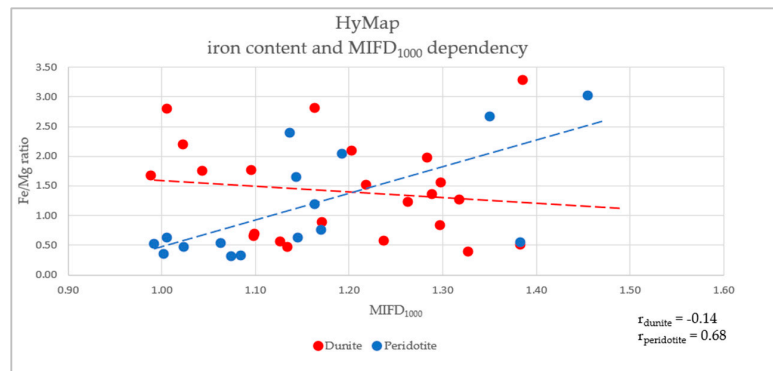
Figure 11. In-situ Fe/Mg ratio of the validation points versus $MIFD_{1000}$ for HyMap data. Dashed lines show the fitted linear trendlines for each lithology with the corresponding color. r values represent the correlation coefficient for each lithology.
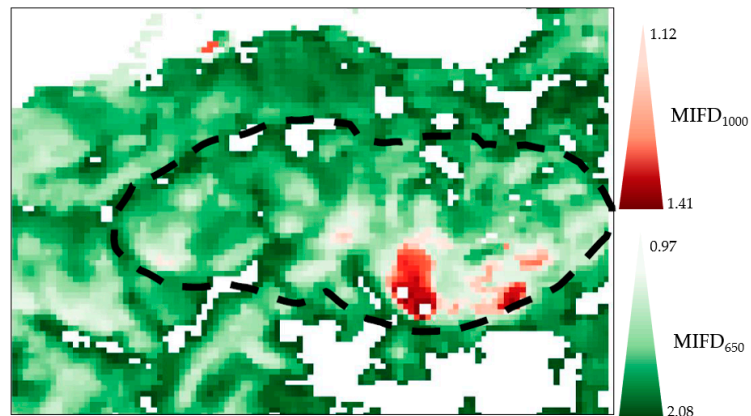

Figure 12. Results of the $MIFD_{650}$ and $MIFD_{1000}$ calculation for the HyMap scene. The red and green colors represent $MIFD_{1000}$ and the $MIFD_{650}$ associated with ferrous and ferric iron, respectively. The black dashed line outlines the study area—the Niaqornarssuit Complex.

Fe^{3+} is the product of Fe^{2+} oxidation, forming secondary iron oxides/hydroxides on weathered rock surfaces. The generally high values for Fe^{3+} at 650 nm are partly associated with the oversaturation of the absorption band by the contribution of vegetation and lichen that partially covers the exposed rock surfaces in the study area. The combination of low $MIFD_{650}$ and high $MIFD_{1000}$ of the eastern dunite block compared to the remaining part of the ultramafic complex suggests that this area is less affected by alteration.

6.2. Integration of Hyperspectral and Magnetic Airborne Data

Result 3—airborne scale: The relationship between the residual magnetic anomaly data, ferrous $MIFD_{1000}$, and ferric $MIFD_{650}$ were analyzed by the forward modeling and inversion results approach. The predicted pseudo-magnetic responses from the Fe ratio ($MIFD_{650}/MIFD_{1000}$) (Figure 13B) are plotted against the high pass filtered residual magnetic anomaly data (Figure 13C) for the modeling approach in Figure 13A. The lithological information in Figura 13A is extracted from a geological map prepared by the company 21st North (Figure 1B).
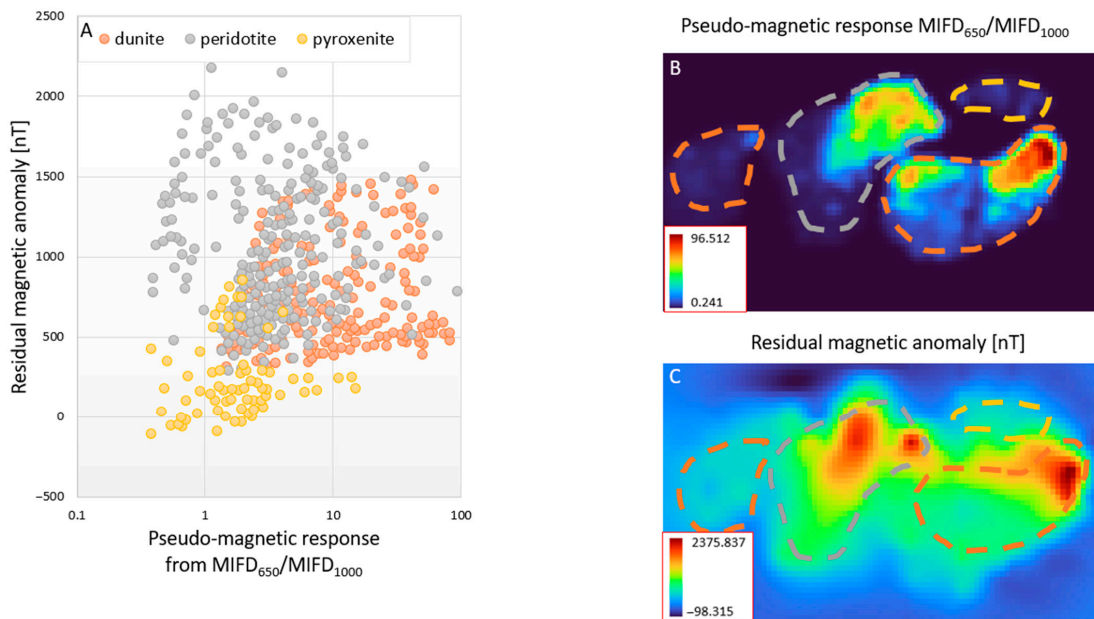
Figure 13. (**A**) Correlation plot between the residual magnetic anomaly data (after IGRF removal and high-pass filtering) and the pseudo-magnetic responses calculated by forward of $MIFD_{650}/MIFD_{1000}$-ratios that are considered as "susceptibility" values at the surface. The lithologies located immediately underneath data point locations are color-coded: dunites (orange), peridotites (gray), and pyroxenites (yellow) (**B,C**) show the calculated pseudo-magnetic response of the $MIFD_{650}/MIFD_{1000}$-ratio and the residual magnetic anomaly map, respectively. The dashed lines sketch the shapes of mapped dunites (orange), peridotites (gray), and pyroxenites (yellow) lithological units.

7. Discussion

This study introduced a novel and robust approach for the geological characterization of an ultramafic complex with elevated magnetic properties based on integrating optical and magnetics datasets. This was achieved by establishing relationships between the residual magnetic anomaly data, ferric/ferrous ratios built from iron absorption feature indices, and lithological information (Figure 13). Primary lithologies in the complex are classified into three lithological groups (dunites, peridotites, and pyroxenites) that can be clearly distinguished based on how magnetic responses and susceptibility estimates correlate with iron absorption features from airborne-based hyperspectral data.

Magnetization depends on the primary magnetite's rock composition and oxidation state [47]. As expected, areas dominated by peridotite are associated with the highest magnetization compared to areas dominated by dunite and pyroxenite. Lower magnetization values for dunites could indicate that the minerals in dunite are less ferruginous than in peridotites, thus containing more Mg.

Furthermore, the iron in dunite is also associated with chromite, which is not influenced by serpentinization and prevents the iron from being released [47]. Chromite has ferromagnetic properties [4], but the magnetization is lower than in magnetite.

Olivine weathers very quickly at and close to the surface and may form paramagnetic hematite that contains only ferric iron [67]. Ferric iron is stable under atmospheric conditions; therefore, the magnetization is lower for strongly weathered than for fresh rocks. In the plot, pyroxenites are weakly magnetized even at a high ferric/ferrous ratio. However, the magnetization increases slightly with increasing ferric iron content.

Considering the relationships in Figure 13, it becomes clear that it is partly difficult to distinguish dunite, peridotite, and pyroxenite-dominated rocks from magnetic data only. Similarly, it is impossible to distinguish pyroxenite and peridotite-dominated rocks only from hyperspectral iron ratios. However, separated clusters for the three lithologies can be observed by the combinations of both data types indicating that lacking diagnostic characteristics in one of the methods are provided by the other method.

The results highlight the benefit of the newly developed approach of combining the MIFD and residual magnetic anomaly data to both (1) detect which lithology is dominated by ferromagnetic magnetite found disseminated as an accessory mineral and (2) estimate how the oxidation state is related to the lithologies at the same time. Our results show that the correlation between residual magnetic anomaly, iron absorption features, and lithology (HSM integration) can be found in proxies. Furthermore, resolution enhancement of the residual magnetic anomaly data helps to unravel nondominant anomalies.

HS data acquired under laboratory settings provide detailed spectral information about the lithological units and rock types in the complex. Against that, a slight change in the illumination or the measuring angle and spectrometer-dependent measurement errors can bias the laboratory results [68,69]. Furthermore, reflectance spectra of an object acquired with different sensors may differ. A reliable mineral classification of HS data in the laboratory should rule out these error sources by applying a consistent spectral library of pure minerals measured using the same instrument utilized for sample analysis [70,71]. One step that needs to be followed before interpreting HS images consisting of two sensors with different spectral ranges (VNIR, SWIR) is the leveling adjustment. Due to this jump effect at 1000 nm, the result of combined VNIR and SWIR detectors can alternate the spectrum [57]. No detector jump is known in the case of HyMap, although the 650 nm absorption band is positioned close to the end of the spectral range, oversaturating the 650 nm absorption feature. Moreover, the 1000 nm band is more pronounced than the 650 nm band in almost all extracted spectra. The reason could be the spectral width of the absorption band close to 1000 nm that can be superimposed by the flanks of water vapor absorption features at 940 and 1130 nm.

It must be noted that the current approaches to combining airborne magnetic and hyperspectral data have several limitations. The magnetic forward modeling and inversion approaches assume that the magnetization only consists of an induced magnetization component (susceptibility) but not of a remanence component. In addition, the resolution of the magnetic method is limited such that the pixel values obtained from the magnetic modeling and inversion approaches have some inaccuracies and, hence, details in the cross-plots are not entirely correct. Another limitation is that a lichen spectrum has an absorption band close to 600 nm and could change the ferric iron absorption feature; moreover, lichen grows on specific lithologies such as peridotites, where the lichen can deteriorate spectral signature in the 400–2000 nm spectral region. In addition, the spectral signatures of ferrous (1000 nm) and ferric (650 nm) iron are distinctive due to the limited lichen cover in this range. Another iron absorption feature at 880 nm has been neglected in the analysis due to the high lichen presence in this spectral range. Therefore, an optimal weighting of the iron features and lichen presence is critical [12].

At multiple scales, e.g., laboratory, airborne-based, satellite-based, and spectral, information may not always be comparable [72]. Our results show that, regardless of scale, both laboratory and airborne-based iron absorption features are stable, making the HSM integration and validation of the results more accessible and feasible.

8. Conclusions and Outlook

This study proposed a novel method to integrate optical remotely sensed and aeromagnetic data using forward modeling and inversion approaches. This study is the first step in the geological analysis of ultramafic complexes based on multidisciplinary research with a strong focus on non-invasive mapping and remote sensing methods. This approach can help to improve the knowledge about the study site's regional lithology and distribution

of rocks and minerals by combining the partly complementary information from hyperspectral image products (Fe iron ratios) and magnetic properties. It is a promising way to analyze a prospect area with high iron-bearing mineral potential and can be considered as a starting point to integrate both data types by performing a more advanced hyperspectral constrained magnetic inversion. Our study on the ultramafic complex demonstrates that detailed airborne HSM-integration results can enhance our understanding of the geological origins of magnetic anomalies. It appears to be particularly useful to thoroughly characterize the distribution of Fe minerals, e.g., for iron-oxide and sulfide deposits.

Such a deposit must retain magnetic properties induced by the presence of both ferrous and ferric iron in minerals. We have shown that the HSM integration can be utilized based on chemical and physical rock properties presenting that the absorption feature depths of ferrous and ferric iron rise with the increasing ferrous and ferric iron amount in the analyzed dataset. The main benefit of the HSM integration is a significantly higher spatial resolution of the hybrid model and, thus, a better exploration possibility.

In our integration of imaging spectroscopy and aeromagnetics at an airborne scale, we used only the near-surface information in the forward modeling and inversion approaches. However, more research is needed to integrate hyperspectral imaging information into magnetic modeling or inversion approaches that consider models of both the near surface and at larger depths and are suited for complete geological characterization. Moreover, it is essential to find proper solutions to incorporate hyperspectral information as meaningful constraints in magnetic inversions. Such constrained inversions will improve the resolution of the magnetic models at shallow depths, making results from cross-plot more reliable and precise.

As a result, several theoretical assumptions and practical considerations should be considered before and while applying this approach:

1. We suggest collecting samples of the whole investigation area, including information about the sample's orientation in the ground, to analyze the laboratory's susceptibility and remanent magnetization. The rock samples should be geochemically analyzed, focusing on whole-rock analysis and titration to determine the rock's ferric and ferrous iron content;
2. Since the magnetic properties can only be regionally correlated with the lithology, more research in a different climate and diverse iron-bearing deposits considering new parameters should prove our approach's robustness.

Accordingly, a precise correlation of magnetic and HS properties is only possible if a proper estimate of the magnetization along the surface is determined using a modeling inversion technique. We can imagine that a magnetic inversion approach constrained by HS surface information is a possible strategy that can solve such problems in the future and better link the lithology at the surface obtained from hyperspectral with the magnetic anomalies determined in the (shallower) subsurface.

Author Contributions: Conceptualization, A.K., S.S., B.H.H. and C.R.; methodology, A.K., B.H.H., S.S., N.K., C.R.; software, A.K., C.R., C.M.; validation, B.H.H., C.R., N.K.; formal analysis, A.K., S.S., B.H.H., C.M.; investigation, A.K., N.K., S.S.; resources, S.S., C.R.; data curation, A.K.; writing—original draft preparation, A.K., B.H.H.; writing—review and editing, B.H.H., S.S., I.B.; visualization, A.K., B.H.H.; supervision, S.S., C.R., U.A., I.B. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

References

1. Mielke, C.; Muedi, T.; Papenfuss, A.; Boesche, N.K.; Rogass, C.; Gauert, C.D.K.; Altenberger, U.; Wit, M.J.D. Multi- and hyperspectral spaceborne remote sensing of the Aggeneys base metal sulphide mineral deposit sites in the Lower Orange River region, South Africa. *S. Afr. J. Geol.* **2016**, *119*, 63–76. [CrossRef]
2. Jackisch, R.; Madriz, Y.; Zimmermann, R.; Pirrtijärvi, M.; Saartenoja, A.; Heincke, B.H.; Salmirinne, H.; Kujasalo, J.-P.; Andreani, L.; Gloaguen, R. Drone-Borne Hyperspectral and Magnetic Data Integration: Otanmäki Fe-Ti-V Deposit in Finland. *Remote Sens.* **2019**, *11*, 2084. [CrossRef]
3. Dentith, M.; Mudge, S.T. *Geophysics for the Mineral Exploration Geoscientist*; Cambridge University Press: Cambridge, UK, 2014.
4. Clark, D.A. Magnetic petrophysics and magnetic petrology: Aids to geological interpretation of magnetic surveys. *J. Aust. Geol. Geophys.* **1997**, *17*, 83–104.
5. Hunt, C.P.; Moskowitz, B.M.; Banerjee, S.K. Magnetic properties of rocks and minerals. In *American Geophysical Union*; Geological Society of America: Boulder, CO, USA, 1995.
6. Till, J.L.; Nowaczyk, N. Authigenic magnetite formation from goethite and hematite and chemical remanent magnetization acquisition. *Geophys. J. Int.* **2018**, *213*, 1818–1831. [CrossRef]
7. Dekkers, M. Magnetic properties of natural goethite-III. Magnetic behaviour and properties of minerals originating from goethite dehydration during thermal demagnetization. *Geophys. J. Int.* **1990**, *103*, 233–250. [CrossRef]
8. Oëzdemir, O.; Dunlop, D.J. Intermediate magnetite formation during dehydration of goethite. *Earth Planet. Sci. Lett.* **2000**, *177*, 59–67. [CrossRef]
9. Hanesch, M.; Stanjek, H.; Petersen, N. Thermomagnetic measurements of soil iron minerals: The role of organic carbon. *Geophys. J. Int.* **2006**, *165*, 53–61. [CrossRef]
10. Ager, C.M.; Milton, N.J.G. Spectral reflectance of lichens and their effects on the reflectance of rock substrates. *Geophysica* **1987**, *52*, 898–906. [CrossRef]
11. Salehi, S.; Thaarup, S.M. Mineral mapping by hyperspectral remote sensing in West Greenland using airborne, ship-based and terrestrial platforms. *Geol. Surv. Den. Greenl.* **2017**, *41*, 47–50. [CrossRef]
12. Salehi, S.; Mielke, C.; Rogass, C. Mapping ultramafic complexes using airborne imaging spectroscopy and spaceborne data in Arctic regions with abundant lichen cover, a case study from the Niaqornarssuit complex in South West Greenland. *Eur. J. Remote Sens.* **2020**, *53*, 156–175. [CrossRef]
13. Rasti, B.; Hong, D.; Hang, R.; Ghamisi, P.; Kang, X.; Chanussot, J.; Benediktsson, J.A. Feature Extraction for Hyperspectral Imagery: The Evolution from Shallow to Deep (Overview and Toolbox). *IEEE Geosci. Remote Sens. Lett.* **2020**, *8*, 60–88. [CrossRef]
14. Kopackova, V.; Koucka, L. Integration of absorption feature information from visible to longwave infrared spectral ranges for mineral mapping. *Remote Sens.* **2017**, *9*, 1006. [CrossRef]
15. Kruse, F. Integrated visible and near-infrared, shortwave infrared, and longwave infrared full-range hyperspectral data analysis for geologic mapping. *J. Appl. Remote Sens.* **2015**, *9*, 096005. [CrossRef]
16. McDowell, M.L.; Kruse, F. Enhanced compositional mapping through integrated full-range spectral analysis. *Remote Sens.* **2016**, *8*, 757. [CrossRef]
17. Notesco, G.; Ogen, Y.; Ben-Dor, E. Integration of hyperspectral shortwave and longwave infrared remote-sensing data for mineral mapping of Makhtesh Ramon in Israel. *Remote Sens.* **2016**, *8*, 318. [CrossRef]
18. Kuras, A.; Brell, M.; Rizzi, J.; Burud, I. Hyperspectral and Lidar Data Applied to the Urban Land Cover Machine Learning and Neural-Network-Based Classification: A Review. *Remote Sens.* **2021**, *13*, 3993. [CrossRef]
19. Buckley, S.J.; Kurz, T.H.; Howell, J.A.; Schneider, D.J.C. Terrestrial lidar and hyperspectral data fusion products for geological outcrop analysis. *Comput. Geosci.* **2013**, *54*, 249–258. [CrossRef]
20. Kirsch, M.; Lorenz, S.; Zimmermann, R.; Tusa, L.; Möckel, R.; Hödl, P.; Booysen, R.; Khodadadzadeh, M.; Gloaguen, R. Integration of Terrestrial and Drone-Borne Hyperspectral and Photogrammetric Sensing Methods for Exploration Mapping and Mining Monitoring. *Remote Sens.* **2018**, *10*, 1366. [CrossRef]
21. Kirsch, M.; Lorenz, S.; Zimmermann, R.; Andreani, L.; Tusa, L.; Pospiech, S.; Jackisch, R.; Khodadadzadeh, M.; Ghamisi, P.; Unger, G.; et al. Hyperspectral outcrop models for paleoseismic studies. *Photogramm. Rec.* **2019**, *34*, 385–407. [CrossRef]
22. Bedini, E.; Rasmussen, T.M. Use of airborne hyperspectral and gamma-ray spectroscopy data for mineral exploration at the Sarfartoq carbonatite complex, southern West Greenland. *Geosci. J.* **2018**, *22*, 641–651. [CrossRef]

23. Jackisch, R.; Lorenz, S.; Kirsch, M.; Zimmermann, R.; Tusa, L.; Pirttijärvi, M.; Saartenoja, A.; Ugalde, H.; Madriz, Y.; Savolainen, M.; et al. Integrated Geological and Geophysical Mapping of a Carbonatite-Hosting Outcrop in Siilinjärvi, Finland, Using Unmanned Aerial Systems. *Remote Sens.* **2020**, *12*, 2998. [CrossRef]
24. Bedini, E. Mapping lithology of the Sarfartoq carbonatite complex, southern West Greenland, using HyMap imaging spectrometer data. *Remote Sens. Environ.* **2009**, *113*, 1208–1219. [CrossRef]
25. Bedini, E. Mineral mapping in the Kap Simpson complex, central East Greenland, using HyMap and ASTER remote sensing data. *Int. J. Remote Sens.* **2012**, *33*, 939–961. [CrossRef]
26. Budkewitsch, P.; Staenz, K.; Neville, R.A.; Sangster, D. Spectral signatures of carbonate rocks surrounding the Nanisivik MVT Zn-Pb mine and implications of hyperspectral imaging for exploration in Arctic environments. In Proceedings of the Ore Deposit Workshop: New Ideas for a New Millennium, Cranbrook, BC, Canada, 5–6 May 2000.
27. Harris, J.R.; Rogge, D.; Hitchcock, R.; Ijewliw, O.; Wright, D. Mapping lithology in Canada's Arctic: Application of hyperspectral data using the minimum noise fraction transformation and matched filtering. *Can. J. Earth Sci.* **2005**, *42*, 2173–2193. [CrossRef]
28. Salehi, S.; Lorenz, S.; Sørensen, E.V.; Zimmermann, R.; Fensholt, R.; Heincke, B.H.; Gloaguen, R. Integration of vessel-based hyperspectral scanning and 3D-photogrammetry for mobile mapping of steep coastal cliffs in the arctic. *Remote Sens.* **2018**, *10*, 175. [CrossRef]
29. Tukiainen, T.; Thorning, L. Detection of kimberlitic rocks in West Greenland using airborne hyperspectral data: The HyperGreen 2002 project. *Greenl. Bull. Geol. Surv. Den.* **2005**, *7*, 69–72. [CrossRef]
30. Tukiainen, T.; Thomassen, B. Application of airborne hyperspectral data to mineral exploration in North-East Greenland. *Greenl. Bull. Geol. Surv. Den.* **2010**, *20*, 71–74. [CrossRef]
31. Jackisch, R.; Heincke, B.H.; Zimmermann, R.; Sørensen, E.V.; Pirttijärvi, M.; Kirsch, M.; Salmirinne, H.; Lode, S.; Kuronen, U.; Gloaguen, R. Drone-based magnetic and multispectral surveys to develop a 3D model for mineral exploration at Qullissat, Disko Island, Greenland. *Solid Earth* **2022**, *13*, 793–825. [CrossRef]
32. Miller, C.A.; Schaefer, L.N.; Kereszturi, G.; Fournier, D. Three-Dimensional Mapping of Mt. Ruapehu Volcano, New Zealand, From Aeromagnetic Data Inversion and Hyperspectral Imaging. *J. Geophys. Res. Solid Earth* **2020**, *125*, e2019JB018247. [CrossRef]
33. Simard, R.L.; Bliss, I.; Vaillancourt, C. *Geological Report on Exploration and Drill Programs 2013—Licenses 2010/17, 2013727 and 2013/28*; NorthernShield Ressources Inc.: West Greenland, Denmark, 2014; p. 43.
34. Geotech. *Report on a Helicopter-Borne Versatile Time-Domain Electromagnetic (VTEMplus) and Horizontal Magnetic Gradiometer Geophysical Survey*; Niaqomarssuit Block: Greenland, Denmark, 2012.
35. Gool, J.A.M.V.; Connelly, J.N.; Marker, M.; Mengel, F.C. The Nagssugtoqidian Orogen of West Greenland: Tectonic evolution and regional correlations from a West Greenland perspective. *NRC Res. Press Web.* **2002**, *39*, 665–686.
36. Gothenborg, J.; Keto, L. Report on the aerial reconnaissance between Sukkertoppen Ice Calot and Nordenskiölds Gletscher. In *Archives of Geological Survey of Denmark and Greenland*; GEUS Report File 20210; Kryolitselskabet Øresund A/S: Copenhagen, Denmark, 1977.
37. Østergaard, C. *21st North—2010 Field Work Qaqortorsuaq (Ikertoq)*, 2011; p. 99.
38. Downes, H. Ultramafic Rocks. In *Encyclopedia of Geology*, 2nd ed.; Alderton, D., Elias, S.A., Eds.; Academic Press: Cambridge, MA, USA, 2021; pp. 69–75.
39. Streckeisen, A.L. Plutonic rocks, classification and nomenclature recommended by the IUGS subcommission on the systematics of igneous rocks. *Geotimes* **1973**, *18*, 26–30.
40. Philpotts, A.R.; Ague, J.J. *Principles of Igneous and Metamorphic Petrology*, 3rd ed.; Cambridge University Press: Cambridge, MA, USA, 2022.
41. Ben-Dor, E.; Irons, J.; Epema, G. Soil reflectance. In *Remote Sensing for the Earth Sciences: Manual of Remote Sensing*; Rencz, A., Ed.; John Wiley & Sons: New York, NY, USA, 1999; pp. 111–188.
42. Lin, J.F.; Speziale, S.; Mao, Z.; Marquardt, H. Effects of the electronic spin transitions of iron in lower mantle minerals: Implications for deep mantle geophysics and geochemistry. *Rev. Geophys.* **2013**, *51*, 244–275. [CrossRef]
43. Bigham, J.M.; Fitzpatrick, R.W.; Schulze, D.G. Iron Oxides. In *Soil Mineralogy with Environmental Applications*; Soil Science Society of America Book Series; Dixon, J.B., Schulze, D.G., Eds.; Wiley: Hoboken, NJ, USA, 2002.
44. Syverson, D.D.; Tutolo, B.M.; Borrok, D.M.; Jr, W.E.S. Serpentinization of olivine at 300 °C and 500 bars: An experimental study examining the role of silica on the reaction path and oxidation state of iron. *Chem. Geol.* **2017**, *475*, 122–134. [CrossRef]
45. Gupta, R.P. *Remote Sensing Geology*; Springer: Berlin/Heidelberg, Germany, 2003.
46. Kokaly, R.F.; Clark, R.N.; Swayze, G.A.; Livo, K.E.; Hoefen, T.M.; Pearson, N.C.; Wise, R.A.; Benzel, W.M.; Lowers, H.A.; Driscoll, R.L.; et al. *USGS Spectral Library Version 7*; USGS: Riston, VA, USA, 2017.
47. Saad, A.H. Magnetic properties of ultramafic rocks from Red Mountain, California. *Geophysics* **1969**, *34*, 974–987. [CrossRef]
48. Oufi, O.; Cannat, M.; Horen, H. Magnetic properties of variably serpentinized abyssal peridotites. *J. Geophys. Res.* **2002**, *107*, EPM 3-1–EPM 3-19. [CrossRef]
49. Bach, W.; Paulick, H.; Garrido, C.J.; Ildefonse, B.; Meurer, W.P.; Humphris, S.E. Unraveling the sequence of serpentinization reactions: Petrography, mineral chemistry, and petrophysics of serpentinites from MAR 15 °N (ODP Leg 209, Site 1274). *Geophys. Res. Lett.* **2006**, *33*, L13306. [CrossRef]
50. Hong, G.; Till, J.L.; Greve, A.; Lee, S.M. New Rock Magnetic Analysis of Ultramafic Cores From the Oman Drilling Project and Its Implications for Alteration of Lower Crust and Upper Mantle. *J. Geophys. Res. Solid Earth* **2022**, *127*, e2022JB024379. [CrossRef]

51. McCollom, T.M.; Klein, F.; Moskowitz, B.; Berquo, T.S.; Bach, W.; Templeton, A.S. Hydrogen generation and iron partitioning during experimental serpentinization of an olivine–pyroxene mixture. *Geochim. Cosmochim. Acta* **2020**, *282*, 55–75. [CrossRef]
52. Maar, G.W.t.; McEnroe, S.A.; Church, N.S.; Larsen, R.B. Magnetic Mineralogy and Petrophysical Properties of Ultramafic Rocks: Consequences for Crustal Magnetism. *Geochem. Geophys. Geosyst.* **2019**, *20*, 1794–1817.
53. Cocks, T.; Jenssen, R.; Stewart, A.; Wilson, I.; Shields, T. The HyMap airborne hyperspectral sensor: The system, calibration and performance. In Proceedings of the 1st EARSeL Workshop on Imaging Spectroscopy, Zurich, Switzerland, 6–8 October 1998.
54. Kruse, F.A.; Boardman, J.W.; Lefkoff, A.B.; Young, J.M.; Kierein-Young, K.S.; Cocks, T.D.; Jenssen, R.; Cocks, P.A. HyMap: An Australian hyperspectral sensor solving global problems-results from USA HyMap data acquisitions. In Proceedings of the 10th Australasian Remote Sensing and Photogrammetry Conference, Adelaide, Australia, 25 August 2000; pp. 18–23.
55. Richter, R. *Atmospheric/Topographic Correction for Airborne Imagery*; DLR—German Aerospace Center: Wessling, Germany, 2010.
56. Miziolek, A. Laser-induced breakdown spectroscopy—An emerging chemical sensor technology for real-time field-portable, geochemical, mineralogical, and environmental applications. *Appl. Geochem.* **2006**, *21*, 730–747.
57. Rogass, C.; Koerting, F.M.; Mielke, C.; Brell, M.; Boesche, N.K.; Bade, M.; Hohmann, C. Translational Imaging Spectroscopy for Proximal Sensing. *Sensors* **2017**, *17*, 1857. [CrossRef]
58. Rogass, C.; Segl, K.; Mielke, C.; Fuchs, Y.; Kaufmann, H. Engeomap—A geological mapping tool applied to the enmap mission. *EARSeL eProc.* **2013**, *12*, 94–100.
59. Stark, P.B.; Parker, R.L. Bounded-Variable Least-Squares: An Algorithm and Applications. *Comput. Stat.* **1995**, *10*, 129–141.
60. Clark, R.N.; Swayze, G.A.; Wise, R.A.; Livo, K.E.; Hoefen, T.M.; Kokaly, R.F.; Sutley, S.J. *USGS Digital Spectral Library splib06a*; Data Series 231; US Geological Survey: Reston, VA, USA, 2007.
61. Pearson, K. On lines and planes of closest fit to systems of points in space. *Philos. Mag. Lett.* **1901**, *2*, 559–572. [CrossRef]
62. Mielke, C.; Boesche, N.K.; Rogass, C.; Kaufmann, H.; Gauert, C.; Wit, M.D. Spaceborne Mine Waste Mineralogy Monitoring in South Africa, Applications for Modern Push-Broom Missions: Hyperion/OLI and EnMAP/Sentinel-2. *Remote Sens.* **2014**, *6*, 6790–6816. [CrossRef]
63. Li, Y.; Oldenburg, D.W. Rapid construction of equivalent sources using wavelets. *Geophysics* **2010**, *75*, L51–L59. [CrossRef]
64. Dilixiati, Y.; Baykiev, E.; Ebbing, J. Spectral consistency of satellite and airborne data: Application of an equivalent dipole layer for combining satellite and aeromagnetic data sets. *Geophysics* **2022**, *87*, G71–G81. [CrossRef]
65. Aster, R.C.; Borchers, B.; Thurber, C.H. *Parameter Estimation and Inverse Problems*; Elsevier Academic Press: Amsterdam, The Netherlands, 2005; p. 301.
66. Menke, W. *Geophysical Data Analysis Discrete Inverse Theory*; Academic Press Limited: Amsterdam, The Netherlands, 1989; Volume 45.
67. Lessovaia, S.N.; Goryachkin, S.; Polekhovskii, Y.S. Soil formation and weathering on ultramafic rocks in the mountainous tundra of the Rai-Iz Massif, Polar Urals. *Eurasian Soil Sci.* **2012**, *45*, 33–44. [CrossRef]
68. Funatomi, T.; Ogawa, T.; Tanaka, K.; Kubo, H.; Caron, G.; Mouaddib, E.; Matsushita, Y.; Mukaigawa, Y. Eliminating temporal illumination variations in whisk-broom hyperspectral imaging. *Int. J. Comput. Vis.* **2022**, *130*, 1310–1324. [CrossRef]
69. Uezato, T.; Yokoya, N.; He, W. Illumination Invariant Hyperspectral Image Unmixing Based on a Digital Surface Model. *IEEE Trans. Image Process.* **2020**, *29*, 3652–3664. [CrossRef] [PubMed]
70. Cardoso-Fernandes, J.; Silva, J.; Dias, F.; Lima, A.; Teodoro, A.C.; Barres, O.; Cauzid, J.; Perrotta, M.; Roda-Robles, E.; Ribeiro, M.A. Tools for Remote Exploration: A Lithium (Li) Dedicated Spectral Library of the Fregeneda–Almendra Aplite–Pegmatite Field. *Remote Sens.* **2021**, *6*, 33. [CrossRef]
71. Koerting, F.M.; Koellner, N.; Kuras, A.; Boesche, N.; Rogass, C.; Mielke, C.; Elger, K.; Altenberger, U. A solar optical hyperspectral library of rare-earth-bearing minerals, rare-earth oxide powders, copper-bearing minerals and Apliki mine surface samples. *Earth Syst. Sci. Data* **2021**, *13*, 923–942. [CrossRef]
72. Hong, Y.; Chen, S.; Chen, Y.; Lindemann, M.; Mouazen, A.M.; Liu, Y.; Guo, L.; Yu, L.; Liu, Y.; Cheng, H.; et al. Comparing laboratory and airborne hyperspectral data for the estimation and mapping of topsoil organic carbon: Feature selection coupled with random forest. *Soil Tillage Res.* **2020**, *199*, 104589. [CrossRef]

Related papers:

Machine learning methods for road edge detection on fused airborne hyperspectral and LiDAR data

Machine learning methods for road edge detection on fused airborne hyperspectral and LiDAR data

Rabin Senchuri[1], Agnieszka Kuras[1], Ingunn Burud[1]

[1]   *Faculty of Science and Technology, Norwegian University of Life Sciences, PB 5003, 1430 Aas, Norway*

MACHINE LEARNING METHODS FOR ROAD EDGE DETECTION ON FUSED AIRBORNE HYPERSPECTRAL AND LIDAR DATA

*Rabin Senchuri[1], **Agnieszka Kuras**[1], Ingunn Burud[1]*

[1]Norwegian University of Life Sciences, Faculty of Science and Technology, PB 5003, 1430 Aas, Norway

ABSTRACT

In the last decades, remote sensing sensors, such as hyperspectral systems or LiDAR scanners, have been used for urban mapping. However, an analysis in the urban environment is very complex in applications, e.g., road detection, city management, and urban planning. One of the important urban features is the detection of the road edges. In this study, an approach on multisensory hyperspectral and LiDAR data fusion (HL-Fusion) is introduced for road edge detection using different machine learning algorithms, such as Support Vector Machines, Random Forests, and Convolutional Neural Networks. The first results show that the Random Forest algorithm outperformed in the experiments on the study area at Oslo's surroundings in Norway. This study opens a window for further investigation on machine learning algorithms and a better understanding of HL-Fusion capabilities.

Index Terms— Hyperspectral, road edge detection, LiDAR, machine learning, data fusion, remote sensing

1. INTRODUCTION

Precise, cost- and time-efficient urban mapping has been an essential task for city management, navigation systems development, and more. One of the advanced methods in urban classification is airborne-based image analysis. In particular, road detection and road edge detection have proven to be crucial in each of those applications. Road edge detection has already been investigated in the computer vision research field using airborne-based RGB [1] and grey-scale images for rural areas [2]. For the automatic road detection approach from RGB-based images, diverse methods have been proposed [3, 4]. Although the RGB systems analysis is a conventional computer vision method, exceeding the visible range (0.4 – 0.7 μm) of the electromagnetic spectrum enables the use of multi- or hyperspectral imagery and dense spectral sampling for the latter [5].

In the last decades, active and passive remote sensing has been widely used for urban analysis, such as landcover mapping, urban planning, urbanization changes, biodiversity, and road detection [6]. Active remote sensing such as Light Detection and Ranging laser scanner (LiDAR) provides geometric and textural properties of targets and precise elevation extraction in a time-efficient way for large areas [7]. The elevation information can be used to differentiate between road and curbside. However, the automated classification of exact road edges based on LiDAR data is challenging due to its low spatial resolution and commonly one available wavelength, limiting the classification of complex urban structures based on their spectral signatures.

In comparison, the benefit of using passive remote sensing, such as hyperspectral imaging, is the ability to recognize the material properties due to its unique absorption features, called spectral fingerprints in the visible and near Infrared (VNIR: 0.4 – 1.0 μm) and the short wave Infrared (SWIR: 1.0 – 2.5 μm) [8,9].

The strategy to apply hyperspectral and LiDAR data fusion (HL-Fusion) has already been proposed by Weinmann et al. [10] for a small urban area. Their approach was to apply color and spectral information from VNIR hyperspectral and shape information from the LiDAR dataset for urban object classification, such as road, buildings, sidewalk, and vegetation. However, combining elevation information from LiDAR, such as normalized Digital Surface Model (nDSM), high spectral and spatial information from hyperspectral data (VNIR and SWIR), can help to deliver robust and accurate road edge mapping that includes spectral-spatial-elevation context [10, 11].

Machine learning algorithms are currently used to classify urban objects based on remotely sensed data. However, different algorithms achieve the best results when detecting different features. Therefore, knowledge is required on the features to be appropriately classified. Standard classifiers in urban landcover classification are Support Vector Machine (SVM) and Random Forest (RF). SVM handles high dimensional hyperspectral data and deals with small training datasets; therefore, it is widely applied in the urban analysis based on hyperspectral data [12], LiDAR [13], and HL-Fusion [14]. RF provides high accuracy of the hyperspectral data classification, high processing speed, retaining relevant spectral information without overfitting

Figure 1 Training dataset in Sandvika, Oslo surroundings in Norway (672x2560 pixels).

Figure 2 Test dataset in Sandvika, Oslo surroundings in Norway (480x1600 pixels).

[15]. Moreover, RF has also been used to LiDAR data [16] and different HL-Fusion methods reviewed by Debes et al. [17].

A more advanced classification method in deep machine learning is Convolutional Neural Network (CNN), applied for urban landcover mapping on hyperspectral data [17] and HL-Fusion [18], among others. CNN automatically learns abstract features and does not require prior knowledge about the class distribution in the hyperspectral scene [19].

In this study, we fuse hyperspectral and LiDAR data to extract road edges in Oslo and its surroundings. We apply machine learning algorithms, such as SVM, RF, and CNN, and compare them to each other in the road edge extraction.

The structure of this work is as follows. Section 2 presents the study area. In section 3, the methodology on the road edge detection based on HL-Fusion is explained. Section 4 provides the results and their discussion. Section 5 concludes the study and shows future perspectives and suggestions for further research.

2. STUDY AREA

The airborne hyperspectral and LiDAR data have been acquired simultaneously by the Terratec AS Company in August 2019 and April 2020. The dataset contained cloud-free airborne-based hyperspectral and LiDAR data over Bærum municipality near Oslo, Norway. The hyperspectral data were acquired using two HySpex sensors: VNIR-1800 (0.4 – 1.0 μm) and SWIR-384 (1.0 – 2.5 μm) with 0.3 and 0.7 m spatial resolution, respectively. The LiDAR data were acquired using ALS70 and Riegl VQ-1560i, with five emitted pulses per m^2 and intensity at 1.064 μm. The hyperspectral signatures were preserved using Nearest-neighbor interpolation. The study area shows a complex urban environment with various urban objects, such as road, vegetation, building, waterbody, train track (Figure 1, 2). In our experiments, we divided the study area into smaller parts due to the large files and high-dimensionality of the data.

3. METHODOLOGY

In this study, the following approach of HL-Fusion for road edge delineation was carried out. Two different HL-Fusions have been applied: 1) LiDAR and radiance data 2) LiDAR and reflectance data. Since the hyperspectral dataset contains data from two different sensors, VNIR and SWIR, the spatial resolution was unified to 0.3 m pixel size. The geocoded radiance data were converted to reflectance, adjusting illumination levels using ATCOR-4 (Atmospheric and Topographic Correction for airborne imagery) software [21]. For the radiance and reflectance data, individually, the Normalized Difference Vegetation Index (NDVI) was applied to mask the study area's vegetation. Further, the Principal Component Analysis (PCA) was applied to hyperspectral VNIR and SWIR data to reduce the high dimensionality of the data and extract spectrally homogeneous regions. The first three principal components (PCs) have been used as input for classification algorithms, covering 99.5 % variance.

The LiDAR-derived features include the intensity values of 1.064 μm from the first return and normalized Digital Surface Model (nDSM). For the HL-Fusion, hyperspectral and LiDAR data were geometrically coregistered [22]. The nDSM was used to mask out the elevated objects, such as trees and buildings. For the road edge delineation purpose, the study area's main urban objects except for roads have been masked in the image, such as train track, waterbody, vegetation (trees, grass), and buildings.

Since we implemented a supervised classification to identify roads, train tracks, vegetation, and waterbody, the labels have been generated manually pixel-wise directly from the training dataset retaining the pixel count similar for each class. For this study, the following machine learning algorithms, SVM, RF, and CNN, were applied to the radiance and the reflectance HL-Fusion data. We split the dataset in training – 70 % and testing – 30 %. For the CNN model, the training data are split into patches of 9x9 pixels each. Three convolution layers with 30 filters and 3x3 filter kernel sizes are applied to each patch. To minimize
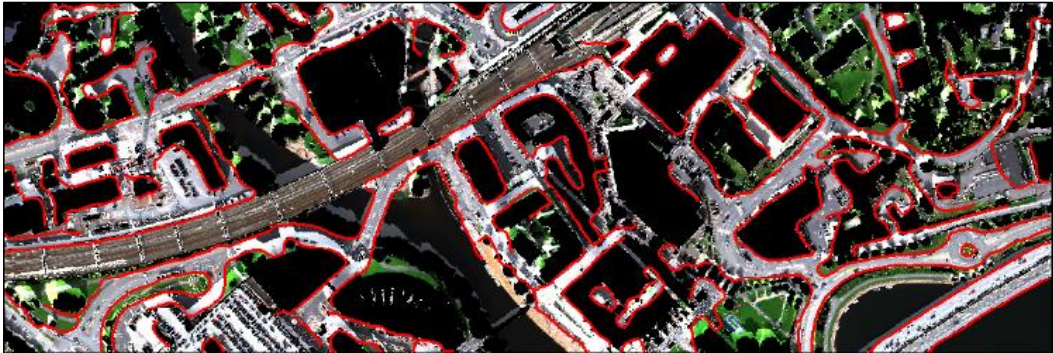
Figure 3 The first results for RF on road edge detection based on HL-Fusion.

Table 1 Comparison of classification results for reflectance (Ref) and radiance (Rad) HL-Fusions. F1 score corresponds to the road class.

	SVM		RF		CNN	
	Rad	Ref	Rad	Ref	Rad	Ref
OA [%]	83	92	**92**	**93**	80	77
F1 [%]	82	77	**88**	**79**	60	57
Time [s]	27.71	5.52	47.64	5.82	450.06	321.5

overfitting and enhance generalization, we add a dropout layer. We chose ReLU as the activation function, "categorical_crossentropy" as the loss function, and Stochastic Gradient Decent (SGD) as the optimization algorithm. Since the CNN training requires a large amount of data, we applied data augmentation by rotating, zooming, and flipping existing images. The final step was to delineate the road edges and produce a map with road lines applying the Canny Edge detector on the image, maintaining road and low vegetation classes.

4. RESULTS AND DISCUSSION

The classification results were evaluated, calculating the overall accuracy (OA), F1-score, and computation time (CT) (Table 1). The initial results show that the RF classifier outperformed for reflectance and radiance HL-Fusions compared to SVM and CNN algorithms, achieving 93 % accuracy (see Table 1). The RF can handle multiclass issues, is less sensitive to noise was already used for HL-Fusion [10]. CNN's poor performance can be related to limited training samples and high-dimensionality of the data, leading to overfitting and longer computing time.

Figure 3 presents one of the first road edge delineation results for reflectance HL-Fusion-based RF classification, achieving 93 % accuracy.

However, there are misclassification spots in the road delineation results (Figure 3). One of the road detection challenges is the lack of identification of smaller roads, which can be caused by the too low resolution of the hyperspectral images and incorrect shadow classification. Another aspect is that the edges of some of the roads are not straight, broken, or misclassified. The main reason is that the road edges are often covered by buildings or trees in airborne-based images due to the inability of hyperspectral sensors to penetrate the surface.

Better F1 score road class results for each classifier for radiance data show that the atmospheric correction must not be required to achieve high urban object classification results. The reason is that any further processing steps can lead to artifacts misclassifying targets of interest. However, the multitemporal analysis results on radiance data may not deliver such a final overall accuracy of the classification since the reflectance data are more reliable and repeatable than radiance data.

5. CONCLUSION AND OUTLOOK

This study evaluated different machine learning classification algorithms for road edge detection based on HL-Fusion. Our main objectives were to provide insights into the capabilities of using multisensor data fusion in urban mapping, considering one of the essential features in many applications, such as road edge detection. The machine learning-based classifier – RF provided the best accuracy results for radiance and reflectance HL-Fusion.

Although deep learning analysis attracts more and more attention, ensemble learning has proved to be the best choice in this classification problem. However, we believe that deep learning is a promising basis for exploring HL-Fusion

further and combining more algorithms in one analysis to improve data reduction, image segmentation, classification, and post-processing of the data.

The next step in this study will be to analyze multitemporal radiance and reflectance data from August 2019, April 2020, and September 2020. The proposed method will be tested in an area with known new roads not covered by the vector data, especially Oslo and its surroundings.

Further research will also explore how the HL-Fusion can be utilized to determine the microtopography along the roads and detect curbsides and other urban microstructures. We also want to explore deep learning algorithms' potential, collecting more training data and other deep learning algorithms to extract time-efficiently deep features in the spectral-spatial context and object-based classification.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] X. Yang, X. Li, Y. Ye, H. Zhang, X. Huang, and B. Zhang, "Road detection via Deep Residual Dense U-Net", *IJCNN 2019, International Joint Conference on Neural Networks*, Budapest, Hungary, 14-19 July 2019.

[2] I. Laptev, H. Mayer, T. Lindeberg, W. Eckstein, C. Steger, and A. Baumgartner, "Automatic extraction of roads from aerial images based on scale-space and snakes", *Machine Vision and Applications*, vol. 12, pp. 23-32, 2000.

[3] Z. Miao, W. Shi, H. Zhang, and X. Wang, "Road centerline extraction from high-resolution imagery based on shape features and multivariate adaptive regression splines", *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 3, pp. 583–587, 2013.

[4] C. Unsalan and B. Sirmacek, "Road network detection using probabilistic and graph theoretical methods", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4441–4453, 2012.

[5] M. Taghizadeh, A.A. Gowen, and C.P. O'Donnell, "Comparison of hyperspectral imaging with conventional RGB imaging for quality evaluation of Agaricus bisporus mushrooms", *Biosystems Engineering*, vol. 108, pp. 191-194, 2011.

[6] J.A. Hepinstall, M. Alberti, and M. Marzluff, "Predicting land cover change and avian community responses in rapidly urbanizing environments", *Landscape Ecology*, vol. 23, pp. 1257-1276, 2008.

[7] A. Serna and B. Marcotegui, "Urban accessibility diagnosis from mobile laser scanning data", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 84, pp. 23-32, 2013.

[8] Y. Gu, Y. Zhang, D. You, and Y. Zhang, "Representative multiple kernel learning for classification in hyperspectral imagery", *IEEE Transactions on Geoscience and Remote Sensing, 2012*.

[9] E. Ben-Dor, "Imaging spectroscopy for urban applications", In *Imaging Spectrometry,* eds. FD van der Meer and SM de Jong, The Netherlands: Kluwer Academic Publishers, pp. 243-281, 2001.

[10] M. Weinmann and M. Weinmann, "Geospatial Computer Vision Based on Multi-Modal Data – How Valuable Is Shape Information for the Extraction of Semantic Information?", *Remote Sensing*, vol. 10, no. 1, 2018.

[11] H. Li, P. Ghamisi, U. Soergel, and X.X. Zhu, "Hyperspectral and LiDAR fusion using deep three-stream convolutional neural networks", *Remote Sensing*, vol. 10, 2018.

[12] Y. Chen, N.M., Nasrabadi, and T.D. Tran, "Hyperspectral Image Classification via Kernel Sparse Representation", *IEEE Transactions on Geoscience and Remote Sensing*, 2011.

[13] S.K. Lodha, E.J. Kreps, D.P. Helmbold, and D. Fitzpatrick, "Aerial LiDAR data classification using support vector machines (SVM)", *The Third International Symposium on 3D Data Processing, Visualization, and Transmission,* 2006.

[14] M. Brell, K. Segl, L. Guanter, and B. Bookhagen, "3D hyperspectral point cloud generation: fusing airborne laser scanning and hyperspectral imaging sensors for improved object-based information extraction", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 149, pp. 200-214, 2019.

[15] S. Amini, S. Homayouni, and A. Safari, "Semi-supervised classification of hyperspectral image using random forest algorithm", *International Geoscience and Remote Sensing Symposium*, 2014.

[16] N. Chehata, L. Guo, and C. Mallet, "Airborne LiDAR feature selection for urban classification using random forests", *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 38, 2009.

[17] C. Debes, A. Merentitis, R. Heremans, J. Hahn, N. Frangiadakis, T. van Kasteren, W. Liao, R. Bellens, A. Pizurica, S. Gautama, W. Philips, S. Prasad, Q. Du, and F. Pacifici, "Hyperspectral and LiDAR data fusion: outcome of the 2013 GRSS data fusion contest", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, 2014.

[18] A. Santara, K. Mani, P. Hatwar, A. Singh, A. Garg, K. Padia, and P. Mitra, "BASS Net: band-adaptive spectral-spatial feature learning neural network for hyperspectral image classification", *IEEE Transactions on Geoscience and Remote Sensing*, 2016.

[19] Y. Chen, C. Li, P. Ghamisi, C. Shi, and Y. Gu, "Deep fusion of hyperspectral and LiDAR data for thematic classification", *International Geoscience and Remote Sensing Symposium*, 2016.

[20] Z. Zhou and J. Gong, "Automated residential building detection from airborne LiDAR data with deep neural networks", *Advanced Engineering Informatics*, 2018.

[21] R. Richter and D. Schläpfer, "Atmospheric/Topographic Correction for Airborne Imagery", 2015.

[22] M. Brell, C. Rogass, K. Segl, B. Bookhagen, and L. Guanter, "Improving Sensor Fusion: A Parametric Method for the Geometric Coalignment of Airborne Hyperspectral and Lidar Data", *IEEE Transactions on Geoscience and Remote Sensing*, vol.

54, no. 6 2016.

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no

116009 / ANDVORDGRAFISK.NO