Norwegian University
of Life Sciences

Predictive Machine Learning on SEM and Hyperspectral Images of Uranium Ore Concentrates (UOCs) for Nuclear Forensics

Isak Biringvad Lande

Environmental Physics and Renewable Energy

Acknowledgements

Abstract

Nuclear and radioactive materials are harmful to individuals, especially when it's not professionally managed. Nuclear and radioactive materials going astray pose a great threat to the general public. The field of nuclear forensics focuses on detecting potentially dangerous materials and determining their origin. This work is a part of the anti-terrorism effort. The investigations can take months before it produces any results, and it's desirable obtain critical information early on. This thesis aims at providing a new method that accurately classifies unknown nuclear samples with high accuracy.

A machine learning (ML) model was developed to provide nuclear forensics with a new tool. The model takes images of uranium samples and learns their discriminative characteristics to determine the true sample class. Two different types of images were used, namely scanning electron microscopy (SEM) and hyperspectral images. These provide different type of information for discrimination, where the former gives morphological information, and the latter gives information about the chemical composition.

Model evaluation and testing revealed that the best performing classifier in terms of accuracy was LDA, and the most informative features were LBP. Both SEM images and hyperspectral images provided satisfactory results when used in their model and could be applied in the field immediately. Further work should focus on making the model more user-friendly. The model performance could be further improved but it is at a satisfactory level in its current state. This thesis demonstrates the usefulness of ML models in the field of nuclear forensics. It is a low-cost, efficient and accurate tool.

Sammendrag

Kjernefysiske og radioaktive materialer er skadelige for mennesker, spesielt når de ikke håndteres av kvalifiserte personer. Kjernefysiske og radioaktive materialer som kommer på villspor, utgjør en stor trussel for allmennheten. Fagfeltet kjernefysisk etterretning fokuserer på å oppdage potensielt farlige materialer og determinere prøvens geografiske opprinnelse. Dette arbeidet er en del av innsatsen mot terrorisme. Undersøkelsene kan ta flere måneder før de gir resultater, og det er ønskelig å redusere varigheten av undersøkelsen. Denne oppgaven tar sikte på å konstruere en ny metode som nøyaktig klassifiserer ukjente kjerneprøver med høy nøyaktighet.

En maskinlæringsmodell ble utviklet for å gi kjernefysisk etterretning et nytt verktøy. Modellen tar bilder av uran-prøver og tilegner seg kunnskap om prøvens karakteristikk. Karakteristikken blir brukt til å skille prøvene fra hverandre og determinere prøvens sanne klasse-tilhørighet. To forskjellige typer bilder ble brukt, nemlig skanning elektronmikroskopi og hyperspektrale bilder. Disse gir forskjellig type informasjon for diskriminering. Førstnevnte gir morfologisk informasjon, og sistnevnte gir informasjon om den kjemiske sammensetningen.

Modellevaluering og testing avdekket at klassifikator-algoritme som hadde best resultat når det gjelder nøyaktighet, var LDA, og de mest informative funksjonene var LBP. Både SEM-bilder og hyperspektrale bilder ga tilfredsstillende resultater når de ble brukt i modellen og kunne brukes i felt. Videre arbeid bør fokusere på å gjøre modellen mer brukervennlig. Modellytelsen kan forbedres ytterligere, men er på et tilfredsstillende nivå per dags dato til å benyttes i felt. Denne oppgaven viser nytten av ML-modeller innen kjernefysisk etterretning. Det er et billig, effektivt og nøyaktig verktøy.

Abbreviations

AI	Artificial Intelligence
AMT	Angle Measure Technique
CC	Colour Category
CI	Confidence Interval
EA	Evolutionary Algorithm
EBR-I	Experimental Breeder Reactor-I
FOS	First Order Statistics
GA	Genetic Algorithm
GLCM	Gray Level Co-Occurence Matrix
GLRLM	Gray Level Run Length Matrix
GLSZM	Gray Level Size Zone Matrix
GNB	Gaussian Naive Bayes
GUI	Graphical User Interface
HEU	Highly Enriched Uranium
HSI	Hyperspectral Imaging
IAEA	International Atomic Energy Agency
IDE	Integrated Development Environment
ITWG	International Technical Working Group
JRC	Joint Research Centre
KNN	K-Nearest Neighbours
LBP	Local Binary Patterns
LDA	Linear Discriminant Analysis
LR	Logistic Regression
ML	Machine Learning
MSC	Multiplicative Scatter Correction
NB	Naive Bayes
n-CV	nested Cross-Validations
NIR	Near InfraRed

NMBU	Norwegian University of Life Sciences
OVO	One-vs-One
PCA	Principal Component Analysis
PLS-DA	Partial Least Squares Discriminant Analysis
PPI	Pixels Per Inch
RF	Random Forest
ROC AUC	Receiver Operating Characteristic Area Under the Curve
SBS	Sequential Backward Selection
SD	Standard Deviation
SEM	Scanning Electron Microscope
SVC	Support Vector Classifier
SVM	Support Vector Machine
UOC	Uranium Ore Concentrates
UV	UltraViolet
Vis	Visible

Table of Contents

1. Introduction

History was made on December the 20th, 1951, when the Experimental Breeder Reactor-I (EBR-I) in the USA become the world's first electricity-producing nuclear power plant (Idaho National Laboratory). Fuelled on uranium, it powered four lightbulbs at that time. Today, uranium powers numerous nuclear plants all around the world (INTERNATIONAL ATOMIC ENERGY AGENCY, 2020) and accounts for about 10 % of the world's electricity produced in 2018 (IAEA, 2020). Despite being a promising solution for combatting global warming (IAEA, 2020), the use of nuclear energy brings its own challenges that must be handled. Uranium, along with other nuclear and radioactive materials, poses risks to living organisms due to its radioactivity and potential use in nuclear weapons (Lin, 2015). There is a growing concern of malevolent actions such as terrorism because of the large quantities of traded nuclear materials, such as the precursors of enriched uranium - uranium ore concentrates (UOC) (Kristo & Tumey, 2012) (Fongaro, Ho, Kvaal, Mayer, & Rondinella, 2016). This material was of interest in this thesis.

Many countries began conducting laboratory analysis on interdicted nuclear and radioactive materials to prevent nuclear terrorism (Kristo & Tumey, 2012), and thereby the field of nuclear forensics arose (Fongaro, Ho, Kvaal, Mayer, & Rondinella, 2016). Nuclear forensics support law enforcement by seeking to answer a variety of questions, one of which this thesis will focus on - the origin of the analysed nuclear material. Different methods of investigation are carried out, and various material characteristics are assessed, such as physical appearance and chemical compositions (Lin, 2015), to determine the origin of the materials.

Nuclear and other radioactive materials can undoubtedly fall into the wrong hands. There has been reported almost 2500 confirmed incidents where these materials have been out of regulatory control solely in a 20 years period from 1993 (Lin, 2015). Even more frightening, unauthorised possession and associated crimes accounted for over 400 of the cases (INTERNATIONAL ATOMIC ENERGY AGENCY, 2015). Furthermore, the number of reported cases increased during the period such that close to half of the incidents happened between 2007 and 2012 (Lin, 2015). The worst-case of illicit trafficking would involve weapons-grade nuclear material, such as plutonium and highly enriched uranium (HEU) at appropriate isotope concentrations (Lin, 2015).

There are indications that more than 17 *kg* and 400 *g* of HEU and plutonium, respectively, were interdicted between 1998 and 2013 (Joint Working Group of the APS and the AAAS). More relevant for this thesis, is a real incident involving UOCs analysed in nuclear forensics back in 1978 when five barrels of this material were stolen (Budinger, Drenski, Varnes, & Mooney, 1980).

Nuclear forensics investigation is a time-consuming process (Joint Working Group of the APS and the AAAS); months of research might be necessary to achieve sufficient levels of confidence in the evaluation of the measurements (Kristo & Tumey, 2012). With more examinations and following aggregated results, the possible origins of interdicted samples are narrowed down to match known materials. Nevertheless, initial assessments give valuable insight that helps to guide the criminal investigation. As time involving an incident is limited, an early indication of the origin of interdicted samples would be of high importance (Kristo & Tumey, 2012).

In hopes of contributing to the established process of nuclear forensics, this thesis presents a method of apparently high accuracy for rapid classification of unknown UOCs. The main objective in this thesis was to utilise a subfield of artificial intelligence (AI), i.e. machine learning (Raschka & Mirjalili, Python Machine Learning, 2019), on scanning electron microscope (SEM) images and hyperspectral images for classifying the origin of UOCs – the precursors of enriched uranium.

SEM images can capture the morphology of UOCs and provide discriminative characteristics of their processing history (Lin, 2015). Hyperspectral images give information of their chemical composition (Plaue, Klunder, Czerwinski, & Hutcheon, 2012) (Klunder, et al., 2012) which can indicate the geological and depositional setting of the originating ore (Kristo & Tumey, 2012). Machine learning algorithms can take this information, along with their known origin, as input for supervised learning to develop models that seek to classify new unseen samples.

The SEM and hyperspectral images were acquired and given access to by the European Commission's science and knowledge service Joint Research Centre (JRC) in Karlsruhe, Germany. JRC Karlsruhe conducts Nuclear Safety and Security research and supports the EU commission with evidence in a variety of areas (EU Commission, 2020).

By the knowledge of the author, no other study has used more UOCs of different origins for image classification than this thesis. Also, this may be the first time that hyperspectral images have been used for classifying UOCs.

The thesis begins by presenting relevant theory in chapter 2, before describing the materials and methods used in chapter 3. Then, a selection of results is presented along with observations in chapter 4. Chapter 5 discusses and assesses the methods and results, with conclusions given in chapter 6.

2. Theory

2.1. Nuclear forensics

2.1.1. Concept

Nuclear forensics seeks to find answers to the origin and intended use of seized nuclear and radioactive materials out of regulatory control. This field of forensics is interdisciplinary and includes sciences such as radiochemistry, nuclear and reactor physics, as well as materials science (Mayer, Wallenius, & Ray, Nuclear forensics—a methodology providing clues on the origin of illicitly trafficked nuclear materials, 2005).

2.1.2. Uranium ore concentrates (UOC)

The nuclear industry is highly dependent on uranium, as no other element that occurs naturally can be processed into reactor fuel (IAEA, 2019). As mentioned in the introduction, UOCs are the precursors for enriched uranium - which are the feed for uranium fuel fabrication. The chemical composition of UOCs is U_3O_8 and is the resulting product of the first step in the nuclear fuel cycle; mining and milling raw uranium ore. The fuel cycle represents the entire process in which raw uranium is processed into fuel, used in reactors, and in the end final disposal of spent fuel and safe storage (IAEA, 2019).

Spent fuel is not necessarily disposed of. It can instead be reprocessed for further use. A fuel cycle where fuel is recycled is called a *closed* fuel cycle, whereas an *open* fuel cycle denotes the process where spent fuel is eventually disposed, as shown in Figure 1 (IAEA, 2001).
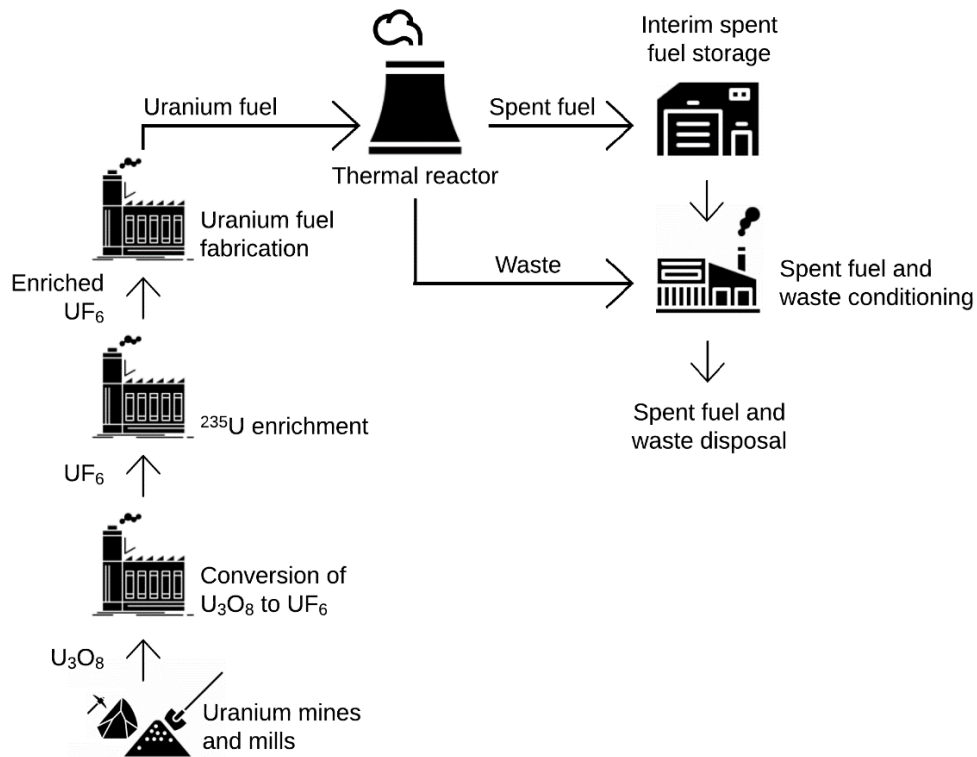
Figure 1: Diagram of an open fuel cycle. Illustration inspired by (IAEA, 2001).

2.1.3. Practices

There is a consensus among the international nuclear forensics' community of which techniques and their order or application that should be used to gain insights into nuclear and radioactive materials (Joint Working Group of the APS and the AAAS). This procedure, recommended by the nuclear forensics International Technical Working Group (ITWG), is presented in Table 1 (IAEA, 2015). Here, possible measures such as physical characterization and chemical composition of the material are included. The sequence of techniques has been established to maximize early valuable information gain (IAEA, 2015) (Joint Working Group of the APS and the AAAS). As can be seen in the table, different measurements are recommended within certain timeframes after the materials arrived at the laboratory, and the total characterization should not take more than two months to complete (IAEA, 2015). High resolution gamma ray spectrometry should be the first analysis performed in the laboratory to ensure quality of on-scene measurements (IAEA, 2015).

Table 1: Techniques and methods in nuclear forensics investigation along with aimed time of completion, from (IAEA, 2015).

Technique/method	Conducted within		
	24 hours	**1 week**	**2 months**
Radiological	Dose rate (α, β, γ, n) Surface contamination Radiography		
Physical characterization	Visual inspection Photography Weight determination Dimensional determination Optical microscopy Density	Microstructure, morphology and other physical characteristics SEM X ray diffraction	Nanostructure, morphology and other physical characteristics TEM
Isotopic analysis	HRGRS	TIMS ICP-MS	SIMS Radioactive counting techniques
Radiochronometry	HRGRS (for Pu)	TIMS ICP-MS	HRGRS (for U) Alpha spectrometry
Elemental/chemical composition	X ray flourescence	ICP-MS Chemical assay FTIR spectrometry SEM/X ray spectrometry IDMS	GC-MS
Traditional forensic science disciplines	Collection of evidence associated with traditional forensic disciplines		Analysis and interpretation of evidence associated with traditional forensic disciplines
Abbreviations:	FTIR	Fourier transform infrared	
	GC–MS	gas chromatography–mass spectrometry	
	HRGRS	high resolution gamma ray spectrometry	
	ICP-MS	inductively coupled plasma mass spectrometry	
	IDMS	isotope dilution mass spectrometry	
	SEM	scanning electron microscopy	
	SIMS	secondary ion mass spectrometry	
	TEM	transmission electron microscopy	
	TIMS	thermal ionization mass spectrometry	

It is crucial that nuclear forensics investigations are initiated immediately following a nuclear or radioactive incident, as this ensures the best quality of sampling for later analyses (IAEA, 2014). The International Atomic Energy Agency (IAEA) has adopted an action plan which nuclear forensics investigations should follow (Joint Working Group of the APS and the AAAS). Figure 2 illustrates the basic scheme of this plan. The plan facilitates the analysis of nuclear and non-nuclear evidence in combination for interpretation (Joint Working Group of the APS and the AAAS).

The order in which the analyses should be conducted is important as some techniques may be destructive (IAEA, 2015). Therefore, the amount of material and gap in information available, as well as possible discriminative signatures such as physical characteristics and chemical composition, must be taken into consideration for planning the forensic investigation (IAEA, 2015).
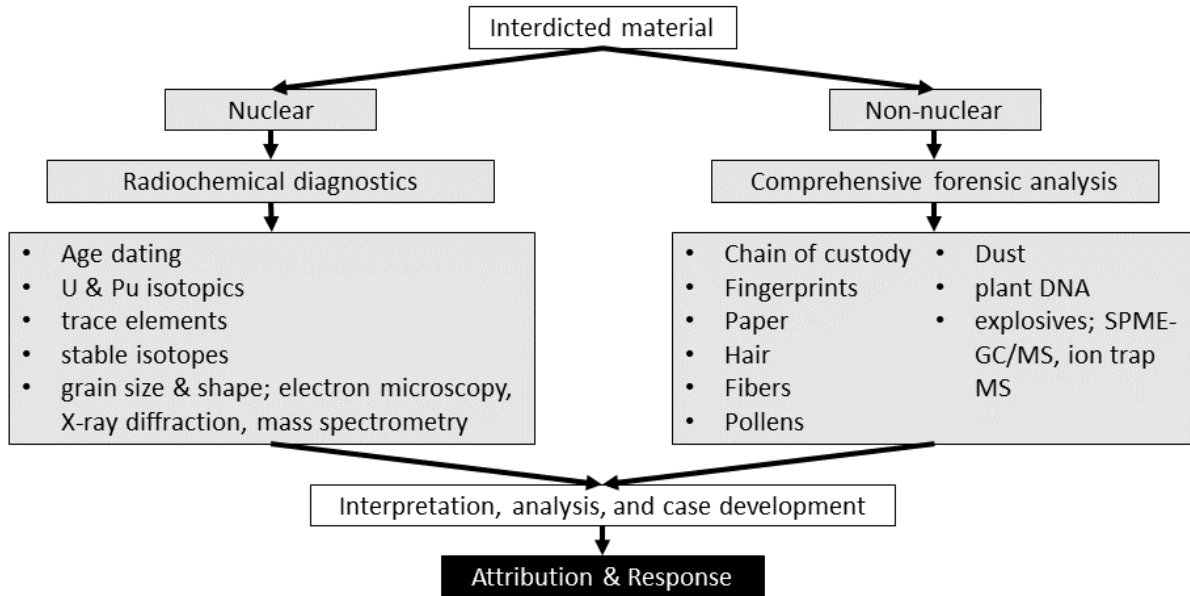
Figure 2: The action plan for nuclear forensic investigation, developed by the ITWG, from (Joint Working Group of the APS and the AAAS).

2.2. Image acquisition techniques

Digital images are created to represent the real world (Burger & Burge, 2016). Assuming rectangular images, they can be treated as matrices where each cell (i.e. pixel) contains information of the colour, and the rows and columns denote the perceived spatial dimensions of the imaged object(s). Therefore, the image resolution is represented by the number of measurements per elements in the real world (Burger & Burge, 2016). Pixels per inch (PPI) is an example of quantifying the resolution.

The pixel values are typically binary with length k, implying that a cell can represent 2^k different numbers. For example, an 8-bit grayscale image has $k = 8$, hence there are $2^8 = 256$ possible pixel values. The values are in the range 0, 1, 2, …, 255, each representing a grayscale intensity where the maximum brightness (i.e. white) is 255 and minimum (i.e. black) is 0. Figure 3 illustrates how a grayscale image is composed of many intensity values. For colour images, additional matrices are created; one matrix of intensity values for each colour. For example, an RGB-image would consist of three matrices, representing the colours red, green, and blue (Burger & Burge, 2016).

Figure 3: Illustration of how a grayscale image is composed of pixels; cells containing intensity values in a matrix. Image object (left) obtained using paid subscription from https://lucid.app/.

2.2.1. Scanning Electron Microscope (SEM)

Scanning electron microscopy is one of the established techniques in nuclear forensics and used for physical characterisation, as seen in Table 1. SEM can magnify an object up to 2,000,000 times (nature research CUSTOM MEDIA and Hitachi High-Technologies, u.d.) revealing intricated details of the structure of the objects. SEM provides the possibility to study the morphology and microstructure of the substance surface, and is therefore used to study UOCs (Mayer, Wallenius,

& Fanghänel, Nuclear forensic science-From cradle to maturity, 2007) (Varga, Wallenius, Mayer, & Meppen, 2011) (Keegan & al., 2014) amongst other things.

The working principle of SEM is to focus and direct the beam of high energy electrons into the specimen and record returning electrons to reconstruct an image. The essential SEM components are presented in Figure 4. The thermionic cathode releases electrons when heated up by an electric current. The electrons are accelerated towards the anode due to a strong electric field between cathode and anode. The electrons, called primary electrons, will form a broad diverging beam from the anode to the specimen if left unchanged. The electromagnetic lens focuses the beam, and electromagnetic deflectors direct the focal point. The incident beam hits the specimen surface and penetrates up to a depth of 1 μm (Khursheed, 2011). The primary electrons collide with the specimen's atoms at different depths and scatter, where some will escape the surface. Primary electrons that collide with specimen atoms at the top surface undergo an inelastic interaction and result in secondary electrons being emitted. The secondary electrons provide information about the surface structure. Primary electrons colliding deeper within the specimen undergo elastic interaction and result in backscatter electrons. The secondary electrons are recorded by a detector and used to build a reconstructed image. The electromagnetic deflectors position the focal point to the top left focal point, and the number of secondary electrons is recorded. This gives the top-left pixel of the resulting image. The scanning process proceeds by scanning all pixels left to right in the first line before it shifts down to the second line repeating the procedure. A focal point with many recorded secondary electrons results in a bright pixel. Fewer recorded secondary electrons result in a dim grey pixel, while no recorded secondary electrons result in a black pixel.
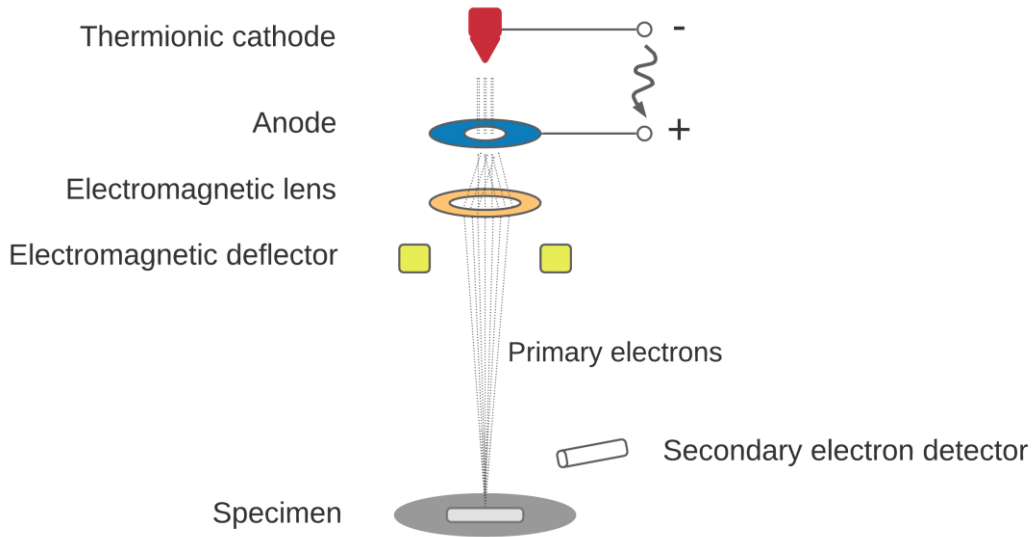
Figure 4: Schematic illustration of Scanning Electron Microscope. Illustration by author.

2.2.2. Hyperspectral imaging (HSI)

The concept of hyperspectral imaging (HSI) is to acquire an image at hundreds of continuous narrow bands in a region of the electromagnetic spectrum (Manolakis, Lockwood, & Cooley, 2016). Many different regions in the electromagnetic spectrum can be used, such as ultraviolet (UV), visible (Vis), and near infrared (NIR) (Amigo, Babamoradi, & Elcoroaristizabal, 2015) (Edelman, Gaston, van Leeuwen, Cullen, & Aalders, 2012). The imaging captures the interaction of incident light on the sample object (Edelman, Gaston, van Leeuwen, Cullen, & Aalders, 2012). Various interactions are possible, and these are illustrated in Figure 5. One of them is the absorption of light in the material and eventually re-emission of the energy as radiation. Absorption is wavelength depended due to the intrinsic properties of the chemical composition of the material (Edelman, Gaston, van Leeuwen, Cullen, & Aalders, 2012).

The hyperspectral image consists of many image matrices, together as a stack, and these compose what is commonly named as a hypercube. Figure 6 illustrates this. As for a grayscale image, the spatial dimensions are the rows and columns in each image matrix, but the third dimension represents the wavelength λ or wavelength band at which the image matrix was captured. The wavelength or band denotes the colour for which the corresponding image matrix contains

intensity values. For each pixel, a spectrum is acquired. The hyperspectral image can thus be represented as a three-dimensional dataset with the number of values equalling the number of rows times the number of columns times the number of bands. There are typically three ways of sampling information for building a hypercube (Manolakis, Lockwood, & Cooley, 2016); pixel-line scanning (pushbroom scanning), pixel by pixel scanning (whiskbroom scanning), and scanning by staring whilst changing wavelength filters.

HSI was primarily used in remote sensing from the start. Since then, the technique has been applied in e.g. pharmaceuticals, medical diagnostics (Edelman, Gaston, van Leeuwen, Cullen, & Aalders, 2012), food sciences, and other fields of research as well as production (Amigo, Babamoradi, & Elcoroaristizabal, 2015). Even more interestingly, HSI has been applied in forensic sciences as it is a non-destructive and non-contact technique (Edelman, Gaston, van Leeuwen, Cullen, & Aalders, 2012). Furthermore, the speed of acquisition, interpretability of both spatial and spectral information, as well as being portable, makes it very usable at the scene of investigation (Edelman, Gaston, van Leeuwen, Cullen, & Aalders, 2012).
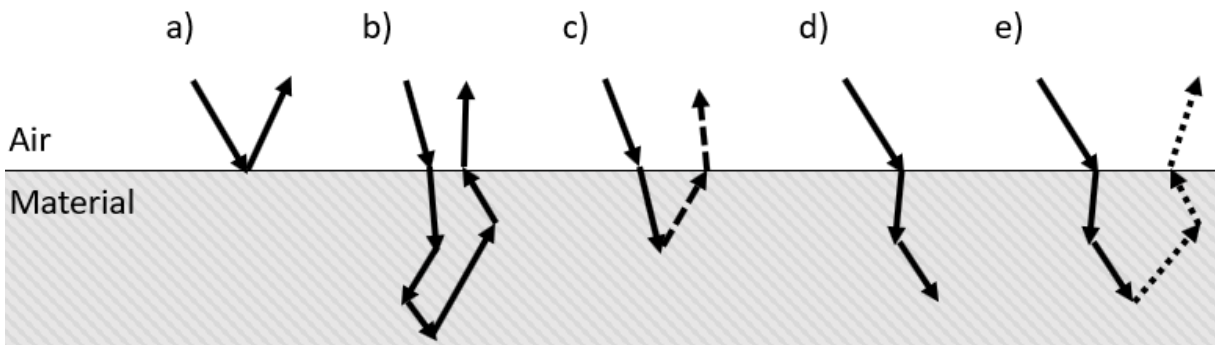
Figure 5: Illustration of the different interactions light can have with a material; a) specular reflection, b) diffuse reflection by elastic scattering, c) emitted Raman shifted light by inelastic scattering, d) absorption, and e) photoluminescence emission by absorption. Inspired by (Edelman, Gaston, van Leeuwen, Cullen, & Aalders, 2012).
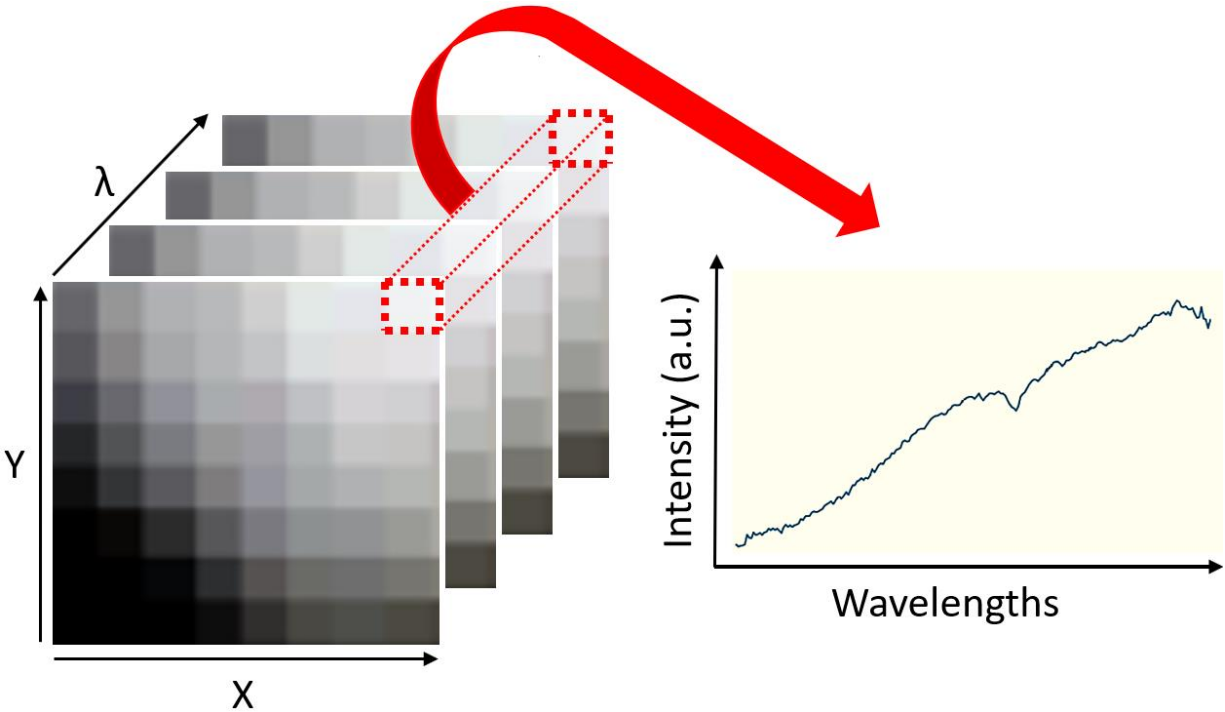
Figure 6: Illustration of a hyperspectral image as a hypercube. Horizontal and vertical axes represent the spatial dimensions, and the depth denote different wavelengths λ. For each pixel, a spectrum has been acquired.

2.3. Machine learning

Artificial intelligence (AI) and one of its sub-fields i.e. machine learning (ML), flourish as buzzwords. But the hype is not necessarily baseless if used intelligently. This thesis has exploited MLs capability of self-learning on labelled (i.e. known) data for classification. This approach lies within *supervised learning*, one of the three types of machine learning. The two other types are unsupervised and reinforcement learning. Principal component analysis (PCA) is an example of an unsupervised learning technique that is popular (Raschka & Mirjalili, Python Machine Learning, 2019).

The idea of supervised learning is to feed an ML algorithm with data and its known labels. These labels would be the desired outcome if those same data were supposed to be predicted. Hence, the generated ML model is supervised in its learning. This trained model will then be applied to new unseen data to predict their labels, as seen in Figure 7. ML models can be used in binary classification with two distinct classes or in multiclass problems with more than two classes. In

12

this thesis, the labels were the origin of the UOCs and will be hereafter referred to as classes. The solved task was a multiclass classification problem, as there were more than two distinct OUC classes. If the desired model output was continuous rather than discrete, it would have been a regression problem (Raschka & Mirjalili, Python Machine Learning, 2019).
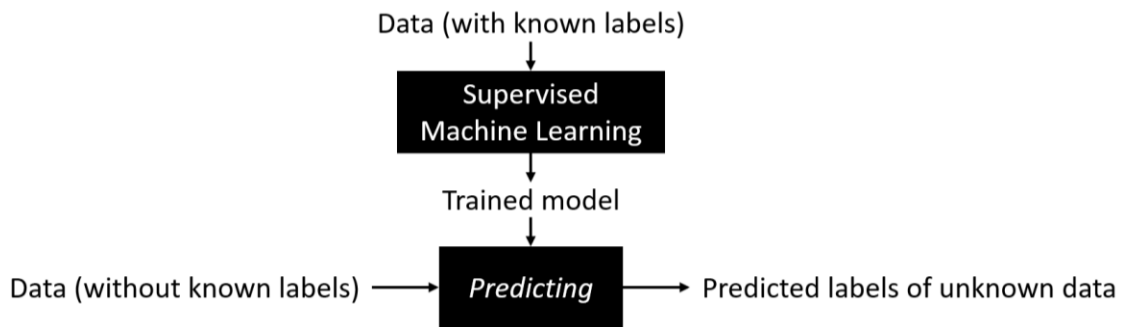
Figure 7: Illustration of supervised training of a ML model for predicting unknown labels of data. Inspired by (Raschka & Mirjalili, Python Machine Learning, 2019).

For clarification, Figure 8 explains notations that were used throughout this thesis. One row in the data matrix was called a *sample* and represented an observation (e.g. an image). The samples in the data were distinguished by their index (or other unique labels). The samples' measurements were listed as features along with their known target (class label).

Figure 8: Notation for referring to the different data structures samples, indices, features, and classes (class labels).

2.3.1. Modelling with Linear Discriminant Analysis (LDA)

Dimensionality reduction is one way to decrease the complexity of a model, thereby preventing overfitting on training data (Raschka & Mirjalili, Python Machine Learning, 2019). Like PCA, linear discriminant analysis (LDA) is a technique for reducing dimensionality and hence lowering generalization error on new data. However, LDA does not maximize the explainable variance in the chosen component axes. Instead, LDA seeks to maximize class separability in the component axes with known class labels of the data. As LDA can be used as a classification algorithm (classifier), it would be regarded as a supervised classifier.

LDA combines Bayes theorem (Equation 1) and multivariate normal distribution (Equation 2) and seek to maximize the linear discriminant functions (Equation 3) to estimate $P(y = k|x)$, the probability that an observation's class y is class k given observation x (sample x) (Hastie, Tibshirani, & Friedman, 2008). This probability itself may not be straight forward to calculate, but Bayes theorem makes it easier:

$$P(y = k|x) = \frac{P(x|y = k)P(y = k)}{P(x)}$$

Equation 1

where $P(x|y = k)$ is the probability of observation x given y is class k, $P(y = k)$ is the probability of y is class k, and $P(x)$ is the probability of observation x occurring. $P(x|y = k)$ is then found by assuming that the class densities are multivariate Gaussian distributed and using the multivariate normal distribution function:

$$f_k(X) = \frac{1}{(2\pi)^{p/2}|\Sigma_k|^{1/2}} e^{-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1}(x-\mu_k)}$$

Equation 2

where p is the number of dimensions (variables), μ_k is the expected value, and Σ_k is the covariance matrix that is assumed to be equal for all classes.

The linear discriminant functions is given as:

$$\delta_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2}\mu_k^T \Sigma^{-1}\mu_k + \log \pi_k$$

Equation 3

The parameters in the Gaussian distribution are estimated by training data where $\hat{\pi}_k = \frac{N_k}{N}$, $\hat{\mu}_k = \sum_{g_i=k} \frac{x_i}{N_k}$, $\hat{\Sigma} = \sum_{k=1}^{K} \sum_{g_i=k} \frac{(x_i - \hat{\mu}_k)(x_i - \hat{\mu}_k)^T}{(N-K)}$, where N is the number of observations, and N_k is the number of observations in class k.

This thesis used the LDA classifier from *scikit-learn* (sklearn.discriminant_analysis.LinearDiscriminantAnalysis, u.d.).

2.3.2. Other classifiers

As will be described in methods (chapter 3.4.3), different classifiers were investigated for the best classification performance on the OUC data. The following classifiers were not chosen for final model development , and thus will not be explained in detail:

- **Logistic regression**, despite of the misleading name, is an algorithm used for classification. It is not surprising that it is popular in industry since it generally achieves satisfying results on classes that are linearly separable (Raschka & Mirjalili, Python Machine Learning, 2019). The parameters in this model are linear and derived from the minimization of the sum of the squared residuals (Kuhn & Johnson, 2013). This thesis used an implementation of this classifier from Sklearn (sklearn.linear_model.LogisticRegression, u.d.).

- The adaptive boosting (**AdaBoost**) algorithm is an ensemble method where many weak models are trained iteratively with reweighting the samples importance, and in the end assembling these to a stronger (better) classifier (Hastie, Tibshirani, & Friedman, 2008). This thesis used an implementation of this classifier from Sklearn (sklearn.ensemble.AdaBoostClassifier, u.d.).

- **Random forest** is also an ensemble algorithm, that is composed of many decision trees (sequences of discriminating questions) (Raschka & Mirjalili, Python Machine Learning, 2019). This thesis used an implementation of this classifier from Sklearn (sklearn.ensemble.RandomForestClassifier, u.d.).

- **Support vector machines** seek to separate data with decision boundaries (hyperplanes) that are the farthest away from each other. The supporting vectors are the training samples closest to the hyperplanes (Raschka & Mirjalili, Python Machine Learning, 2019). This thesis used an implementation of this classifier from Sklearn (sklearn.svm.SVC, u.d.).

- **Gaussian Naïve Bayes** is similar to LDA; it is based on Bayes theorem but assumes a univariate normal distribution (Hastie, Tibshirani, & Friedman, 2008). This thesis used an implementation of this classifier from Sklearn (sklearn.naive_bayes.GaussianNB, u.d.).

- **K-nearest neighbours classifier** is a *lazy learner* because it remembers the training data instead of learning a discriminative function; a new sample will be assigned to the majority class among the *k* nearest neighbouring training samples (Raschka & Mirjalili, Python Machine Learning, 2019). This thesis used an implementation of this classifier from Sklearn (sklearn.neighbors.KNeighborsClassifier, u.d.).

3. Material and Methods

3.1. Software

The laptop *EliteBook 8560w* (Intel® Core™ i7-2860QM CPU @ 2.50GHz 2.50 GHz with 16.0 GB RAM, and Windows 10 Pro 64-bit operating system) was used for analysis. This implies that the hardware required for replicating the study is not particularly financially limiting. Computational time will vary depending on hardware specifications. In this study, the most long-lasting analyses took roughly 12 hours to complete and were, therefore, run overnight. But, once the results from these analyses are stored - one can classify an unknown sample in a matter of seconds.

The analysis was conducted with the programming language *Python* (v. 3.7.4) through the integrated development environment (IDE) *Spyder* (v. 3.3.6). Solely open-source software was used throughout the entire study.

3.1.1. Codes for data analysis

Table 2 lists the python codes used for data analysis in this study, and Figure 9, 2 and 3 visualize their dependency of each other. Most codes were written by the author, except for two codes for pre-processing spectra and two codes for splitting data.

The pre-processing code for multiplicative scatter correction (MSC) was found on the *NIRPY Research Blog* by Daniel Pelliccia (Pelliccia, Two scatter correction techniques for NIR spectroscopy in Python, 2018) and is also available at his *GitHub* repository (nevernervous78, u.d.).

An implementation of the proposed baseline correction with Asymmetric Least Squares (Eilers & Boelens, 2005) was found on the open community *Stack Overflow* (Python baseline correction library, u.d.).

Each of the two PY files *fcn_SGS_split* and *fcn_SGKF_split* contained a cross-validation object that were not implemented in the library *scikit-learn* at the time of conducting this study. These were found on respective *GitHub* repositories (scikit-learn/sklearn/model_selection/_split.py, u.d.) and (scikit-learn Stratified GroupKFold #13621, u.d.). They were needed in order to include sample group information when splitting data. More about this is given in chapter 3.4.2.

Table 2: List of python codes used for data analysis. The first column denotes the numbering of the codes used in the IDE Spyder, the second column contains the PY file name, column three gives a short description of the code, and the two last columns respectively report whether or not the codes were written by the author and, if they were, how many lines were written.

No.	PY file name	Description	Coded by author	Lines of code
1	fcn_baseline	Function that returns baseline corrected spectra given input spectra.	No	-
2	fcn_clf_vs_classes	Function that calculate statistics for each combination of classifier and class.	Yes	151
3	fcn_confstat	Function that returns a dictionary containing the calculated accuracy for each class given input confusion matrix.	Yes	38
4	fcn_dropcol	Function that implements the algorithm of Sequential Backward Selection (SBS) by using the function from fcn_mean_score.py (see file no. 11). This implementation was made to enable tracking of the prediction of each sample at all times.	Yes	127
5	fcn_featg_vs_classes	Function that calculates statistics for each combination of feature group and class.	Yes	155
6	fcn_featg_vs_clf	Function that calculates statistics for each combination of feature group and classifier.	Yes	150
7	fcn_first_derivative	Function that returns a first derivative transformed array of absorbance values given an array of absorbance values and respective array of corresponding wavelengths.	Yes	26
8	fcn_keep_labels	Function that returns a dataset containing the features that were specified in an input list of features, given input list and dataset.	Yes	23
9	fcn_make_dataset	Function that imports the extracted features for SEM images at one magnification and concatenates them together into one dataset which the function returns.	Yes	111
10	fcn_make_dataset_allmagn	Function that imports the extracted features for SEM images at all magnifications and concatenates them together into one dataset which the function returns.	Yes	209
11	fcn_mean_score	Function that returns the training and validation performance of running LDA over a specified number or random splits on a dataset. The function also makes it possible to track what each sample has been classified as.	Yes	105
12	fcn_msc	Function that returns a multiplicative scatter corrected dataset of spectra given input dataset of spectra.	No	-
13	fcn_nCV	Function that runs n-CV and returns a dictionary of confusion matrices, all the selected hyperparameter sets, and average time for each inner fold in the n-CV, for all combinations of classifiers and feature groups.	Yes	171
14	fcn_predict	Function for predicting unknown UOC sample.	Yes	167
15	fcn_remove_corr_feats	Function that returns a correlation filtered dataset given input dataset and correlation threshold.	Yes	38
16	fcn_screening_make_pressed_dataset	Function that is a modification of fcn_make_dataset.py to fit the dataset originating from the pressed UOC samples.	Yes	111
17	fcn_SGKF_split	This code provides a Stratified K-Folds iterator with non-overlapping groups.	No	-
18	fcn_SGS_split	This code provides a Stratified GroupShuffleSplit cross-validator-	No	-
19	run_main	Main code for running the analysis of feature selection and performance estimation of models developed on test sets.	Yes	1026
20	run_predict	Main code for predicting an unknown UOC sample.	Yes	30
21	run_screening	Main code for creating three heatmaps based on the n-CV. One where classifiers versus feature groups are shown, another where classifiers versus classes are shown, and the last where feature groups versus classes are shown. The frequency of hyperparameter sets used in outer folds in n-CV are also created, along with confusion matrices.	Yes	311
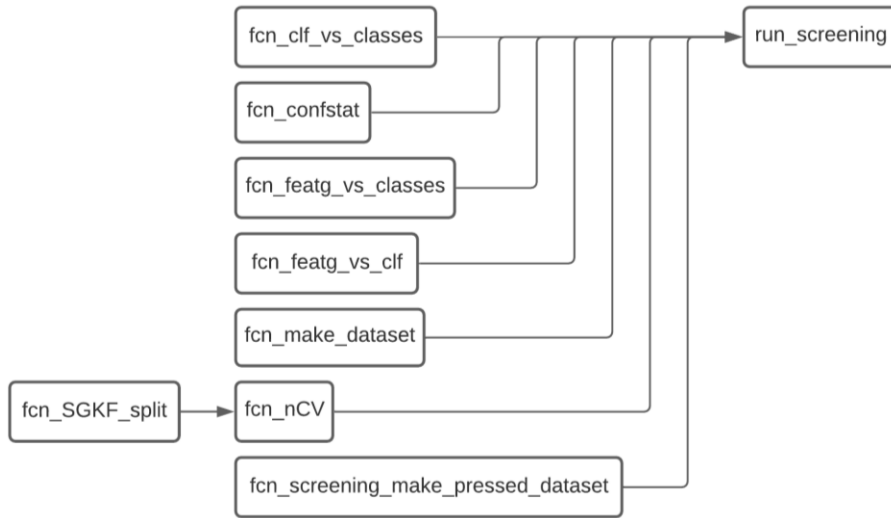
Figure 9: The dependencies of codes used for conducting the initial investigation of performance using different classifiers and feature groups on data from SEM images of pressed and unpressed black UOCs. The PY file names are written inside each box. For example, run_screening.py imports functionality from fcn_nCV.py. Functionality is also imported from fcn_SGKF_split.py into fcn_nCV.py.
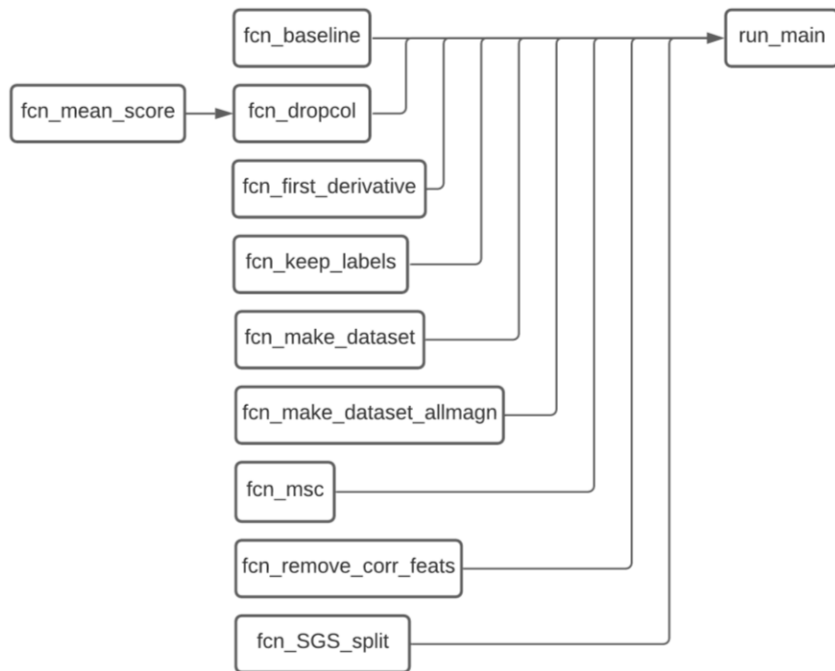
Figure 10: The dependencies of codes used for conducting the feature selection and model performance testing. The PY file names are written inside each box. For example, functionality from fcn_baseline.py is imported into run_main.py.
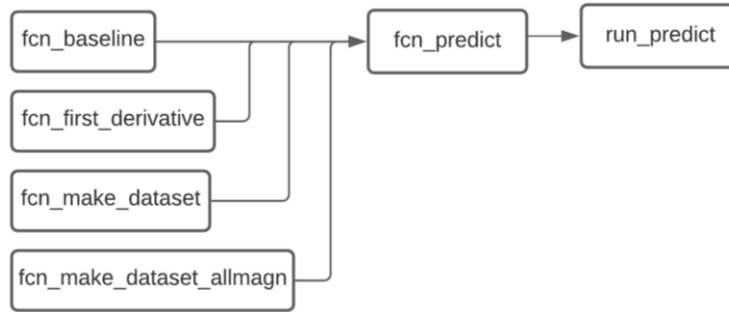
Figure 11: The dependencies of codes for predicting an unknown sample. The PY file names are written inside each box. The functionality of fcn_baseline.py, fcn_first_derivative.py, fcn_make_dataset.py, and fcn_make_dataset_allmagn.py is imported into fcn_predict.py, which in turn run_predict.py imports.

3.1.2. Codes for extracting features from images

Table 3 summarizes the different processes for acquiring features. The features included in the datasets for the SEM images were from the feature groups first order statistics (FOS), local binary patterns (LBP), gray level co-occurence matrix (GLCM), gray level size zone matrix (GLSZM), gray level run length matrix (GLRLM), and angle measure technique (AMT). The features included in the datasets for the hyperspectral images were exclusively spectra. The dependencies of the code for acquiring these spectra are shown in Figure 12. These codes are described in Table 4 along with the code for extracting LBP features.

The AMT features were extracted by the plugin *jAMT Explorer* (Kvaal) in the open source Java image processing program *ImageJ*. Knut Kvaal, *Norwegian University of Life Sciences* (NMBU) authored the code. Both settings "max scale" and "random samples" were set to 500, "statistical method" set to "MA" and "unfold type" set to "Spiral".

The LBP features were extracted with a code written by Linn E. Sogn and Anja K. Smit in their master's thesis in 2018 (Smit & Sogn, 2018). This code was modified to fit this study, also to be compatible with recent upgraded packages.

The other features, FOS, GLSZM, GLRLM, and GLCM, were acquired by using the feature extraction functionality from the *Biorad* project developed by Ahmed Albunni. The project is available on his GitHub repository (Albunni, u.d.) and uses the open-source package *pyradiomics* (van Griethuysen, et al., 2017).

Table 3: The process for extracting different feature groups from images. The groups are given in the top row and their corresponding process for acquiring these features.

Type of features	AMT	LBP	FOS, GLSZM, GLRLM, GLCM	Spectra
Process for feature extraction	In *ImageJ*, the plugin *jAMT Explorer is run* on a stack of images with max scale at 500 and 500 random samples	The PY file *extract_lbp_features_MODIFIED* is run	The PY file *feature_extraction* (from the *Biorad* project at (https://github.com/ahmedalbuni/biorad) is run in *Anaconda Prompt* with "glcm_distance 1,3,10,15,20"	The PY file *run_hyperspectral* is run

fcn_hyperspctral_mean_cube_dataset → run_hyperspectral
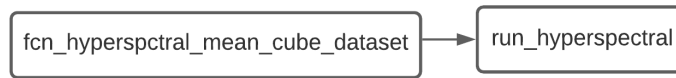
Figure 12: The dependencies of codes used for extracting spectra and creating datasets from hyperspectral images. The PY file names are written inside each box. The functionality from fcn_hyperspectral_mean_cube_dataset.py is imported into run_main.py.

Table 4: List of python codes used for feature extraction. The first column denotes the numbering of the codes used in the IDE Spyder, the second column contains the PY file name, column three gives a short description of the code, and the two last columns respectively report whether or not the codes were written by the author and, if they were, how many lines were written.

No.	PY file name	Description	Coded by author	Lines of code
22	extract_lbp_features_MODIFIED	This code was made by Linn E. Sogn and Anja K. Smit in 2018. The code extracts local binary pattern features from images and saves them as a dataset. This code is modified by the author to fit this study.	Modified	-
23	fcn_hyperspctral_mean_cube_dataset	Function that returns the extracted dataset of spectra.	Yes	173
24	run_hyperspectral	Main code for extracting the dataset of spectra from selected hyperspectral images.	Yes	101

3.2. Investigated materials

3.2.1. Uranium Ore Concentrates

The investigated uranium ore concentrates (UOCs) originated from a total of 48 different facilities across 17 countries, with varying chemical composition. They were available at EU Commission JRC, Karlsruhe.

Table 5 lists the UOCs used in this study, along with information of their origin, chemical composition, label used in the analysis, and whether they were acquired with scanning electron microscope (SEM) or hyperspectral imaging (HSI). A priori information used for colour categorization of the UOCs was provided by Lorenzo Fongaro, EU Commission JRC Karlsruhe. In his paper *Image texture analysis and colorimetry for the classification of uranium ore concentrate powders* a method for classifying UOCs by their colour was proposed. This was done by applying hierarchical clustering on measurements taken by a spectrophotometer on the UOCs (Marchetti, et al., 2019). The paper reports that six colour classes were found. In this study, UOCs from five of these classes are investigated. To prevent confusion, the colour classes from the study will be named colour categories throughout this thesis, while classes refer to the sample's origin.

The UOCs for each facility were distributed into sample holders as preparation before image acquisition as illustrated in Figure 5. Most of the facilities were represented by three sample holders, but not all UOCs had enough quantity to be distributed into more than two sample holders.

As Table 5 shows, both pressed and unpressed UOCs were used. The meaning of "pressed" is that sample preparation of UOCs included some sort of pressing (Fongaro, Ho, Kvaal, Mayer, & Rondinella, 2016). Images of unpressed UOCs were acquired after it was concluded that the pressed sample preparation practice was suboptimal, according to Lorenzo Fongaro. In this study, the pressed UOCs are only used in the initial analysis along with unpressed UOCs belonging to the same facility for comparison.
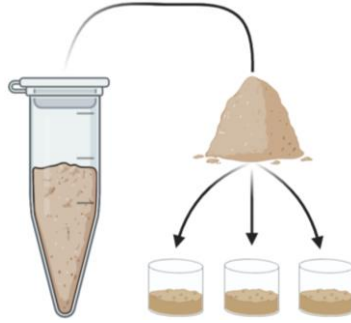
Figure 13: Illustration of the distribution of UOC into three sample holders.

Table 5: The UOCs investigated in this study. Information of their respective colour category, labels used in the analysis, origin, chemical composition, whether they were acquired with SEM or HSI, and ID No. is shown. The ID number is important to discriminate the hyperspectral images from each other as their filenames include this number but no abbreviation for origin. There are three columns containing label names because the image filenames were not named consistently. The abbreviation n.d. means "not declared". It was observed that the labelling of UOC from Rossing (Namibia) seems to be wrong as both prefixes "Nia" and "Sa" occurred. The labels are derived from the filenames of the images.

	Pressed SEM labels	Unpressed SEM Labels	HSI Labels	Country	Facility	Composition	SEM	HSI	ID No.
Colour category 1 (Black-Dark Brown)	NiaRos	SaROS		Namibia	Rossing	Peroxide	Yes	No	5
			Zam_Min	Zambia	Mindola	Peroxide	No	Yes	6
	USAPet	UsPET		USA	Pathfinder	Peroxide	Yes	No	9
	SAfPal	SaPAL		S.Africa	Palabora	Peroxide	Yes	No	13
	AusQue	AuQUE		Australia	Queensland	Peroxide	Yes	No	17
	CanKel	CaKEL	Can_Key	Canada	Key Lake	Peroxide	Yes	Yes	21
	ChiHeY	ChHEY	Chi_Hen	China	Hengyang	Peroxide + Oxide	Yes	Yes	22
	YugSpB	YuSPB		Yogoslavia	Spisak Black	Peroxide	Yes	No	26
	AusOID	AuOLD	Aus_Oly	Australia	Olympic Dam	Peroxide	Yes	Yes	28
	USAAtl	UsATL		USA	Atlas	Peroxide + Oxide	Yes	No	57
	AusMak	AuMAK		Australia	Mary Kathleen	Peroxide	Yes	No	58
	USAFAP	UsFAP		USA	Federal American Partners	Peroxide	Yes	No	62
	SAfNuf	SaNUF		S.Africa	Nufcor	Peroxide	Yes	No	69
	RusTex	RuTEC	Rus_Tec	Russia	Techsnab	Peroxide	Yes	Yes	70
Colour category 3 (Light Brown-Dark Yellow)		Yeelir		Australia	Yeelirre	n.d.	Yes	No	4
		Cotter	USA_Cot	USA	Cotter	n.d.	Yes	Yes	7
		BeCong		Belgian	Congo	Hydroxide	Yes	No	14
		Nucleb	Bra_Nuc	Brazil	Nuclebras	ADU	Yes	Yes	15
		Wismut	Ger_Wis	Germany	Wismut	ADU	Yes	Yes	18
		CaDyno	Can_Dyn	Canada	Dyno	n.d.	Yes	Yes	20
			Can_Sun	Canada	Sunnar	Hydroxide	No	Yes	25
			Can_Far	Canada	Faraday	Hydroxide	No	Yes	27
		SpisYe	Yog_Spi	Yogoslavia	Spisak-Yellow	ADU	Yes	Yes	29
		Fallsc	USA_Fal	USA	Falls City	n.d.	Yes	Yes	31
		USAESI		USA	ESI	ADU	Yes	No	34
		macass		Canada	Macassa	Hydroxide	Yes	No	35
		Anacon		USA	Anaconda	ADU	Yes	No	36
		RadiHi	Aus_Rad	Australia	Radium Hill	ADU	Yes	Yes	40
		RumJun	Aus_Run	Australia	Run Jungle	n.d.	Yes	Yes	65
			Yog_Rud	Yogoslavia	Rudnik	ADU	No	Yes	72
Colour category 4 (Dark Yellow-Yellow)		KMcGee		USA	Kerr McGee	ADU	Yes	No	2
		SpaGen	Spa_Gen	Spain	Gen	ADU	Yes	Yes	3
		GabEFI		Gabon	EFI(Mouand)	ADU	Yes	No	8
			Can_Nor	Canada	North Span	Hydroxide	No	Yes	41
		Ellwel	Ger_Hel	Germany	Helwiler	ADU	Yes	Yes	45
		Ransta	Swe_Ran	Sweden	Ranstadt	SDU	Yes	Yes	53
		Deniso		Canada	Denison	ADU	Yes	No	56
		Millik	Can_Mil	Canada	Milliken Lake	ADU	Yes	Yes	64
		southa	Aus_S A	Australia	S Alligator	Hydroxide	Yes	Yes	73
		Romani	Rum_Rum	Rumania	Rumania	SDU + Oxide	Yes	Yes	76
Colour category 5 (Yellow-Light Yellow)		StanRo		Canada	Stamrock	ADU	Yes	No	10
		HDelft		Holland	Delft	ADU	Yes	No	11
		ElMesq		USA	El Mesquite	n.d.	Yes	No	54
		USDawn		USA	Dawn	ADU	Yes	No	63
Color category 6 (Light Yellow-White)		Irigar	USA_Iri	USA	Irigaray	Peroxide	Yes	Yes	33
		UMobil	USA_Mob	USA	Mobil	Peroxide	Yes	Yes	38
		Rabbla		Canada	Eldore(Rabbit Lake)	Peroxide	Yes	No	51
		EverYe	USA_Eve	USA	Everestr-Yellow	Peroxide	Yes	Yes	52

3.2.2. Image acquisition

3.2.2.1. Scanning electron microscope images

Unpressed SEM images were acquired from five regions for each sample holder, as illustrated in Figure 14. Within these regions, at least three images at different magnification were acquired. These three images originating from the same area in the sample holder overlapped, as depicted in Figure 15. Some UOCs were acquired at the four magnifications 100x, 250x, 500x, and 1000x. Only three magnifications per sample were used in this study, due to inconsistent practices. By doing this, the UOCs belonging to the same colour category had the same origin across the acquired magnifications. The underlying cause of inconsistent acquisition across the magnifications was that during 1000x magnification particles of the UOC started moving due to charging effects induced by the SEM.
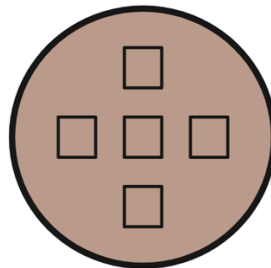
Figure 14: Illustration of the regions within a sample holder where SEM images were acquired, looking down onto the sample holder. The outer circle ilustrates the sample holder wall and the squares within are the regions.
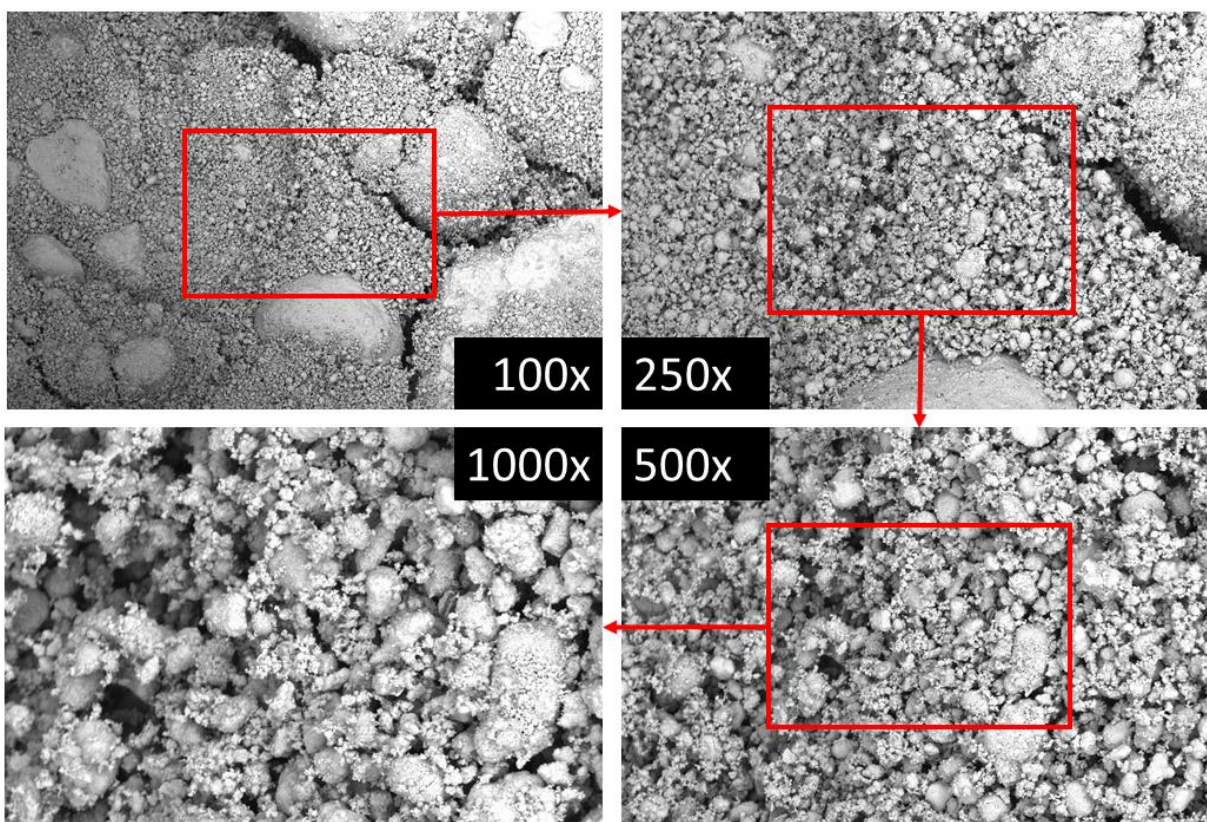
Figure 15: Example of SEM images acquired from the same region in one sample holder at the different magnifications. The magnification of each image is showed in the centre corners of the images. The red regions mark the area in which the next magnified image represents. The UOC imaged originates from the facility Key Lake in Canada.

3.2.2.2. *Hyperspectral images*

One hyperspectral image was acquired for each sample holder (containing unpressed UOC) with the push-broom hyperspectral camera *Specim FX17e*, with 224 bands in the range of 936 *nm* to 1720 *nm*. Four sub-images were picked out from each hyperspectral image for use in the analysis, Figure 16 illustrates the process. This was done in the same code as where spectra were extracted from the hyperspectral images.
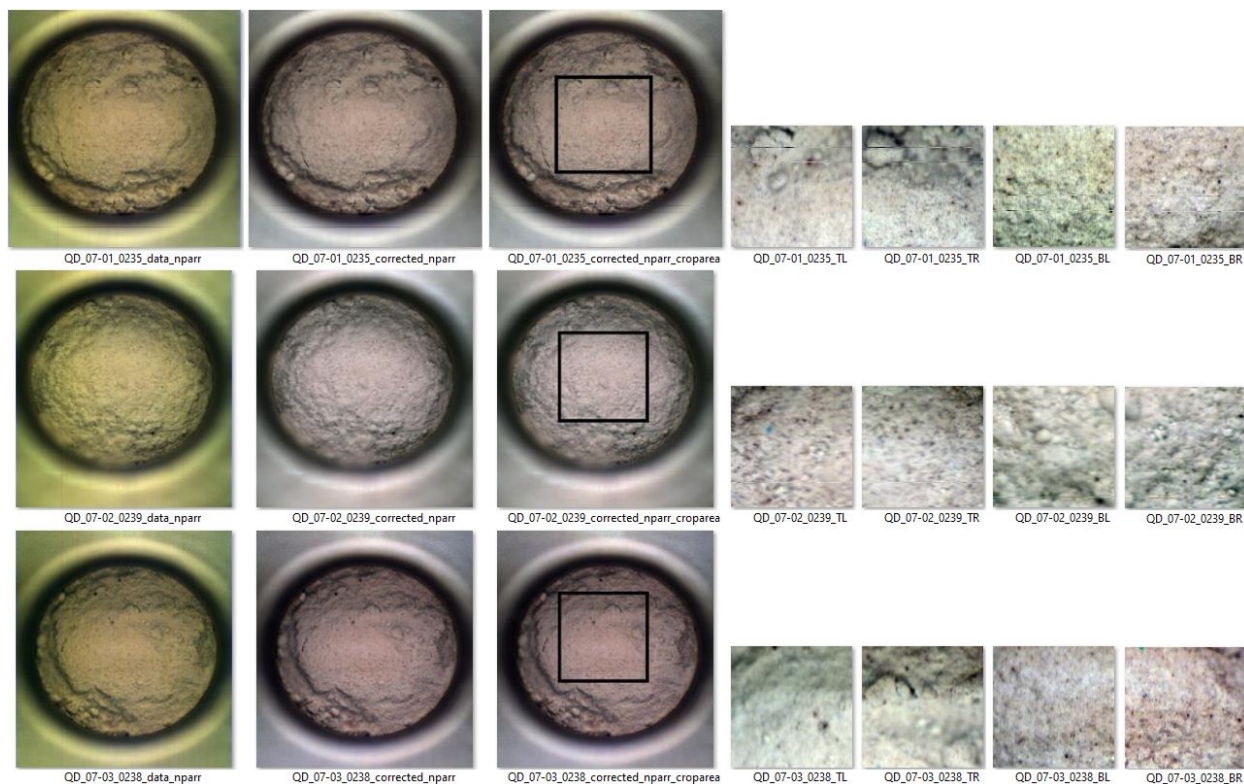
Figure 16: The process of acquiring the hyperspectral images for the analysis from the raw hyperspectral image. The columns of images show (from left to right) the raw hyperspectral image, the reflectance calibrated image, then the same image again but with a square selecting a cropped area, and the last four images are sub-images from the cropped area. These equally sized sub-images from the cropped area were used for feature extraction. The two last characters (suffix) in the image name denotes the location of the sub-image. T means top, B means bottom, L means left, and R means right. All hyperspectral images are shown with the RGB bands [20, 120, 220].

3.3. Creation of datasets for analysis

3.3.1. Preparation of images

After cropping and rescaling, the acquired SEM images of pressed UOCs had pixel dimensions of 1225 x 925. The dimensions for the unpressed UOCs are listed in Table 6. The images were used as they were received, in regards to their dimensions. No cropping or rescaled were done, which is the reason for the differences in the number of pixels per dimension between magnifications and colour categories.

Table 6: The pixel dimensions for each acquired set of SEM images for each colour category and magnification of unpressed UOCs.

Magnification	CC1	CC3	CC4	CC5	CC6
100x	1500 x 1000	767 x 513	3000 x 2000	767 x 512	767 x 511
250x	1500 x 1000	768 x 512	3000 x 2000	767 x 510	767 x 512
500x		768 x 516	3000 x 2000	767 x 510	767 x 513
1000x	1500 x 1000				

The images were pre-processed in order to centre the image histogram and create a better foundation for cross image comparison. The was done by using the plugin *Stack Meancenter5* (Kvaal) in *ImageJ*. The plugin centred the images by subtracting the pixel values by their average pixel value, then dividing the pixel values by their standard deviation (SD). Therefore, the minimum pixel value was subtracted from all pixels. The average, minimum, and SD was calculated for each image individually.

3.3.2. Feature extraction

Features were extracted from different algorithms. Their type of algorithm grouped the extracted features. These were denoted as feature groups and used in the initial assessment of classifiers.

3.3.2.1. Spectra

The spectrum used as data for each sample was obtained following a series of pre-processing steps as outlined in chapter 3.4.1.2. First, the raw spectra were reflectance calibrated and then the hyperspectral image was divided into four sub-images. Each of the sub-image spectra was averaged over all spatial pixels followed by conversion to absorbance. The code used for creating the datasets from hyperspectral images is presented in chapter 3.1.2.

3.3.2.2. First order statistics (FOS)

First order statistics features were extracted by using the GitHub repository (Albunni, u.d.) as described in chapter 3.1.2. These features describe the distribution of images´ pixel intensity values by different measures (Radiomic Features, u.d.). Table 42 in the appendix lists each of the extracted first order features used.

3.3.2.3.   Local Binary Pattern (LBP)

Local Binary Pattern features were extracted by using a code originally developed by Smit and Sogn (Smit & Sogn, 2018), which uses the open-source package *scikit-image* for calculating these features (Local Binary Pattern for texture classification, u.d.). The package refers to the article (Ojala, Pietikäinen, & Mäenpää, 2002) for the method of LBP, and it presents a way to measure local texture.

Given a number of neighbouring points p evenly distributed on a circle of pixels with radius r from a centre pixel, the neighbouring pixels are thresholded by the centre pixel so that pixels with an intensity value greater than the centre pixel are set to 1 and otherwise to 0. Then, the neighbouring thresholded pixel values around the circle (in a consistent way), are unfolded to a binary number and then translated to a natural number. This is the calculated LBP value, which is assigned to the centre pixel as a new pixel value. This is repeated for all pixels in the image. Then, a histogram of all these LBP values are created and used for further calculations (Pietikäinen, Hadid, Zhao, & Ahonen, 2011).

The code used the "uniform" method for calculating the features, which means that an improved rotation and gray scale invariant method was used (Pietikäinen, Hadid, Zhao, & Ahonen, 2011). Also, 11 pairs of neighbours p and radius r were chosen. These were $p = (i + 1)$ x 4 and $r = (i + 1)$ for all i in $\{i \in \mathbb{Z} \mid 0 \leq i \leq 10\}$.

3.3.2.4.   Gray level Co-occurence Matrix (GLCM)

Gray level Co-occurence Matrix features were extracted by using the GitHub repository (Albunni, u.d.) as described in chapter 3.1.2. GLCM is a invariant method for quantifying the number of

times two pixels of intensity value i and j separated by a distance d appear along a specific angle (Radiomic Features, u.d.). The distances used in this study were 1, 3, 10, 15, and 20, and are included in the added suffix to the feature names. The X in the suffix "_d_X" denotes the pixel distance. The prefix "glcm_" was also added. This was done to make it easier to distinguish different feature groups. Table 44 in the appendix lists each of the extracted GLCM features that were used.

3.3.2.5.   Gray Level Size Zone Matrix (GLSZM)

Gray Level Size Zone Matrix features were extracted by using the GitHub repository (Albunni, u.d.) as described in chapter 3.1.2. GLSZM is a rotation invariant method for quantifying zones of horizontal, vertical, and diagonal neighbouring pixels of the same intensity value (Radiomic Features, u.d.). Table 44 in the appendix lists each of the extracted features used. The prefix "glszm_" was added to these feature names in the dataset to make it easier to distinguish different feature groups.

3.3.2.6.   Gray Level Run Length Matrix (GLRLM)

Gray Level Run Length Matrix features were extracted by using the GitHub repository (Albunni, u.d.) as described in chapter 3.1.2. GLRLM is a method for quantifying the number of consecutive pixels of the same intensity value along a specific angle, it is not rotation invariant (Radiomic Features, u.d.). Table 44 in the appendix lists each of the extracted features used. The prefix "glrlm_" was added to these feature names in the dataset to make it easier to distinguish different feature groups.

3.3.2.7.   Angle Measure Technique (AMT)

The Angle Measure Technique is a method for measuring the texture in an image by using a spectral approach (Halstensen, Kvaal, & Esbensen, 2019). An image is unfolded from 2-D into a

1-D array before further processing. Then, for each point A_n of n randomly selected points, a circle with radius of a scale s is drawn. This circle intersects the unfolded array at two points, B_n and C_n. The angle $C_nA_nB_n$ is then calculated. The n angles calculated at scale s is then averaged, resulting in a feature of scale s. This is repeated for many scales. In this study the scales $s = (i + 2)$ for all i in $\{i \in \mathbb{Z} \mid -2 \leq i \leq 66\}$ were used, resulting in 35 features as listed in Table 43. This method for measuring texture has previously been used in the field of nuclear forensics (Fongaro, Ho, Kvaal, Mayer, & Rondinella, 2016).

3.3.3. Datasets created

An overview of the resulting datasets originating from SEM images of pressed and unpressed UOCs, and HSI UOCs are shown in Table 7, Table 8 and Table 9. Four datasets were constructed from each of the colour categories from SEM images, while one dataset was constructed for each of the categories from the hyperspectral images. The number of samples and features are listed in the tables. The table information should be considered along with Figure 8 in theory, chapter 2.3. The datasets are illustrated in terms of dimensions and content; the number of samples equals the number of rows, and the number of features, along with class labels (and even more information such as sample holder ID for grouping) are the columns. As described in section 1.2.1, most of the **unpressed** UOCs were distributed into three sample holders. Furthermore, each of the sample holders, five images were acquired at each magnification using SEM, and one image per sample holder using HSI. For example, a dataset composed by features extracted from the black-dark brown (CC1) coloured UOCs at any magnification had 195 samples; 13 (classes) $\times$ 3 (sample holders) $\times$ 5 (images). Throughout the available data, some UOCs had fewer sample holders and even missing images.

Table 7: Overview of the number of classes, samples, features, and features in each feature group, in all datasets originating from **pressed** UOC SEM images. The colour category given in the leftmost column is CC1, and the magnification of images included in the dataset are given in the second column from the left.

	Magnification in the dataset	No. of classes	No. of samples	Total no. of features	No. of features in each feature group									
					FOS	LBP	AMT	GLCM 1	GLCM 3	GLCM 10	GLCM 15	GLCM 20	GLRLM	GLSZM
CC1	250x	14	280	491	18	286	35	24	24	24	24	24	16	16

Table 8: Overview of the number of classes, samples, features, and features in each feature group, in all datasets originating from **unpressed** UOC SEM images. The colour categories are shown in the leftmost column, and the magnification of images included in the dataset are given in the second column from the left. There are four datasets per colour category, one for each included set of magnification(s) given by the second column.

	Magnifications in the dataset	No. of classes	No. of samples	Total no. of features	No. of features in each feature group									
					FOS	LBP	AMT	GLCM 1	GLCM 3	GLCM 10	GLCM 15	GLCM 20	GLRLM	GLSZM
CC1	100x / 250x / 1000x	13	195	491	18	286	35	24	24	24	24	24	16	16
	100x + 250 x + 1000x			1473	54	858	105	72	72	72	72	72	48	48
CC3	100x / 250x / 500x	13	194	491	18	286	35	24	24	24	24	24	16	16
	100x + 250x + 500x			1473	54	858	105	72	72	72	72	72	48	48
CC4	100x / 250x / 500x	9	120	491	18	286	35	24	24	24	24	24	16	16
	100x + 250x + 500x			1473	54	858	105	72	72	72	72	72	48	48
CC5	100x / 250x / 500x	4	50	491	18	286	35	24	24	24	24	24	16	16
	100x + 250x + 500x			1473	54	858	105	72	72	72	72	72	48	48
CC6	100x / 250x / 500x	4	53	491	18	286	35	24	24	24	24	24	16	16
	100x + 250x + 500x			1473	54	858	105	72	72	72	72	72	48	48

Table 9: Overview of the number of classes, samples, and features (which are wavelengths) in the datasets originating from the hyperspectral images of unpressed UOCs. The colour categories are shown by the leftmost column. There is one dataset per colour category as the images were acquired at solely one magnification in contrast to the SEM images.

	No. of classes	No. of samples	Total no. of features (spectral bands)	Spectral range	Spectral sampling (step in wavelength)
CC1	5	56	224	935.61 nm - 1720.23 nm	3.5 nm
CC3	11	124	224	935.61 nm - 1720.23 nm	3.5 nm
CC4	7	84	224	935.61 nm - 1720.23 nm	3.5 nm
CC6	3	28	224	935.61 nm - 1720.23 nm	3.5 nm

3.4. Roadmap to developing predictive models

The process for finding the most promising classifier for UOCs classification was found using the black coloured UOCs of both pressed and unpressed samples at magnification 250x. It was thought that the black-dark brown coloured UOCs were the most difficult to discriminate, and it was therefore thought that a classifier able to classify these classes would perform at least as good on the other colour categories. However, it was understood from prior analysis (Raschka & Mirjalili, Python Machine Learning, 2019) that there was no guarantee that the best classifier found on the two datasets would perform any better on the other colour categories or different magnifications. Also, the performance of the chosen classifier on spectral datasets was unknown. Nevertheless, as a start, the best classifier on the two datasets chosen in the beginning was tried on all datasets. The investigated classifiers are given in section 3.4.3.

The upper part of Figure 17 illustrates the process of arriving at the best (most promising) classifier. Each of the two datasets of pressed and unpressed UOCs underwent nested cross-validation for testing the performance of using different classifiers on different feature groups of data. This process is described in more detail in section 3.4.3. The best classifier was considered to be the one that on a reasonably good trade-off between performance and hyperparameter stability, performed well.

The best performing classifier was then used for model training in the process of finding optimized sets of features for each dataset. In the end, the classifier was trained on optimized feature sets for each dataset, as final models, and their performance was tested on hold-out test data. This process is described in more detail in section 3.4.4.
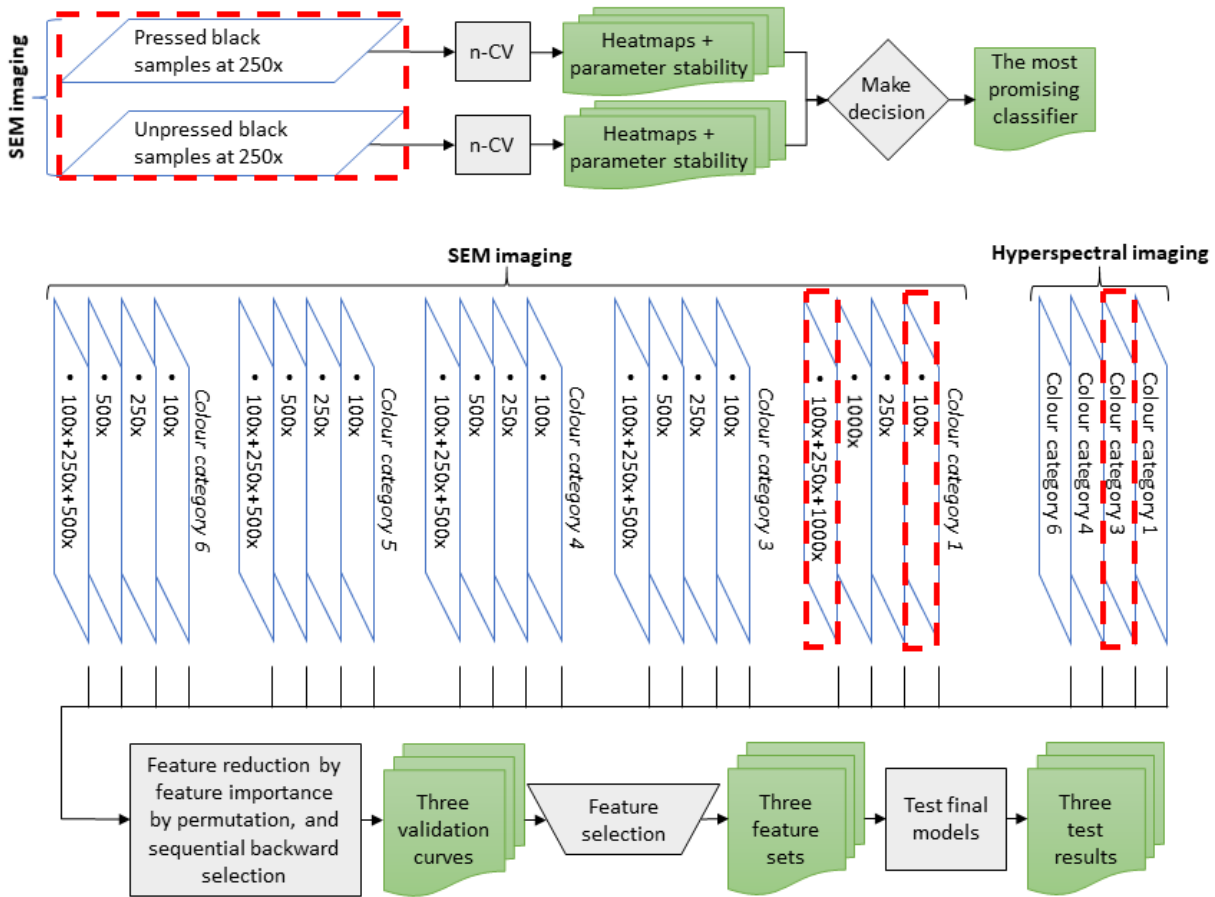
Figure 17: Visualisation of the steps in the analysis and datasets used. The red dotted marked areas show the datasets used for the results included in the results chapter. The grey areas denote the approach used and the green areas show the obtained results.

3.4.1. Pre-processing

3.4.1.1.   Standardization

Standardization is a widely used technique in machine learning for scaling features (Raschka & Mirjalili, Python Machine Learning, 2019). There are many classifiers that are optimized for features with values that are somewhat normally distributed. This is achieved by centring the feature values to a mean of zero and scaling them to unit variance, on each feature independently. Or in other words,

$$\chi_{std}^{(i)} = \frac{\chi^{(i)} - \mu_\chi}{\sigma_\chi}$$

where $\chi_{std}^{(i)}$ is the standardized value i of feature χ, μ_χ is the mean value of feature χ, and σ_χ is the standard deviation of feature χ (Raschka & Mirjalili, Python Machine Learning, 2019). The analysis used the class *StandardScaler* from *scikit-learn* (sklearn.preprocessing.StandardScaler, u.d.). The features in the datasets originating from the SEM images were processed as in the described method above, but the spectral datasets were processed using $\sigma_\chi = 1$ for all spectra.

3.4.1.2. *Preparing spectral data for analysis*

The raw spectra acquired for each hyperspectral sample were pre-processed before the spectra were extracted and inserted into datasets. The spectra used in this study were first reflectance calibrated and then converted to absorbance followed baseline correction, as described in chapters 3.4.1.2.1 and 3.4.1.2.2, respectively. Other pre-processing methods, such as MSC, were also investigated. These are described in chapter 3.4.1.2.3. Figure 18 illustrates the process of arriving at the used baseline corrected spectra from the absorbance spectra, as well as the other methods of processing spectra that were examined.
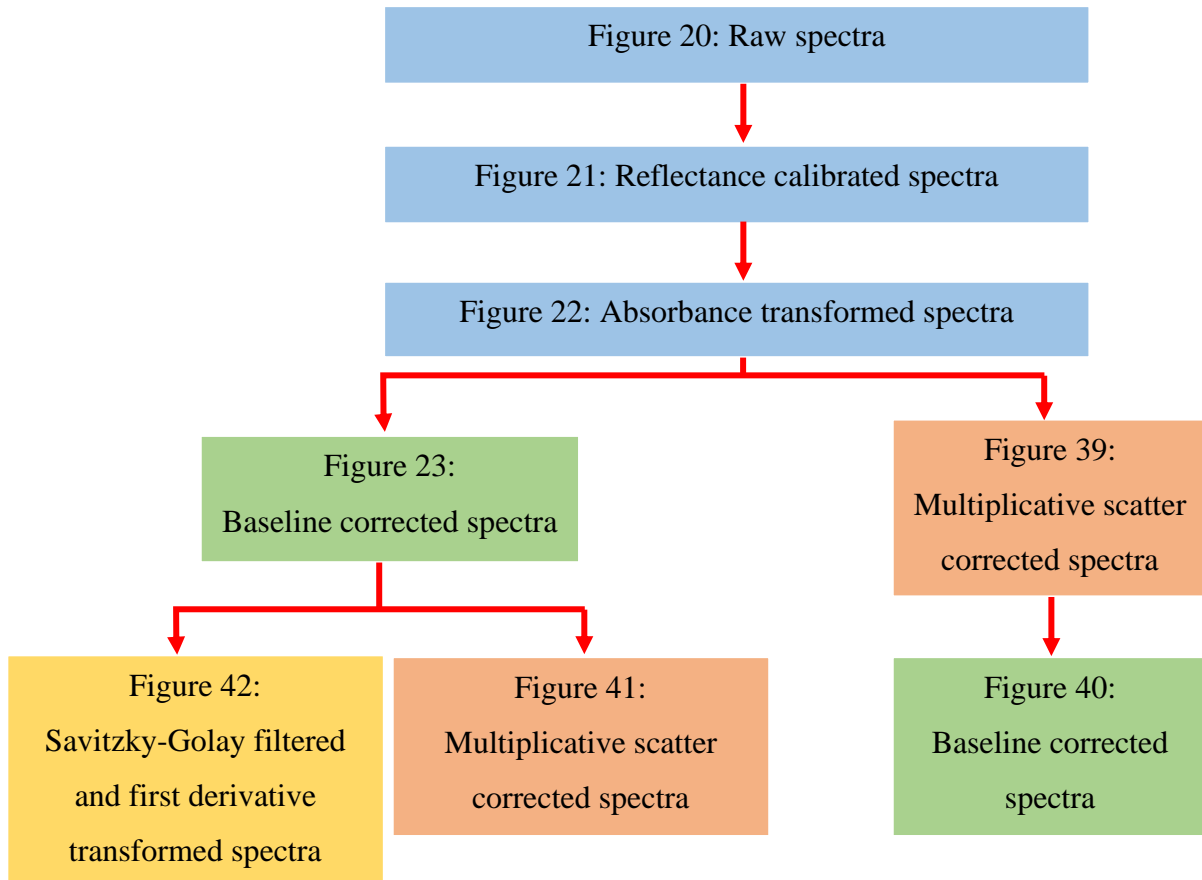
Figure 18: Overview of the processing steps starting on the raw spectra. Each square (except the top one) includes a description of the processing method used on the previous square and refers also to the corresponding figures illustrating the change of the spectra of colour category 3. The change of shape of the spectra can be observed following this flow chart of processes along with the referred figures.

3.4.1.2.1. Reflectance calibration and conversion to absorbance

The reflectance calibrated hyperspectral image *I*, also called as the relative reflectance image, was calculated as

$$I = \frac{I_0 - D}{W - D}$$

where I_0 is the raw reflectance image, *W* represents the white reference image, and *D* the dark reference image (Basantia, Nollet, & Kamruzzaman, 2019). Thereafter, the absorbance image I_A was calculated as

$$I_A = log_{10}(I)$$

where *I* is the relative reflectance image (Basantia, Nollet, & Kamruzzaman, 2019). The effect of reflectance calibrating a hyperspectral image on the appearance can be clearly seen in Figure 19. Vertical noise appears to be removed. The change in spectra when reflectance calibrating the raw spectra can be seen in Figure 20 and Figure 21.
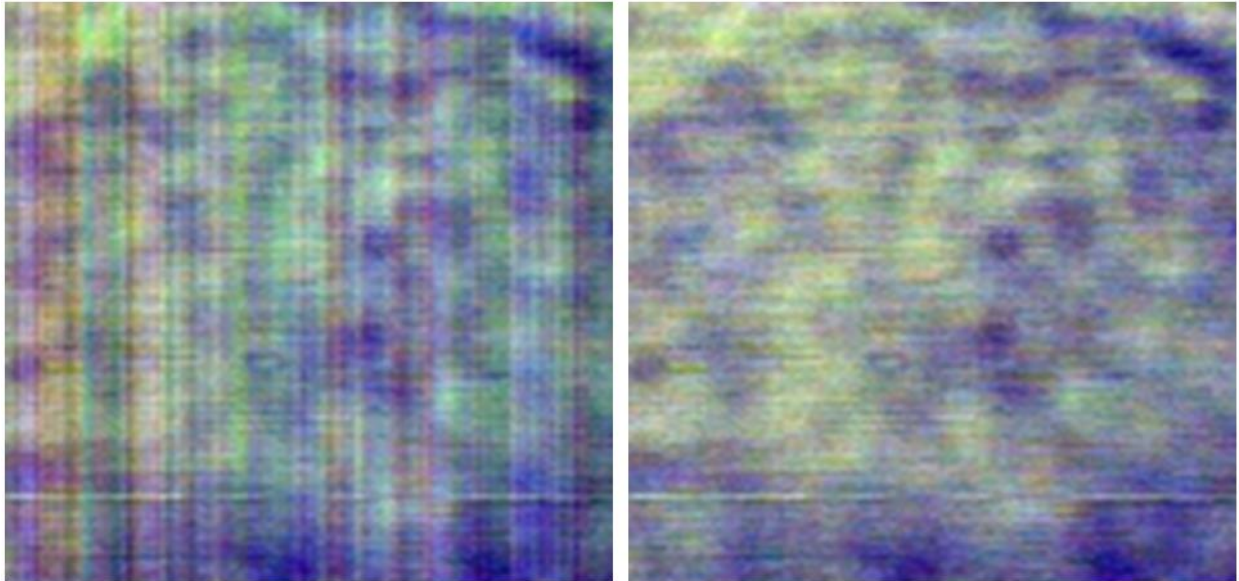
Figure 19: The effect on the spatial appearance of reflectance calibrating a hyperspectral image. The image to the left is the raw image (named QD_15-01_0065_BR) and the image to the right has been calibrated using the white and dark reference image. As hyperspectral images contain pixel intensities for many bands of wavelengths, these images were visualized by choosing the RGB band indices [20, 120, 220].
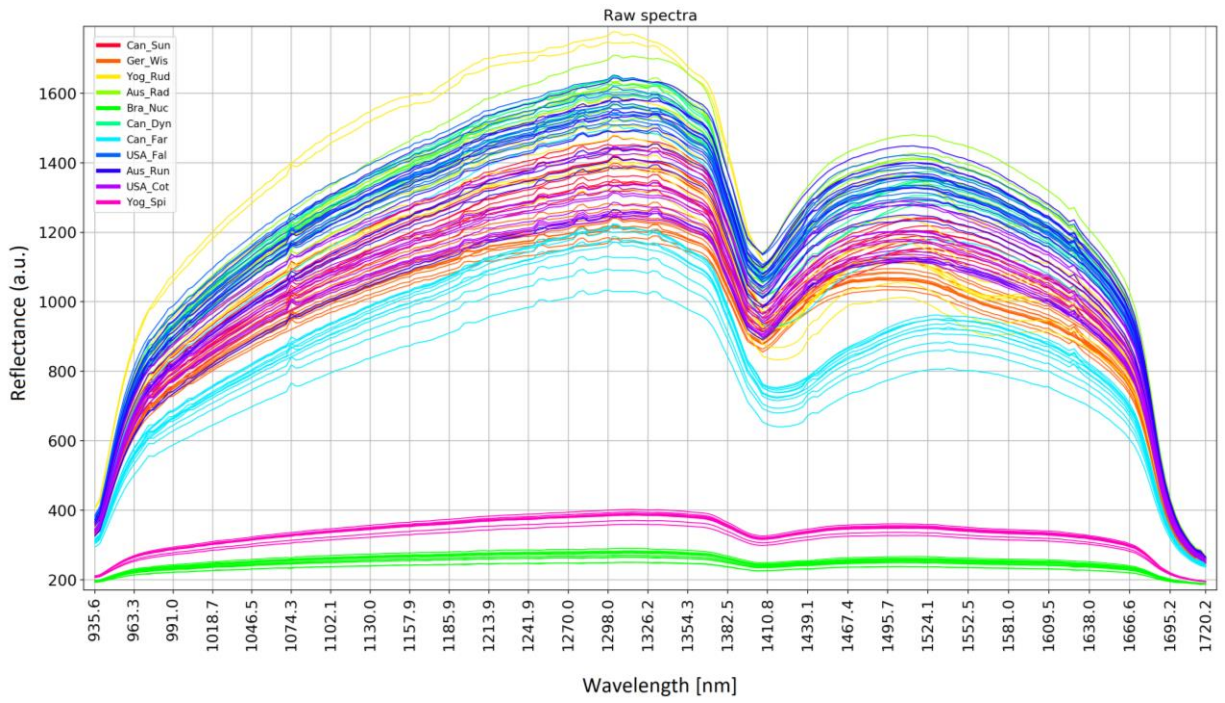
Figure 20: The raw spectra of samples in the colour category 3 dataset. The vertical axis denotes reflectance in arbitrary units and the horizontal axis denotes wavelengths in nm.
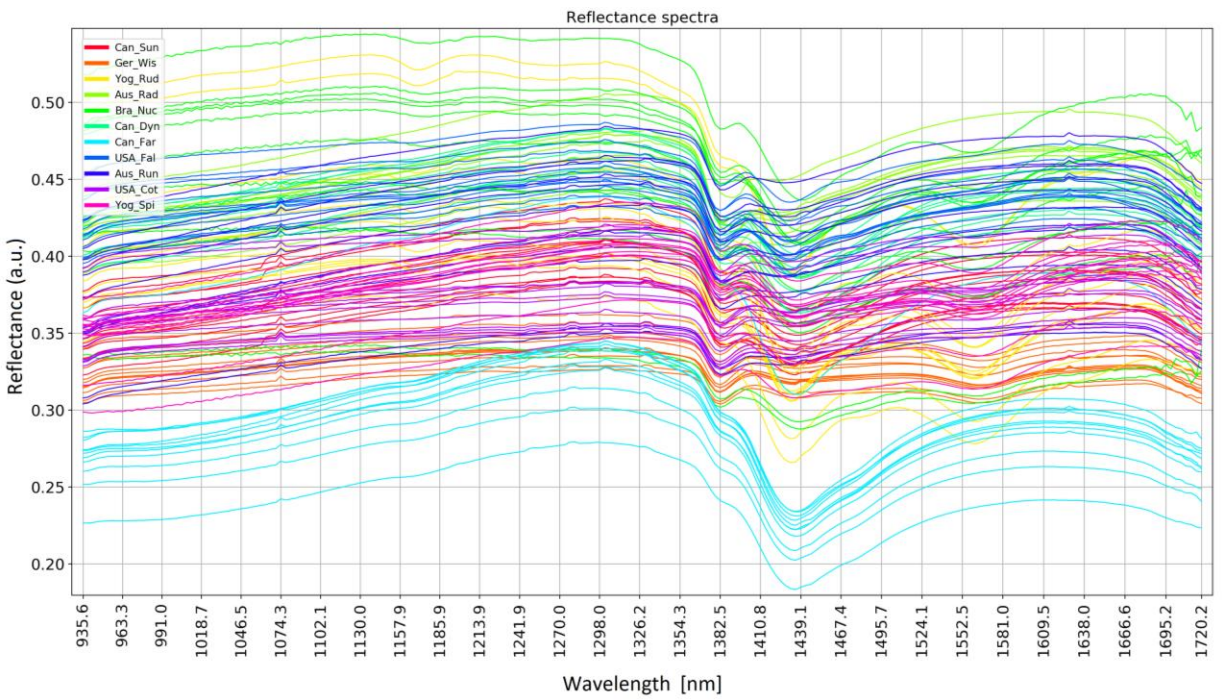
Figure 21: The reflectance calibrated spectra of samples in the colour category 3 dataset. The vertical axis denotes relative reflectance in arbitrary units and the horizontal axis denotes wavelengths in nm.
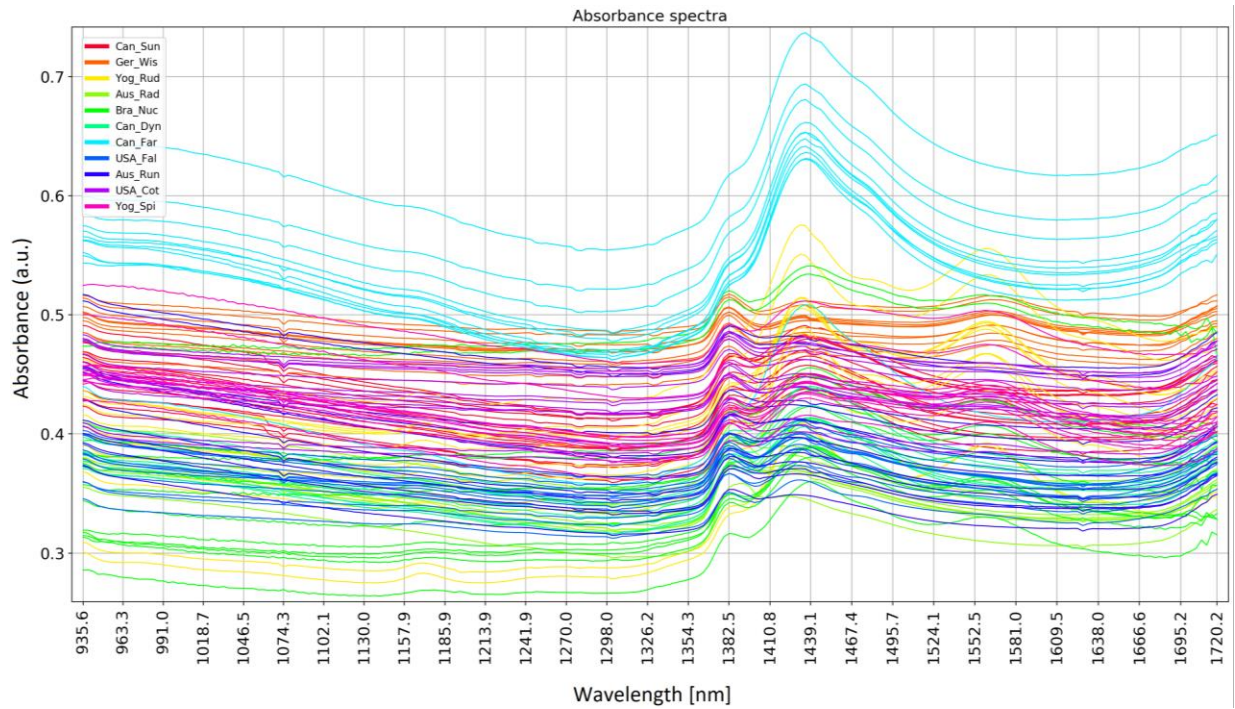
Figure 22: The absorbance converted spectra of samples in the colour category 3 dataset. The vertical axis denotes reflectance in arbitrary units and the horizontal axis denotes wavelengths in nm.

3.4.1.2.2.   Baseline correction

Spectral baseline correction was performed on all spectra in the dataset, using baseline estimations (Eilers & Boelens, 2005). The calculated baseline for each spectrum was subtracted from the spectrum itself. As can be seen by comparing Figure 22 and Figure 23, the flat ranges of the spectra were shifted downwards when the baseline estimations were subtracted.
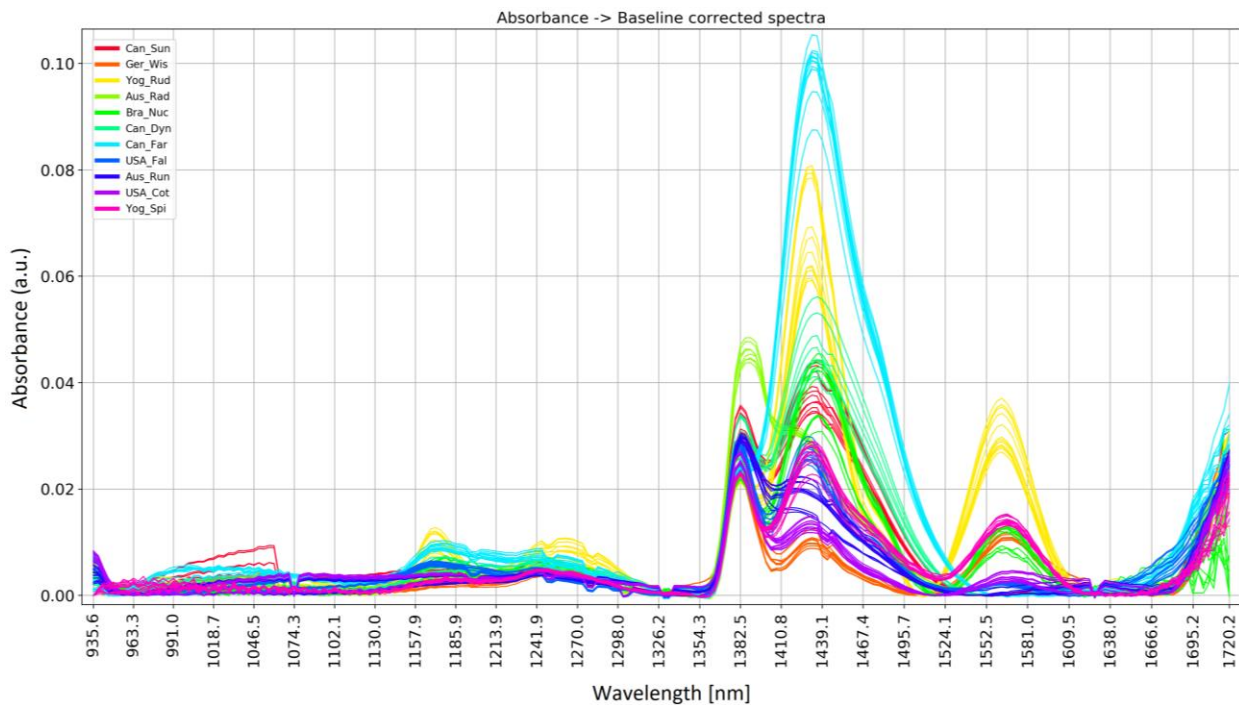
Figure 23: Baseline correction applied on the absorbance converted spectra of samples in the colour category 3 dataset. The vertical axis denotes absorbance in arbitrary units and the horizontal axis denotes wavelengths in nm.

3.4.1.2.3.   Other briefly investigated processing techniques

Other steps applying various techniques for processing spectra were examined, but most only to the extent of visualization. Results of applying the Savitzky-Golay smoothing filter (scipy.signal.savgol_filter, u.d.) and first derivative transformation on baseline corrected absorbance spectra are included in Figure 42 appendix I.

MSC were not used for any analysis due to the arose challenges of ensuring no data leakage (Schutt & O'Neil, 2014). Moreover, since baseline correction was used on the absorbance spectra first and gave excellent results, it was decided to keep the processing as simple as possible – and thus not applying MSC.

3.4.2. Splitting the datasets

Three methods were used to split the data in this study. The first method consisted of randomly picking out a hold-out test data before beginning the analysis. The last two methods were the PY files *fcn_SGS_split* and *fcn_SGKF_split* found on GitHub repositories (scikit-learn/sklearn/model_selection/_split.py, u.d.) and (scikit-learn Stratified GroupKFold #13621, u.d.), respectively. These files contained the cross-validation objects *StratifiedGroupShuffleSplit* and *StratifiedGroupKFold, respectively*.

StratifiedGroupKFold returned folds which had the same percentage of samples for each class (origin of UOC). Furthermore, the method ensured that samples originating from a particular group (sample holder) did not appear in the training and test fold at the same time.

StratifiedGroupShuffleSplit returned folds stratified by the groupings of the data. The percentage of groups in each class were equal.

3.4.3. Initial screening of classifiers on the black-dark brown coloured UOC samples

Several classifiers underwent a screening to find the best one. The screening was done by performing nested cross-validations (Raschka & Mirjalili, Python Machine Learning, 2019). This was done on datasets consisting of different feature groups.

The investigated classifiers from *scikit-learn* were *LogisticRegression* (LR), *SVC* (SVM), *GaussianNB* (NB), *LinearDiscriminantAnalysis* (LDA), *RandomForestClassifier* (RF), *AdaBoost*, and *KNeighborsClassifier* (KNC).

These classifiers (except NB and LDA) has tuneable parameters that have to be manually specified a priori use as they are not automatically optimised from the data (Raschka & Mirjalili, Python Machine Learning, 2019). These so-called hyperparameters can improve model performance if chosen correctly.

The investigated feature groups were:

- *AMT*: containing AMT features
- *FOS*: containing FOS features
- *GLCM 1*: containing GLCM features with distance = 1 (see 3.3.2.4 for info. about distance)
- *GLCM 3*: containing GLCM features with distance = 3
- *GLCM 10*: containing GLCM features with distance = 10
- *GLCM 15*: containing GLCM features with distance = 15
- *GLCM 20*: containing GLCM features with distance = 20
- *GLCM all*: containing all features from the other GLCM groups
- *GLRLM*: containing GLRLM features
- *GLSZM*: containing GLSZM features
- *LBP*: containing LBP features (see 3.3.2.3 for parameters)
- *all*: containing all of the features from all of the groups

Nested cross-validation (n-CV) is a method for finding the best classifier among several classifiers (Raschka & Mirjalili, Python Machine Learning, 2019). It works by splitting the original dataset into one training and one test set (fold) *m* times, and each of these training folds are split to a training and validation fold *n* times. Thus, one has *m* outer folds and *n* inner folds. Figure 24 illustrates the n-CV folds used in this study, three outer and two inner folds were used. The idea is that inner loop of folds finds the best hyperparameter set for a classifier trained on the training fold and validated on the validation fold. The best hyperparameter set is then used for training the classifier in the outer training fold and then tested on the test fold. This is repeated for each of the outer folds. The three test folds together cover all the samples, which means that after going through the outer loop once, all samples have been used as a test sample once. From n-CV a confusion matrix was obtained for the dataset.

n-CV was performed for each combination of classifiers and feature groups, giving 33 combinations. This resulted in 33 confusion matrices which were averaged. For each time a n-CV was performed, a different random state initialization was given to the shuffling functionality in *StratifiedGroupKFold* which created the outer and inner folds. The shuffling was only applied when the outer folds were created. This was done to prevent results to be dependent on the specific split of the dataset (Raschka & Mirjalili, Python Machine Learning, 2019).
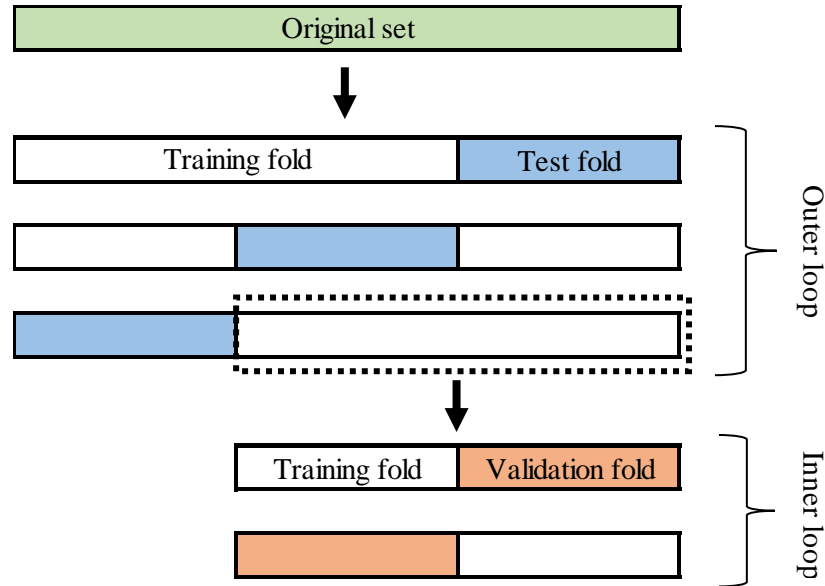
Figure 24: Illustration of the folds in nested cross-validation (n-CV). This is a 3x2 n-CV. The inner loop finds the best classifier hyperparameters, and the outer loop train and validate a model using the found optimal hyperparameter set.

The average confusion matrix for each combination of classifiers and feature groups was used for calculating the performance accuracy and standard deviation.

The average accuracy for the classes, and the standard deviation of the class accuracies, for each confusion matrix were calculated and used to create a heatmap of performance on the dataset for each combination of classifier and feature group.

All classifier confusion matrices were averaged to create a performance heatmap for all feature group and class in that dataset. Likewise, all feature group confusion matrices were averaged to create a performance heatmap for all classifiers and classes in that dataset.

The two inner folds returned the set of best-performing hyperparameters to each of the three outer folds. Consequently, thee different sets of hyperparameters could have been used. This could also vary for each time the n-CV was performed with different samples in each fold. If many different sets of hyperparameters had been used (and no specific set dominated), then that could suggest that the classifier was not stable as the chosen sets were dependent on the dataset splits. All the used hyperparameter sets were tracked to calculate how frequently they were selected.

The total number of times a hyperparameter set is used is the number of outer folds times the number of times the n-CV was performed. The n-CV was run 33 times and thus giving 3 x 33 = 99 times a hyperparameter set was used. Furthermore, the time taken to finish the outer folds in each combination of classifier and feature group were tracked and averaged.

3.4.4. Feature optimization for each dataset and final estimate of model performance

As outlined in Figure 17, the dataset analysed was split three times into a training set and hold-back test set. All samples in the three test sets were unique. The test set consisted of one single sample from each class. For each of these three training sets, an optimised feature set was determined, resulting in three feature sets.

The features sets were determined by first estimate the importance of each feature. This were done by using the function *Feature Importance Permutation* (Feature Importance Permutation, u.d.) from the Python library *Mlxtend*, with the parameter *num_rounds* = 100. The process was repeated 100 times with different random splits for training and validation tests using *StratifiedGroupShuffleSplit*. All the estimated importance values were averaged for each feature, and then the features were sorted by their importance.

The estimated 30 most important features were the starting point for the next step. Here, the author implemented the feature selection algorithm Sequential Backward Selection (SBS) inspired by *Mlxtend* (Sequential Feature Selector, u.d.). The creation of training and validation sets inside this function was done again by *StratifiedGroupShuffleSplit* 100 times splitting randomly. It was desirable to track the classification for each sample for all repeated times using different training and validation sets. To achieve this a new code was implemented and used instead of the already-implemented function from *Mlxtend*. The new code for tracking allowed misclassified samples to be tracked and check if the misclassifications were random of consistent.

The result of doing the SBS was a validation curve in which the accuracy of performance was represented on the vertical axis, and the number of features used along the horizontal axis. This curve was visually inspected to find the best combination of the number of features, and the trade-

off between validation accuracy in relation to training accuracy (i.e. bias-variance trade-off) which denotes the compromise between an under- and overfitting model (Raschka & Mirjalili, Python Machine Learning, 2019). The chosen number of features were listed in a results table containing the features for that particular performance. Then, the classifier was trained on all samples (except the hold-out test data) using the chosen features, and the performance of the trained final model was estimated on the hold-out test data.
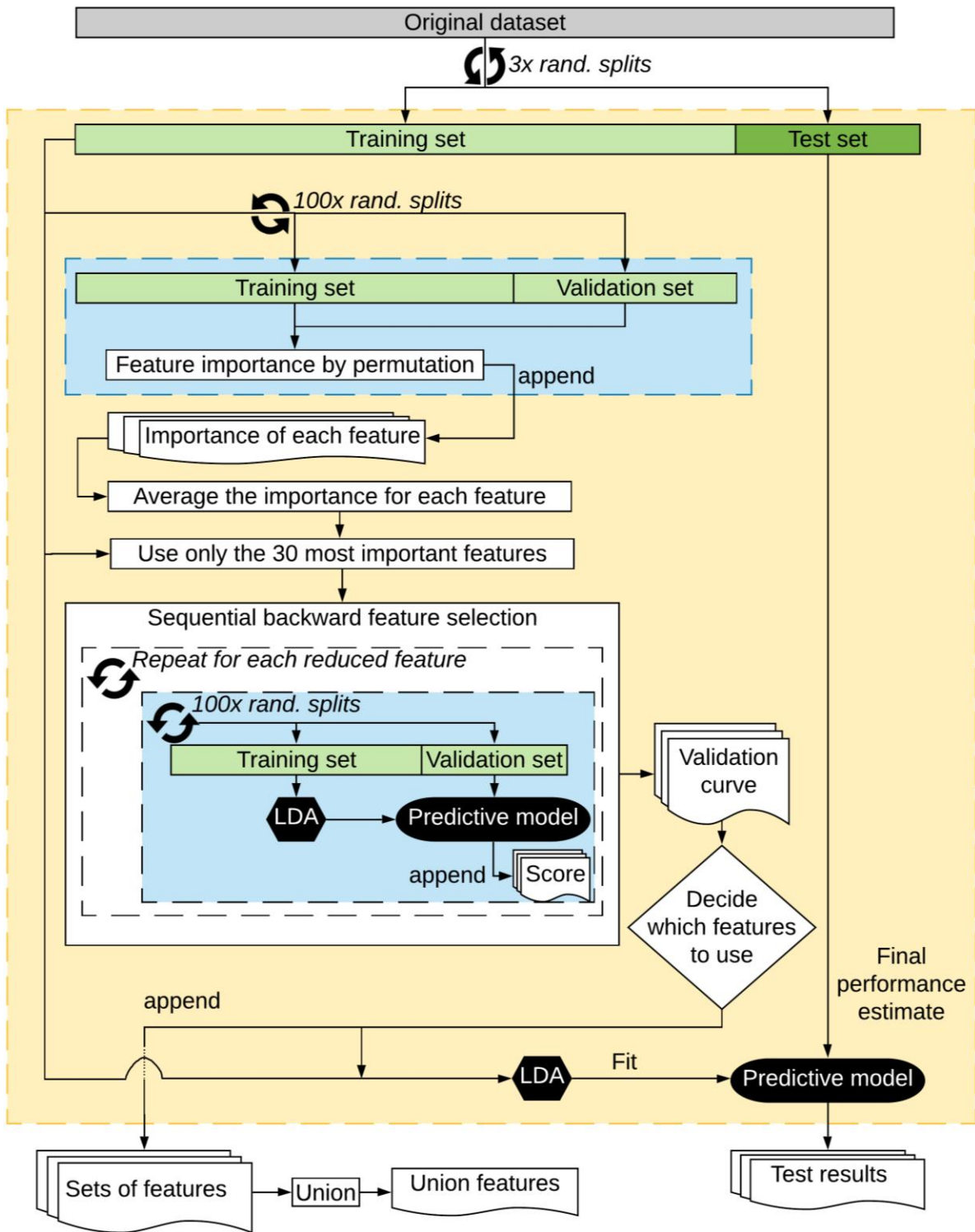
Figure 25: Overview of the process of finding optimized feature sets for a dataset and estimating their performance on hold-out test data.

3.4.5. Examination of initial feature reduction to reduce computational time

An algorithm was developed to reduce the number of features based on their correlation, as seen in example Figure 18. The first step was to calculate the correlation between the features. Then all the correlation values were thresholded by a given number; all cells containing a value greater than the threshold were swapped with value 1, the rest were set to 0. After that, all values in each column (except the diagonal value) were counted. Then, columns with the greatest count were removed along with its mirrored row, meaning that a feature was removed. In step 1 in Figure 18, column and row *A* are highlighted in yellow, meaning that this feature had the most correlations above the threshold. Furthermore, in step 2, this feature is now greyed out to represent that it has been removed. Now, the remaining matrix will be counted as before, and then the next feature will be dropped corresponding to the column with the greatest count. This is repeated until no features had correlations above the threshold. As can be seen in step 7, five features remain in the end, and it is these features that would have been extracted from the original dataset, thus reducing the number of features.

After this algorithm was completed, a more efficient procedure was discovered. A more time-efficient calculation would be to keep the features which had a sum of zero in the columns below the diagonal.
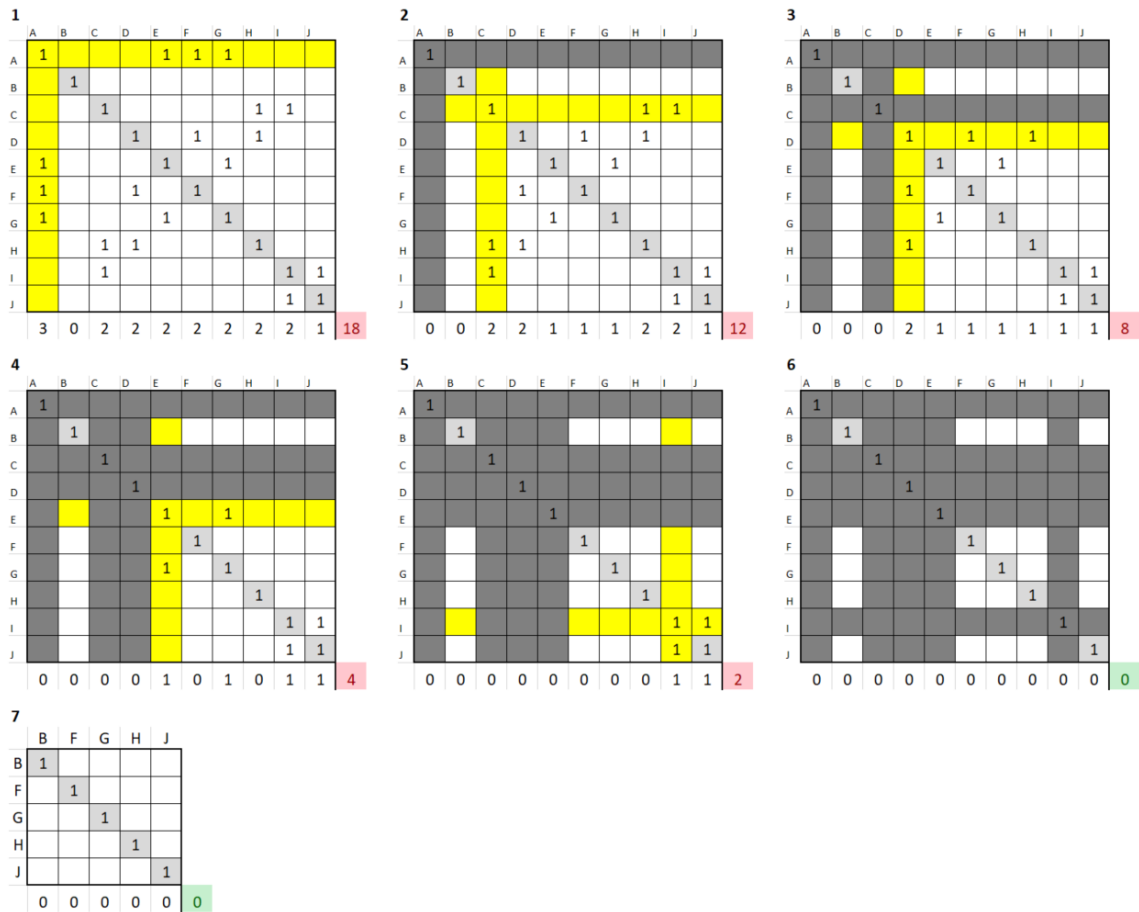
Figure 26: Visual example of how feature removal of the developed correlating features algorithm works, given a correlation matrix which has been thresholded for a chosen value of correlation. The number on the top left of each matrix denotes the steps in the algorithm. The letters along the outside of rows and columns denote specific features. A number of value 1 inside a cell on row X and column Y means that the correlation of the features on X and Y had a Pearson's correlation value greater than the threshold. The numbers below each matrix is a summation of the 1s which are not on the diagonal. The number outside the bottom right corner of the matrices is a summation of the numbers to the left.

3.4.6. Preparing for the prediction of unknown samples

Additional code was created to enable the developed model to classify new unseen UOCs samples. The information needed to predict these unknown samples was their colour category and type of image acquisition (SEM with magnification or HSI). Then the code would retrieve the features that were optimized on the same colour category and type of image acquisition of the known samples from the analysis (the features in this set would be the necessary features to extract from the unknown samples before running this code). Then this feature set would be extracted from the known UOC samples matching the same colour category and type of image acquisition and used

as training data for the classifier to be the best in the analysis. The trained model would then be used to classify the unknown samples by applying this model on the unknown sample data of the same features as the model was trained on. The unknown sample must consist of the same image resolution and have undergone the same pre-processing to be applicable for the predictive model.

4. Results

In this chapter, a selection of results is presented to illustrate the major findings of this study. The remaining results are omitted due to a large number of results, but additional results are referred to and presented in appendix II.

The first step was to determine the best classifier, referring to the top of Figure 17 in chapter 3.4. the most promising classifier was selected from the performance heatmaps and parameter stability. The two datasets pressed and unpressed black- dark brown (CC1) UOC were used. The heatmaps and parameter stabilities were based on nested cross-validation (n-CV) runs on each dataset separately. The second step was to build models, using the best classifier, on the datasets individually. This was done by selecting optimised feature sets.

4.1. Classifier selection – Initial assessment of the classifiers

Classification results are presented in three heatmaps. These maps illustrate (1) the mean accuracy of the classifiers using different feature groups, (2) the mean accuracy for each class using different classifiers, and (3) the average accuracy achieved by the classifiers for each class using different feature groups. Also, the confusion matrix for the classifier giving the highest performance metrics using all features is shown for both datasets.

4.1.1. Pressed black UOC SEM images at 250x magnification

In Figure 27 it can be seen that only the classifiers LR, SVM, and LDA obtained an accuracy over 90 %. The accuracy achieved by these three classifiers was 94 % - 95 % when all feature groups were used together. These classifiers achieved an accuracy of 90 % - 91 % using only LBP features. The classifier AdaBoost achieved the poorest performance on average across the feature groups. Classification based on AMT features resulted in the lowest accuracy scores on average across the

51

different classifiers. The combination given the poorest performance was AdaBoost with AMT features at 31 % accuracy.

Classifiers	FOS	LBP	GLSZM	GLRLM	GLCM 1	GLCM 3	GLCM 10	GLCM 15	GLCM 20	GLCM all	AMT	all
LR	0.70 (0.21) [24.9]	0.91 (0.11) [76.1]	0.81 (0.10) [23.7]	0.77 (0.15) [24.9]	0.71 (0.25) [27.3]	0.70 (0.25) [27.0]	0.70 (0.28) [25.9]	0.70 (0.28) [25.6]	0.70 (0.28) [25.4]	0.83 (0.17) [45.9]	0.56 (0.22) [32.2]	0.94 (0.08) [112.1]
SVM	0.69 (0.21) [14.2]	0.90 (0.11) [22.6]	0.79 (0.13) [12.4]	0.71 (0.18) [12.6]	0.70 (0.24) [12.9]	0.71 (0.23) [12.9]	0.69 (0.25) [12.9]	0.68 (0.26) [13.0]	0.66 (0.27) [13.0]	0.80 (0.17) [16.8]	0.60 (0.22) [13.3]	0.94 (0.10) [33.3]
RF	0.64 (0.24) [681.1]	0.82 (0.15) [765.1]	0.78 (0.15) [680.5]	0.60 (0.22) [679.2]	0.61 (0.25) [680.5]	0.65 (0.25) [680.8]	0.66 (0.23) [681.8]	0.63 (0.24) [684.0]	0.64 (0.25) [683.1]	0.71 (0.24) [730.4]	0.55 (0.23) [690.2]	0.87 (0.14) [801.3]
NB	0.61 (0.24) [2.8]	0.72 (0.16) [3.3]	0.74 (0.19) [2.8]	0.58 (0.24) [2.8]	0.59 (0.24) [2.8]	0.59 (0.27) [2.9]	0.64 (0.25) [2.9]	0.63 (0.25) [2.8]	0.61 (0.27) [2.8]	0.67 (0.24) [3.0]	0.54 (0.25) [2.9]	0.81 (0.18) [3.8]
KNC	0.57 (0.24) [14.0]	0.82 (0.13) [49.7]	0.73 (0.13) [13.0]	0.58 (0.24) [13.0]	0.60 (0.24) [14.3]	0.60 (0.24) [14.3]	0.63 (0.25) [14.4]	0.62 (0.25) [14.3]	0.60 (0.26) [14.4]	0.69 (0.26) [28.6]	0.57 (0.23) [15.7]	0.85 (0.16) [82.4]
LDA	0.71 (0.22) [4.2]	0.91 (0.08) [12.3]	0.86 (0.13) [3.8]	0.77 (0.19) [3.7]	0.77 (0.19) [3.8]	0.78 (0.21) [3.8]	0.73 (0.25) [3.9]	0.72 (0.26) [3.9]	0.72 (0.25) [3.9]	0.87 (0.14) [6.7]	0.58 (0.21) [5.3]	0.95 (0.08) [34.9]
AdaBoost	0.41 (0.11) [26.4]	0.65 (0.10) [73.5]	0.50 (0.12) [26.1]	0.39 (0.11) [26.1]	0.41 (0.11) [27.7]	0.39 (0.11) [27.8]	0.40 (0.13) [27.9]	0.40 (0.13) [27.9]	0.42 (0.11) [27.7]	0.46 (0.10) [45.0]	0.31 (0.07) [29.6]	0.65 (0.08) [110.4]

Feature groups

Figure 27: Heatmap showing the classification performance for different combinations of classifiers (rows) and feature groups (columns). Three values are shown for each classifier and feature group combination. The top value is accuracy, the middle value inside round brackets is standard deviation (SD), and the bottom value inside square brackets is the average time in seconds for going through one outer fold in the n-CV. The accuracies and SDs were calculated from the sample class accuracies for each classifier and feature group combination. The colouring is based on the value of accuracy and given in the colour bar on the right.

Figure 2 shows the classification performance obtained by each classifier for the different sample classes. Here, the accuracy is averaged across the feature groups and the standard deviation (SD) is taken over all the different feature groups. This heatmap gives an indication of how different classifiers manage to classify specific classes on average over the feature groups. The classes AusMak, AusOlD, USAFAP, and USAPet were challenging to classify for all classifiers compared to the other classes. The variation is presented inside the rounded brackets where SD varied from 4 % up to 33 % across the feature groups. On average, LDA gave the highest accuracy for classifying the classes, except for classes SafNuf and YugSpB. They were classified with higher accuracy by the NB classifier, and with NB, SVM, and LR, respectively.

The last heatmap, Figure 29, shows how specific classes were classified given different feature groups, on average across different classifiers. It can be observed that LBP was the only feature group that obtained an accuracy above 50 % for all classes, except for the feature group "all". However, the two feature groups also contained all LBP features. GLSZM also gave reasonably high accuracy, where only one class had an accuracy below 50 % (i.e. 47 %).
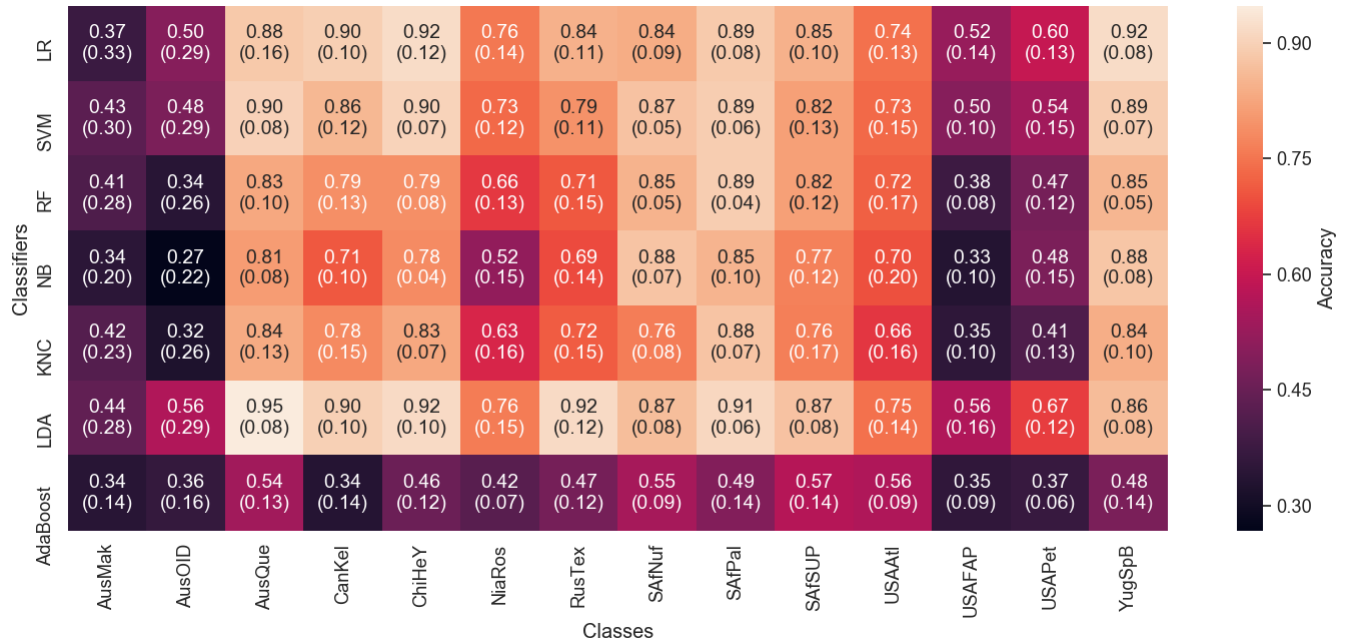
Figure 28: Heatmap showing the classification performance on each of the sample classes (columns) using different classifiers (rows). The top value is accuracy and the bottom value inside round brackets is SD. The accuracy and SD for each classifier and sample class combination were calculated from each class sample accuracies over all feature groups. The colouring is based on the value of accuracy and given in the colour bar on the right.
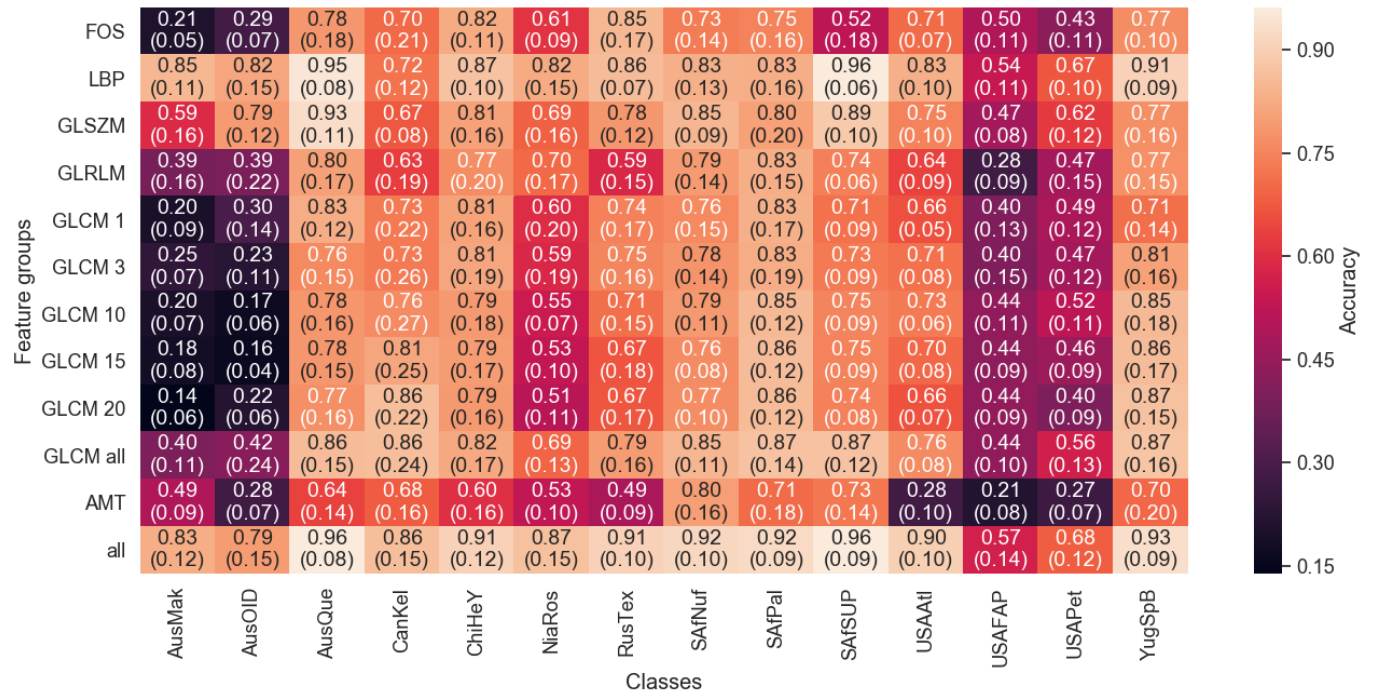
Figure 29: Heatmap showing the classification performance on each of the sample classes (columns) using different feature groups (rows). The top value is accuracy and the bottom value inside round brackets is the SD. The accuracy and SD for each feature group and sample class combination were calculated from each class sample accuracies over all classifiers. The colouring is based on the value of accuracy and given in the colour bar on the right.

Table 10 show how frequently particular classifier hyperparameters were used by LR, SVM and LDA for both LBP features and all feature groups together. Hyperparameter frequency was investigated to check the stability of classifiers. Depending on performance, stable classifiers can be favourable above better-performing classifiers. Neither LR nor SVM had a combination of hyperparameters that occured more than 50 % of the times, indicating that these classifiers were somewhat unstable. For LDA, the solver "lsqr" and shrinkage "auto" were selected. LDA has no other tuneable parameters (sklearn.discriminant_analysis.LinearDiscriminantAnalysis, u.d.).

Table 10: Overview of how frequently different hyperparameter sets were used in the outer folds in the n-CV using LR (left), SVM (middle) and LDA (right). The feature groups used are given in the left-most column, the last column gives the occurrences, and the remaining columns give hyperparameters.

LR

feature group	__C	__solver	occur.
LBP	10	newton-cg	35%
	1	newton-cg	29%
	100	newton-cg	19%
	1000	newton-cg	8%
	0.1	newton-cg	8%
all	1	newton-cg	30%
	10	newton-cg	27%
	100	newton-cg	15%
	0.1	newton-cg	15%
	1000	newton-cg	8%
	0.01	newton-cg	4%

SVM

feature group	__C	__gamma	__kernel	occur.
LBP	1	-	linear	48%
	0.1	-	linear	26%
	100	scale	rbf	18%
	10	scale	rbf	7%
all	0.1	-	linear	46%
	10	scale	rbf	32%
	100	scale	rbf	11%
	0.01	-	linear	10%

LDA

feature group	__shrinkage	__solver	occur.
LBP	auto	lsqr	97%
	-	lsqr	3%
all	auto	lsqr	100%

As seen in Figure 27 and Figure 28 the classifier LDA consistently achieved high classification accuracy. In addition, the setting lsqr solver and auto shrinkage were selected in 97 % of the times when LDA was used, indicating that LDA was a stable and consistent classifier. Therefore, LDA was considered to be the most promising classifier for this dataset.

In Table 11 the confusion matrix obtained for LDA using all features groups together is shown. In general, the majority of samples for most classes were classified correctly. However, on average, in four out of 20 times USAFAP samples were incorrectly predicted as AusQue, perhaps suggesting that these misclassified samples had similarities to the AusQue class. A problem with showing only a confusion matrix is that one cannot see if the same four samples are being misclassified each time, or if the misclassifications happened at random in the class. Consistent misclassifications could be treated as outliers, but this has not been examined in more detail than looking into the tracking of sample classifications. Table 20 on page 66 in chapter 4.2.1.1 give an example to see the predicted class of each sample. This gives insight into if only a few samples are consistently misclassified or not.

Table 11: Confusion matrix obtained for LDA for sample classification based on all feature groups together. Values were averaged over all the n-CV runs using different combinations of samples in training and test folds. The first column gives the true classes, and the top row denotes the predicted classes. There were 20 samples for all classes in this dataset, which means that each row adds up to 20. For example, the top cell in the first column gives the true class AusMak, the remaining cells on the same row tell how many times AusMak samples were predicted as the classes specified in the top row. On average, 19.2 AusMak samples were correctly classified, but 0.1 samples were classified as SAfSUP and 0.8 samples as USAFAP of the total 20 samples.

True class		Predicted class													
		AusMak	AusOID	AusQue	CanKel	ChiHeY	NiaRos	RusTex	SAfNuf	SAfPal	SAfSUP	USAAtl	USAFAP	USAPet	YugSpB
	AusMak	19.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.8	0.0	0.0
	AusOID	0.0	20.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	AusQue	0.0	0.0	20.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	CanKel	0.0	0.0	0.2	16.1	0.0	0.0	0.0	0.0	0.0	0.0	2.7	1.0	0.0	0.0
	ChiHeY	0.0	0.0	0.0	0.0	20.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	NiaRos	0.6	0.0	0.0	0.0	0.0	19.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	RusTex	0.0	0.0	0.0	0.0	0.0	0.0	20.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	SAfNuf	0.0	0.0	0.0	0.0	0.0	0.0	0.0	19.8	0.2	0.0	0.0	0.0	0.0	0.0
	SAfPal	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	20.0	0.0	0.0	0.0	0.0	0.0
	SAfSUP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	20.0	0.0	0.0	0.0	0.0
	USAAtl	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	19.4	0.4	0.0	0.0
	USAFAP	0.0	0.0	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.6	0.3	0.0
	USAPet	2.1	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	1.6	16.0	0.0
	YugSpB	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	20.0

4.1.2. Unpressed black UOC SEM images at 250x magnification

Figure 30 shows the performance of different classifiers used in combination with different feature groups on the unpressed black-dark brown coloured UOC SEM dataset. LDA in combination with each of the two feature groups LBP and all features had the highest average accuracies at respectively 80 % and 75 %. This is 14 % and 9 % higher than the third highest accuracy, achieved by LR and SVM using LBP features. In general LBP features outperformed the other feature groups. As LDA outperformed the other classifiers, the occurrences of different hyperparameter combinations for the other classifiers for this dataset were not investigated.

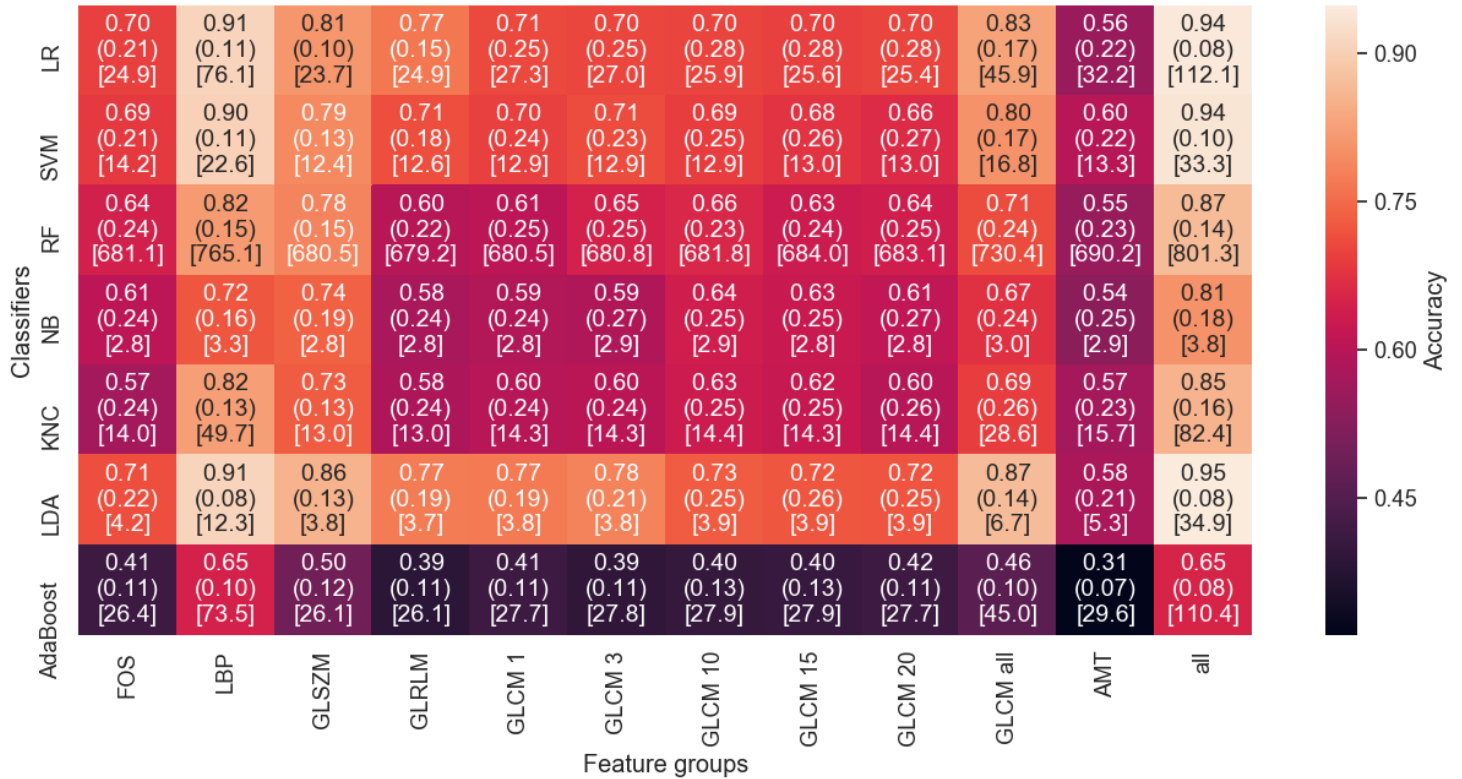| Classifiers	FOS	LBP	GLSZM	GLRLM	GLCM 1	GLCM 3	GLCM 10	GLCM 15	GLCM 20	GLCM all	AMT	all
LR	0.28 (0.32) [19.7]	0.66 (0.24) [54.7]	0.38 (0.29) [19.6]	0.32 (0.33) [19.1]	0.32 (0.33) [20.4]	0.33 (0.32) [20.4]	0.33 (0.31) [20.8]	0.34 (0.32) [20.8]	0.35 (0.33) [20.8]	0.40 (0.31) [34.3]	0.34 (0.30) [23.0]	0.59 (0.27) [76.7]
SVM	0.29 (0.31) [11.4]	0.66 (0.25) [16.8]	0.31 (0.32) [10.4]	0.29 (0.30) [10.5]	0.30 (0.32) [10.5]	0.30 (0.32) [10.4]	0.29 (0.32) [10.4]	0.31 (0.32) [10.4]	0.32 (0.33) [10.4]	0.36 (0.31) [12.7]	0.32 (0.28) [10.8]	0.58 (0.26) [22.9]
RF	0.31 (0.35) [646.9]	0.52 (0.27) [712.8]	0.32 (0.34) [649.1]	0.29 (0.32) [648.8]	0.28 (0.33) [650.5]	0.30 (0.31) [650.1]	0.29 (0.32) [650.3]	0.31 (0.32) [649.2]	0.32 (0.34) [649.7]	0.36 (0.31) [680.6]	0.29 (0.32) [657.5]	0.52 (0.27) [745.8]
NB	0.29 (0.33) [1.9]	0.36 (0.25) [2.2]	0.28 (0.27) [1.8]	0.25 (0.28) [1.8]	0.26 (0.32) [1.8]	0.28 (0.30) [1.8]	0.28 (0.31) [1.8]	0.28 (0.31) [1.8]	0.30 (0.33) [1.8]	0.29 (0.30) [1.9]	0.29 (0.29) [1.8]	0.40 (0.26) [2.5]
KNC	0.29 (0.32) [10.4]	0.50 (0.26) [29.2]	0.27 (0.32) [9.7]	0.26 (0.31) [9.7]	0.28 (0.32) [10.4]	0.29 (0.32) [10.4]	0.28 (0.30) [10.3]	0.31 (0.32) [10.3]	0.30 (0.33) [10.3]	0.33 (0.32) [17.7]	0.29 (0.30) [11.0]	0.45 (0.30) [47.0]
LDA	0.33 (0.31) [2.9]	0.80 (0.19) [10.1]	0.44 (0.30) [2.8]	0.33 (0.33) [2.9]	0.35 (0.30) [3.0]	0.36 (0.30) [2.9]	0.36 (0.31) [2.9]	0.37 (0.31) [3.0]	0.37 (0.31) [3.0]	0.49 (0.28) [4.9]	0.34 (0.30) [3.5]	0.75 (0.22) [29.6]
AdaBoost	0.19 (0.18) [23.5]	0.30 (0.15) [54.2]	0.18 (0.08) [23.1]	0.18 (0.11) [23.0]	0.18 (0.12) [24.0]	0.20 (0.10) [24.0]	0.20 (0.13) [24.0]	0.22 (0.15) [24.0]	0.23 (0.17) [24.0]	0.22 (0.14) [35.6]	0.18 (0.08) [25.5]	0.30 (0.15) [77.7]

Feature groups

Figure 30: Heatmap showing the classification performance for different combinations of classifiers (rows) and feature groups (columns). Three values are shown for each classifier and feature group combination. The top value is accuracy, the middle value inside round brackets is standard deviation (SD), and the bottom value inside square brackets is the average time in seconds for going through one outer fold in the n-CV. The accuracies and SDs were calculated from the sample class accuracies for each classifier and feature group combination. The colouring is based on the value of accuracy and given in the colour bar on the right.

Figure 31 shows the performance for different classifiers on each class, on average across different feature groups. Here, ChHEY and RuTEC were the classes classified with the highest accuracy, whereas especially AuQUE samples were difficult to classify.

Figure 32 shows the average performance of classifiers using different feature groups for the classes. Here, ChHEY and RuTEC were classified with an accuracy above 75 % for every feature group. On average, LBP and "all" features seemed to give the highest accuracy.
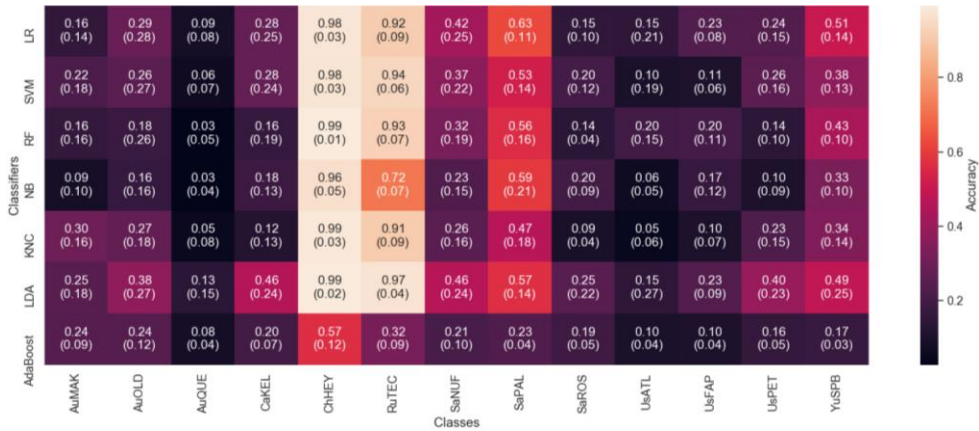
Figure 31: Heatmap showing the classification performance on each of the sample classes (columns) using different classifiers (rows). The top value is accuracy and the bottom value inside round brackets is SD. The accuracy and SD for each classifier and sample class combination were calculated from each class sample accuracy over all feature groups. The colouring is based on the value of accuracy and given in the colour bar on the right.
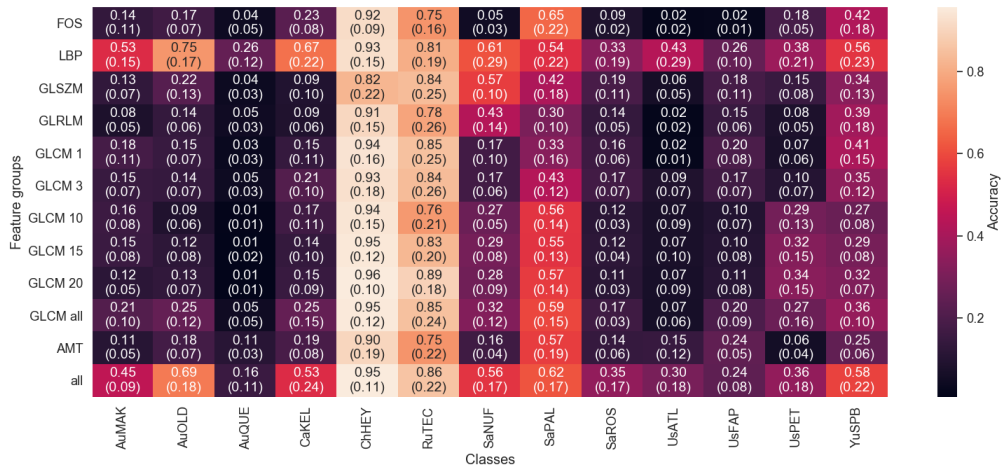
Figure 32: Heatmap showing the classification performance on each of the sample classes (columns) using different feature groups (rows). The top value is accuracy and the bottom value inside round brackets is the SD. The accuracy and SD for each feature group and sample class combination were calculated from each class sample accuracy over all classifiers. The colouring is based on the value of accuracy and given in the colour bar on the right.

As seen in Figure 5, the classifier LDA achieved the highest classification accuracy on average over the classes in this dataset. Table 12 shows a confusion matrix for the classification

performance achieved using LDA on all feature groups together. Most classes were classified correctly, especially the AuOLD, CaKEL, ChHEY, and YuSPB classes. However, UsFAP and AuQUE were often misclassified, with respectively only 4.3 and 5.7 correct predictions out of 15 possible, on average. A sample of UsFAP was almost twice as often classified as SaROS than the correct class UsFAP.

Table 12: Confusion matrix obtained for LDA for sample classification based on all feature groups together. Values were averaged over all the n-CV runs using different composition of samples in training and test folds. The first column gives the true classes, and the top row denotes the predicted classes. There were 15 samples in each class in this dataset, which means that each row adds up to 15. For example, the last cell in the first column says YuSPB, the remaining cells on the same row tell how many times YuSPB samples have been predicted as the classes given by the top row. On average, 14.8 YuSPB samples were correctly classified, but 0.1 samples as AuMAK and 0.2 samples were classified as AuQUE of the total 15 samples.

		Predicted class											
True class		AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
	AuMAK	9.1	0.3	0.1	0.0	0.0	0.0	1.9	0.0	1.5	0.0	1.6	0.6	0.0
	AuOLD	1.0	14.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	AuQUE	2.9	0.0	5.7	0.4	0.0	0.0	0.2	0.0	3.8	0.0	2.0	0.0	0.0
	CaKEL	0.2	0.4	0.0	14.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0
	ChHEY	0.0	0.0	0.0	0.0	15.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	RuTEC	0.0	0.0	0.0	0.0	0.0	15.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	SaNUF	1.9	0.0	0.2	0.0	0.0	0.0	11.9	0.3	0.0	0.0	0.1	0.6	0.0
	SaPAL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.9	0.0	0.0	0.0	3.1	0.0
	SaROS	1.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	10.8	0.0	3.0	0.0	0.0
	UsATL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.5	9.8	2.6	0.0	0.1
	UsFAP	0.4	0.0	0.3	0.0	0.0	0.0	0.0	0.0	8.0	2.0	4.3	0.0	0.0
	UsPET	0.1	0.0	0.0	0.0	0.0	0.0	0.4	3.7	0.2	0.0	0.8	9.8	0.0
	YuSPB	0.1	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	14.8

4.1.3. Screening of pressed versus unpressed UOCs samples

LDA was selected as the best classifier on the black-dark brown coloured samples based on the two screenings. This classifier provided the highest overall classification accuracies and was consistent with regards to classifier parameter settings. In addition, as LDA has no hyperparameters that have to be optimised, it requires a shorter time to develop models. Black-dark brown coloured samples are considered the most challenging samples to classify and were therefore used as a quality check. Furthermore, pressed samples appeared to be the best method to

prepare samples for imaging given the classification accuracy presented in the two previous sub-chapters. Throughout the rest of the chapter, results based on unpressed UOCs only are presented.

4.2. Model building for SEM images for unpressed UOC samples

Based on the previous sections, LDA was chosen as the most promising classifier due to its high performance and consistent selection of settings as well as its speed. The next step was to find a feature set for each dataset that provided high accuracy with the smallest number of features. This section will examine results obtained for the unpressed UOC dataset, colour category 1 (CC1) at 100x magnification and the concatenation of datasets for all magnifications for this colour category, namely 100x, 250x and 1000x. The remaining results can be found in appendix II. An overview of the overall results for all colour categories will also be given in section 4.2.2.

A final model was developed for each of the available datasets, as outlined in Figure 17 (and the detailed Figure 25). Each CC got four final models, one for each magnification and one for all magnifications combined (the concatenated datasets of all magnification available for each CC). The final models consisted of the LDA classifier and a unique, optimised feature set.

4.2.1. Colour category 1

4.2.1.1. 100x magnification

As feature selection was preformed thrice, each time with a different hold-out test set, three validation curves were obtained. These are depicted in Table 13, along with the chosen selected feature sets and the union set of these. As the validation curves follow the SBS algorithm, the graph should be read from right to left, as one after one feature is removed. In Table 13, six features were chosen for each of the three graphs, all of which were LBP features.

For all three graphs, it can be observed that both training and validation accuracy were stable from 30 down to about 7 features, before a sharp decrease to 1 feature. Also, while discarding the first handful of features, the accuracy increased slightly. In the second test, there is a small plateau in validation accuracy at 4 and 3 features, and in the third test at 5, 4, and 3 features. The difference between the training and validation accuracies, given by the red curve, had a dip at three features.

Table 14 shows the distribution of the features selected by SBS sorted into feature groups for each of the three runs. FOS, LBP, GLCM 1, and GLCM 3 were the only feature groups occurring in all three tests. LBP had the most features included in all three runs with different test set. GLSZM, GLRLM, GLCM 20, and AMT were never selected in any run.

Table 15, Table 16 and Table 17 show the test results using the LBP features listed in the far right in Table 13. For example, looking at the 4th row in Table 15, the image C1AuQUEU3O8100x.tif, was assigned to the correct class AuQUE with 99.75% probability and incorrectly assigned to RuTEC and YuSPB with 0.09% and 0.15% probability, respectively. The class assigned with the highest probability will be chosen as the predicted class. For Table 15, this means that the images belonging to the classes CaKEL, SaNUF, SaPAL, SaROS, and UsPET were incorrectly classified. Only 8 out of 13 samples were correctly classified, resulting in an accuracy of 62 %. The three tables show that for all tests the images of SaNUF and SaPAL samples were misclassified at all times.

Table 18 reports the average assignment of probabilities of each sample for the three previous tables. If predictions were made on these averaged probabilities, 11 out of 13 samples would have been correctly classified – an accuracy of 85 %. Table 19 reports the SD of the assigned probabilities in Table 15, 9 and 10, and gives an indication of how much the probabilities in Table 18 varies.

Table 13: Validation curves obtained during feature selection on the unpressed UOC of the CC1 (at 100x) dataset. Each curve was obtained with a different hold-out test set. The first column denotes which run of the feature selection the curve on the corresponding row belongs to, the second column contains validation graphs for the respective runs, and the third column contains the selected features for these runs. The bottom row gives the union of features of the third column. The horizontal and vertical axes of the validation graphs are respectively the number of features and accuracy. The graphs should be read from right to left since there was a reduction of features, the number of features was reduced from 30 to 1. There are three curves in each of the three validation graphs. The blue curve is the training accuracy, green is validation accuracy, and the red is their difference. The light blue and green shaded bands are their corresponding SD.

Which run	Validation curves	Features selected
First		6 features in total: L4_1_4 L44_11_1 L4_1_5 L16_4_1 L24_6_25 L12_3_5
Second		6 features in total: L32_8_1 L4_1_4 L4_1_5 L12_3_5 L16_4_1 L40_10_41
Third		6 features in total: L28_7_1 L4_1_4 L4_1_5 L4_1_2 L16_4_1 L12_3_5
Union features	L44_11_1, L4_1_4, L4_1_2, L24_6_25, L28_7_1, L40_10_41, L32_8_1, L16_4_1, L12_3_5, L4_1_5	

Table 14: Overview of the distribution of the 30 most important features among feature groups for each run on the CC1 (at 100x) dataset. For example, in the first feature selection run, 26 of the 30 most important features belonged to the LBP feature group.

	How many of the 30 most important features belongs to each feature group									
Which test	FOS	LBP	GLSZM	GLRLM	GLCM 1	GLCM 3	GLCM 10	GLCM 15	GLCM 20	AMT
First	1	26	0	0	1	2	0	0	0	0
Second	4	22	0	0	2	2	0	0	0	0
Third	1	21	0	0	1	4	2	1	0	0

Table 15: Prediction matrix of assigned probabilities for the first hold-out test set of the CC1 (at 100x) dataset. The first column denotes the names of the original images that have been classified, and the second column denotes the true class of these samples. The rest of the columns on the first row denote the predicted classes.

First test files	Pred. / True	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
C1AuMAKU3O8100x.tif	AuMAK	62.68%	0.03%	0.01%	0.00%	0.00%	0.00%	35.05%	0.21%	0.33%	0.00%	0.00%	1.51%	0.18%
A3AuOLDU3O8100x.tif	AuOLD	0.34%	99.40%	0.00%	0.07%	0.00%	0.00%	0.00%	0.00%	0.13%	0.01%	0.00%	0.05%	0.00%
C1AuQUEU3O8100x.tif	AuQUE	0.00%	0.00%	99.75%	0.00%	0.00%	0.09%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.15%
A2CaKELU3O8100x.tif	CaKEL	0.01%	0.04%	51.05%	45.46%	0.00%	3.08%	0.16%	0.00%	0.02%	0.00%	0.03%	0.00%	0.17%
A5ChHEYUO2_100x.tif	ChHEY	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A2RuTECMix_100x.tif	RuTEC	0.00%	0.00%	0.02%	0.00%	0.00%	99.88%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%
C1SaNUFU3O8100x.tif	SaNUF	45.46%	0.00%	0.00%	0.00%	0.00%	0.00%	7.40%	0.08%	43.11%	0.07%	0.20%	3.24%	0.44%
C4SaPALU3O8100x.tif	SaPAL	2.25%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	38.33%	0.04%	0.00%	0.00%	59.38%	0.00%
C3SaROSU3O8100x.tif	SaROS	0.19%	0.00%	0.00%	0.01%	0.00%	0.00%	0.01%	0.00%	40.94%	1.11%	57.69%	0.00%	0.03%
A3UsATLU3O8100x.tif	UsATL	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.95%	0.04%	0.00%	0.00%
C2UsFAPU3O8100x.tif	UsFAP	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.68%	0.01%	99.30%	0.00%	0.00%
A1UsPETMix_100x.tif	UsPET	0.77%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	37.78%	60.06%	1.13%	0.27%	0.00%
A4YuSPBUH__100x.tif	YuSPB	0.40%	0.00%	0.06%	0.00%	0.00%	0.01%	2.16%	0.00%	0.64%	0.00%	0.03%	0.00%	96.70%

Table 16: Prediction matrix of assigned probabilities for the second hold-out test set of the CC1 (at 100x) dataset. The first column denotes the names of the original images that have been classified, and the second column denotes the true class of these samples. The rest of the columns on the first row denote the predicted classes.

Second test files	Pred. / True	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
A4AuMAKU3O8100x.tif	AuMAK	60.55%	0.00%	0.00%	0.00%	0.00%	0.00%	20.60%	0.40%	7.80%	0.00%	0.01%	10.63%	0.01%
B5AuOLDU3O8100x.tif	AuOLD	0.03%	99.97%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A3AuQUEU3O8100x.tif	AuQUE	0.18%	0.00%	24.55%	0.26%	0.00%	0.00%	0.09%	0.00%	8.49%	0.00%	66.43%	0.00%	0.01%
B4CaKELU3O8100x.tif	CaKEL	0.00%	0.01%	0.01%	99.98%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B3ChHEYUO2_100x.tif	ChHEY	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C5RuTECMix_100x.tif	RuTEC	0.00%	0.00%	0.21%	0.00%	0.00%	99.79%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B2SaNUFU3O8100x.tif	SaNUF	0.00%	0.00%	0.06%	0.00%	98.70%	1.24%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A5SaPALU3O8100x.tif	SaPAL	10.81%	0.00%	0.00%	0.00%	0.00%	0.00%	2.32%	3.32%	3.21%	0.02%	0.00%	80.33%	0.00%
B3SaROSU3O8100x.tif	SaROS	2.25%	0.00%	0.00%	0.00%	0.00%	0.00%	0.08%	0.00%	82.22%	5.28%	10.07%	0.08%	0.00%
B4UsATLU3O8100x.tif	UsATL	0.08%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	10.92%	88.16%	0.69%	0.14%	0.00%
B5UsFAPU3O8100x.tif	UsFAP	2.88%	0.00%	0.00%	0.00%	0.00%	0.00%	0.25%	0.00%	70.36%	0.19%	26.32%	0.01%	0.00%
C3UsPETMix_100x.tif	UsPET	1.79%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	2.17%	0.03%	0.00%	0.00%	95.99%	0.00%
B4YuSPBUH__100x.tif	YuSPB	10.21%	0.10%	0.05%	0.00%	0.00%	0.00%	10.69%	0.00%	4.00%	0.00%	0.07%	0.01%	74.86%

Table 17: Prediction matrix of assigned probabilities for the third hold-out test set of the CC1 (at 100x) dataset. The first column denotes the names of the original images that have been classified, and the second column denotes the true class of these samples. The rest of the columns on the first row denote the predicted classes.

Third test files	Pred. / True	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
B5AuMAKU3O8100x.tif	AuMAK	51.82%	0.03%	0.01%	0.00%	0.00%	0.00%	2.57%	0.00%	41.37%	0.19%	1.28%	2.42%	0.30%
C4AuOLDU3O8100x.tif	AuOLD	0.16%	91.37%	0.02%	0.16%	0.00%	0.00%	0.00%	0.00%	5.11%	0.11%	3.06%	0.00%	0.01%
B1AuQUEU3O8100x.tif	AuQUE	0.00%	0.00%	0.23%	0.06%	0.00%	99.71%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C5CaKELU3O8100x.tif	CaKEL	0.00%	0.01%	0.00%	99.99%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C4ChHEYUO2_100x.tif	ChHEY	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B5RuTECMix_100x.tif	RuTEC	0.00%	0.93%	0.35%	1.30%	0.00%	97.42%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A2SaNUFU3O8100x.tif	SaNUF	17.05%	77.10%	1.49%	0.00%	0.00%	0.20%	3.90%	0.00%	0.05%	0.00%	0.00%	0.04%	0.17%
B3SaPALU3O8100x.tif	SaPAL	2.17%	0.00%	0.00%	0.00%	0.00%	0.00%	0.15%	39.09%	0.01%	0.00%	0.00%	58.59%	0.00%
A4SaROSU3O8100x.tif	SaROS	40.11%	0.00%	0.07%	0.00%	0.00%	0.00%	20.75%	0.00%	13.96%	0.01%	0.24%	0.97%	23.89%
C2UsATLU3O8100x.tif	UsATL	3.54%	0.01%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	10.66%	73.26%	0.41%	12.12%	0.00%
A3UsFAPU3O8100x.tif	UsFAP	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.71%	5.50%	92.78%	0.00%	0.00%
B2UsPETMix_100x.tif	UsPET	3.25%	0.00%	0.00%	0.00%	0.00%	0.00%	0.26%	18.48%	0.01%	0.00%	0.00%	78.00%	0.00%
C4YuSPBUH__100x.tif	YuSPB	0.01%	0.00%	0.43%	0.00%	0.00%	0.01%	0.22%	0.00%	0.00%	0.00%	0.00%	0.00%	99.33%

Table 18: The average assigned probabilities for the three hold-out test sets of the CC1 (at 100x) dataset, for each combination of true and predicted class.

Pred. / True	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
AuMAK	58.35%	0.02%	0.01%	0.00%	0.00%	0.00%	19.40%	0.20%	16.50%	0.07%	0.43%	4.85%	0.16%
AuOLD	0.17%	96.91%	0.01%	0.08%	0.00%	0.00%	0.00%	0.00%	1.75%	0.04%	1.02%	0.02%	0.00%
AuQUE	0.06%	0.00%	41.51%	0.10%	0.00%	33.27%	0.03%	0.00%	2.83%	0.00%	22.14%	0.00%	0.05%
CaKEL	0.00%	0.02%	17.02%	81.81%	0.00%	1.03%	0.05%	0.00%	0.01%	0.00%	0.01%	0.00%	0.06%
ChHEY	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RuTEC	0.00%	0.31%	0.19%	0.43%	0.00%	99.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.03%
SaNUF	20.84%	25.70%	0.52%	0.00%	32.90%	0.48%	3.77%	0.03%	14.39%	0.02%	0.07%	1.09%	0.20%
SaPAL	5.08%	0.00%	0.00%	0.00%	0.00%	0.00%	0.82%	26.91%	1.08%	0.01%	0.00%	66.10%	0.00%
SaROS	14.19%	0.00%	0.03%	0.00%	0.00%	0.00%	6.95%	0.00%	45.71%	2.13%	22.67%	0.35%	7.97%
UsATL	1.21%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	7.19%	87.12%	0.38%	4.09%	0.00%
UsFAP	0.96%	0.00%	0.00%	0.00%	0.00%	0.00%	0.08%	0.00%	24.25%	1.90%	72.80%	0.00%	0.00%
UsPET	1.94%	0.00%	0.00%	0.00%	0.00%	0.00%	0.09%	6.89%	12.60%	20.02%	0.38%	58.08%	0.00%
YuSPB	3.54%	0.03%	0.18%	0.00%	0.00%	0.01%	4.36%	0.00%	1.55%	0.00%	0.03%	0.00%	90.30%

Table 19: SD of the assigned probabilities across the three hold-out test sets of the CC1 (at 100x) dataset, for each combination of true and predicted class.

Pred. / True	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
AuMAK	4.70%	0.01%	0.00%	0.00%	0.00%	0.00%	13.29%	0.16%	17.85%	0.09%	0.60%	4.10%	0.12%
AuOLD	0.13%	3.93%	0.01%	0.06%	0.00%	0.00%	0.00%	0.00%	2.38%	0.05%	1.44%	0.02%	0.00%
AuQUE	0.08%	0.00%	42.36%	0.11%	0.00%	46.98%	0.04%	0.00%	4.00%	0.00%	31.31%	0.00%	0.07%
CaKEL	0.00%	0.01%	24.06%	25.70%	0.00%	1.45%	0.07%	0.00%	0.01%	0.00%	0.01%	0.00%	0.08%
ChHEY	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RuTEC	0.00%	0.44%	0.14%	0.61%	0.00%	1.14%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.05%
SaNUF	18.75%	36.34%	0.69%	0.00%	46.53%	0.54%	3.02%	0.04%	20.31%	0.04%	0.09%	1.52%	0.18%
SaPAL	4.06%	0.00%	0.00%	0.00%	0.00%	0.00%	1.06%	16.69%	1.50%	0.01%	0.00%	10.07%	0.00%
SaROS	18.35%	0.00%	0.03%	0.00%	0.00%	0.00%	9.76%	0.00%	28.07%	2.27%	25.09%	0.44%	11.25%
UsATL	1.65%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	5.09%	10.92%	0.27%	5.68%	0.00%
UsFAP	1.36%	0.00%	0.00%	0.00%	0.00%	0.00%	0.12%	0.00%	32.60%	2.55%	32.98%	0.00%	0.00%
UsPET	1.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.12%	8.25%	17.80%	28.31%	0.53%	41.54%	0.00%
YuSPB	4.72%	0.05%	0.18%	0.00%	0.00%	0.01%	4.55%	0.00%	1.75%	0.00%	0.03%	0.00%	10.97%

Figure 7 displays the original images for the test samples used for the class UsPET (top row), and one randomly chosen image of the classes UsATL and SaROS (bottom row). This was done, as an example, to illustrate the intra- and inter-class differences between images. Both C3UsPETMix_100x.tif and B2UsPETMIX_100x.tif were correctly classified with respectively assigned probabilities of 96 % (Table 16) and 78 % (Table 17), whereas A1UsPETMix_100x.tif weas only assigned to its true class with 0.3 % probability (Table 15). Instead, this misclassified sample was assigned with 60 % and 38 % probability of belonging to respectively classes UsATL and SaROS. The two correctly classified UsPET images were both assigned to UsATL and SaROS with 0 % probability. These images had large structures, in contrast to the fine and coarse texture in the UsATL and SaROS images, respectively. A blend of both fine and coarse texture appears to be present in the misclassified UsPET image, suggesting the reason why it could have been assigned to the wrong classes.
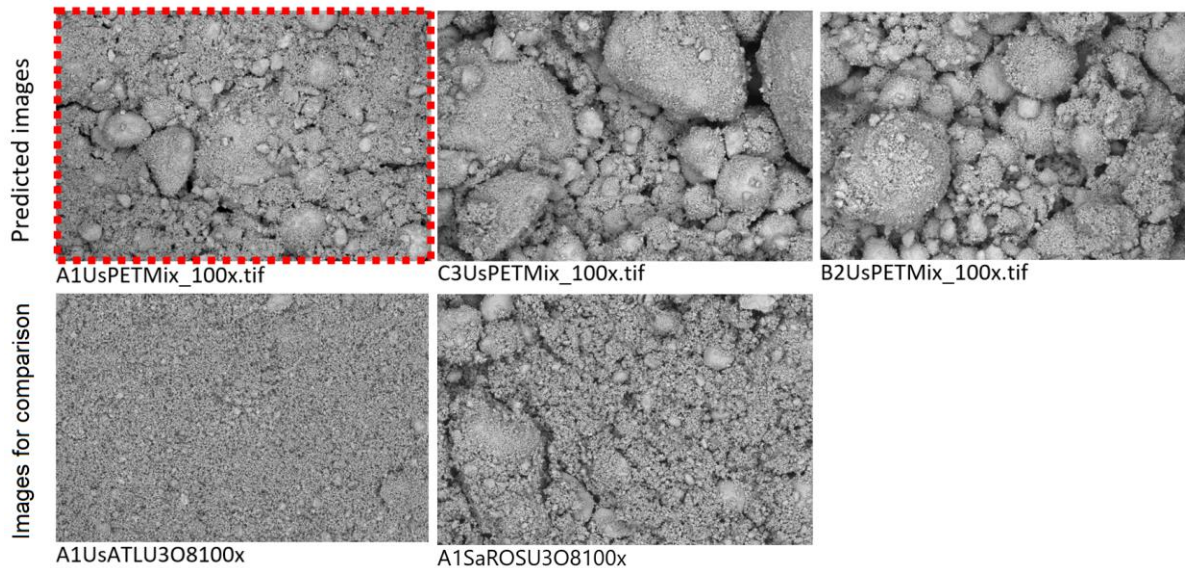
Figure 33: A selection of images, as an example, to illustrate the intra- and inter-class differences between images and its effect on classification. The top row shows the three images of class UsPET that were classified in the three test runs. The two bottom images belonging to the classes UsATL and SaROS are included for comparison. The top left image, which is marked with red dots, was misclassified. It was wrongly attributed to class UsATL with 60 % probability and SaROS with 38 % probability.

Code was implemented to track predictions of each sample throughout the feature reductions reported in Table 13. A record of image misclassifications could provide insight into whether misclassifications within a class happened at random or if only a few samples were consistently misclassified. As an example, Table 20 provides this information for the class SaNUF in the first run using the six LBP features specified in the first row of Table 6. It was observed that five samples accounted for the majority of misclassifications. The nine remaining samples were correctly classified at least 88 % of the time. This table indicates that the five misclassified samples A2SaNUFU3O8100x.tif, B2SaNUFU3O8100x.tif, C3SaNUFU3O8100x.tif, C4SaNUFU3O8100x.tif and C5SaNUFU3O8100x.tif should be examined for further analysis.

Table 20: In-depth information of sample predictions for the class SaNUF in the first feature reduction run using the six LBP features specified in the first row of Table 6. The stated percentage reports how often the respective samples have been classified as which class, as percentage of the total number of times the respective samples have been classified. For example, in the second row, the sample originating from image A1SaNUFU3O8100x.tif was correctly classified as SaNUF in 93 % of the times this specific sample was classified. Also, this sample were incorrectly classified as YuSPB in 7% of the cases.

filenames	True classes	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
A1SaNUFU3O8100x.tif	SaNUF							93%						7%
A2SaNUFU3O8100x.tif	SaNUF	14%	38%	41%				7%						
A3SaNUFU3O8100x.tif	SaNUF							100%						
A4SaNUFU3O8100x.tif	SaNUF							100%						
A5SaNUFU3O8100x.tif	SaNUF							100%						
B1SaNUFU3O8100x.tif	SaNUF							100%						
B2SaNUFU3O8100x.tif	SaNUF					97%	3%							
B3SaNUFU3O8100x.tif	SaNUF							100%						
B4SaNUFU3O8100x.tif	SaNUF							100%						
B5SaNUFU3O8100x.tif	SaNUF							100%						
C2SaNUFU3O8100x.tif	SaNUF							88%						12%
C3SaNUFU3O8100x.tif	SaNUF	100%												
C4SaNUFU3O8100x.tif	SaNUF	78%								22%				
C5SaNUFU3O8100x.tif	SaNUF	95%								5%				

4.2.1.2.   All magnifications

Table 21, 15, and 16 were acquired from the three tests using the three concatenated datasets originating from the magnifications 100x, 250x, and 1000x of the unpressed black-dark brown coloured UOCs. In Table 21, auold, cakel, saros, and uspet samples were misclassified in the first test. The auold sample was wrongly assigned to the class rutec with more than 99 % probability of belonging to this class, whereas for the cakel sample, the assigned probability to the correct class was 3 % lower than wrong assignment as auque with 46 % probability. Table 22 shows that three misclassifications were made in the second test. Also, only one sample was ever assigned to sapal, with only a 6 % probability. Table 23 shows that there was only one misclassification in the third test.

In contrast to the predictive accuracy based on the averaged assigned probabilities for samples at 100x magnification in Table 18, the concatenation of magnifications (Table 24) gave one less misclassified class.

Table 26 shows how many features of the 30 initial features in the validation curve runs, belong to the different feature groups and at which magnification they were extracted. Most of the selected features belonged to the LBP group regardless of magnification. In contrast to

Table 14 for samples at 100x magnification, features from three different magnifications were included in this dataset. Table 27 shows which features occurred more than once among the three magnifications in each run. This table was made to seek insight into if some features managed to extract useful information independently of magnification.

Table 21: Prediction matrix of assigned probabilities for the first hold-out test set on the concatinated dataset (CC1) of features from all magnifications. The first column denotes the names of the original images that have been classified, and the second column denotes the true class of these samples. The rest of the columns on the first row denote the predicted classes.

First test	Pred. / True	aumak	auold	auque	cakel	chhey	rutec	sanuf	sapal	saros	usatl	usfap	uspet	yuspb
c1aumaku3o8100x.ti	aumak	79.18%	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	2.03%	0.01%	0.00%	18.53%	0.20%
a3auoldu3o8100x.ti	auold	0.00%	0.22%	0.00%	0.01%	0.00%	99.76%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
c1auqueu3o8100x.ti	auque	0.12%	3.53%	75.76%	0.27%	0.00%	0.00%	0.00%	0.00%	10.36%	0.01%	1.43%	0.00%	8.52%
a2cakelu3o8100x.ti	cakel	0.00%	6.63%	45.95%	43.04%	0.00%	0.74%	0.00%	0.00%	1.19%	0.00%	1.37%	0.00%	1.08%
a5chheyuo2_100x.ti	chhey	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
a2rutecmix_100x.ti	rutec	0.00%	0.00%	0.00%	0.56%	0.00%	99.42%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%
c1sanufu3o8100x.ti	sanuf	0.04%	0.00%	0.00%	0.00%	0.02%	0.00%	89.32%	0.57%	0.11%	0.00%	0.14%	0.21%	9.58%
c4sapalu3o8100x.ti	sapal	0.06%	0.00%	0.00%	0.00%	0.00%	0.00%	0.03%	62.90%	0.00%	0.00%	0.00%	36.98%	0.02%
c3sarosu3o8100x.ti	saros	0.33%	0.38%	42.10%	0.05%	0.00%	0.00%	0.00%	0.00%	38.09%	4.72%	13.29%	0.00%	1.03%
a3usatlu3o8100x.ti	usatl	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.18%	91.82%	7.99%	0.00%	0.00%
c2usfapu3o8100x.ti	usfap	0.00%	0.00%	1.66%	0.11%	0.00%	0.00%	0.00%	0.00%	0.39%	0.45%	97.39%	0.00%	0.01%
a1uspetmix_100x.ti	uspet	0.49%	0.00%	0.00%	0.00%	0.00%	0.00%	1.01%	3.53%	2.01%	0.33%	77.49%	14.94%	0.20%
a4yuspbuh__100x.ti	yuspb	1.13%	2.27%	3.29%	0.08%	0.00%	0.01%	0.00%	0.00%	10.29%	0.01%	0.89%	0.00%	82.02%

Table 22: Prediction matrix of assigned probabilities for the second hold-out test set on the concatinated dataset (CC1) of features from all magnifications. The first column denotes the names of the original images that have been classified, and the second column denotes the true class of these samples. The rest of the columns on the first row denote the predicted classes.

Second test	Pred. / True	aumak	auold	auque	cakel	chhey	rutec	sanuf	sapal	saros	usatl	usfap	uspet	yuspb
a4aumaku3o8100x.ti	aumak	56.53%	0.02%	0.08%	0.00%	0.00%	0.00%	0.01%	0.00%	21.89%	0.03%	0.04%	0.25%	21.14%
b5auoldu3o8100x.ti	auold	0.00%	3.78%	27.49%	36.81%	0.00%	31.92%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
a3auqueu3o8100x.ti	auque	0.12%	6.79%	40.04%	0.02%	0.00%	0.00%	0.00%	0.00%	20.76%	0.00%	21.78%	0.00%	10.49%
b4cakelu3o8100x.ti	cakel	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
b3chheyuo2_100x.ti	chhey	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
c5rutecmix_100x.ti	rutec	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
b2sanufu3o8100x.ti	sanuf	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
a5sapalu3o8100x.ti	sapal	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	99.97%	0.00%	0.01%	0.00%	0.00%	0.00%	0.01%
b3sarosu3o8100x.ti	saros	2.71%	6.14%	0.23%	0.00%	0.00%	0.00%	0.05%	0.00%	58.99%	0.50%	30.30%	0.00%	1.08%
b4usatlu3o8100x.ti	usatl	3.58%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	12.79%	83.28%	0.11%	0.01%	0.23%
b5usfapu3o8100x.ti	usfap	1.73%	5.10%	5.49%	0.00%	0.00%	0.00%	0.43%	0.00%	53.17%	0.02%	24.63%	0.00%	9.44%
c3uspetmix_100x.ti	uspet	2.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	6.01%	0.01%	0.01%	0.00%	91.63%	0.00%
b4yuspbuh__100x.ti	yuspb	4.36%	0.64%	9.17%	0.00%	0.00%	0.00%	0.00%	0.00%	19.11%	0.00%	0.24%	0.00%	66.48%

Table 23: Prediction matrix of assigned probabilities for the third hold-out test set on the concatinated dataset (CC1) of features from all magnifications. The first column denotes the names of the original images that have been classified, and the second column denotes the true class of these samples. The rest of the columns on the first row denote the predicted classes.

Third test	Pred. / True	aumak	auold	auque	cakel	chhey	rutec	sanuf	sapal	saros	usatl	usfap	uspet	yuspb
b5aumaku3o8100x.ti	aumak	76.70%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	15.59%	0.71%	0.07%	6.31%	0.61%
c4auoldu3o8100x.ti	auold	0.02%	96.56%	0.22%	0.01%	0.00%	0.00%	0.02%	0.00%	0.56%	0.00%	2.61%	0.00%	0.00%
b1auqueu3o8100x.ti	auque	0.00%	30.68%	64.80%	4.52%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
c5cakelu3o8100x.ti	cakel	0.00%	0.28%	0.00%	99.72%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
c4chheyuo2_100x.ti	chhey	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
b5rutecmix_100x.ti	rutec	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
a2sanufu3o8100x.ti	sanuf	0.00%	0.03%	0.00%	0.00%	0.00%	0.00%	99.97%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
b3sapalu3o8100x.ti	sapal	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	96.13%	0.00%	0.00%	0.00%	3.87%	0.00%
a4sarosu3o8100x.ti	saros	6.13%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.19%	0.09%	0.00%	0.01%	93.58%
c2usatlu3o8100x.ti	usatl	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	99.95%	0.00%	0.01%	0.01%
a3usfapu3o8100x.ti	usfap	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.54%	1.25%	98.20%	0.00%	0.00%
b2uspetmix_100x.ti	uspet	4.41%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.11%	0.00%	0.00%	0.00%	94.48%	0.00%
c4yuspbuh__100x.ti	yuspb	0.08%	0.00%	1.08%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	98.84%

Table 24: The average assigned probabilities for the three hold-out test sets on the concatenated CC1 dataset, for each combination of true and predicted class.

True \ Pred.	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
AuMAK	70.81%	0.02%	0.03%	0.00%	0.00%	0.00%	0.00%	0.01%	13.17%	0.25%	0.04%	8.36%	7.32%
AuOLD	0.01%	33.52%	9.24%	12.28%	0.00%	43.89%	0.01%	0.00%	0.19%	0.00%	0.87%	0.00%	0.00%
AuQUE	0.08%	13.67%	60.20%	1.60%	0.00%	0.00%	0.00%	0.00%	10.37%	0.00%	7.74%	0.00%	6.34%
CaKEL	0.00%	2.31%	15.32%	80.92%	0.00%	0.25%	0.00%	0.00%	0.40%	0.00%	0.46%	0.00%	0.36%
ChHEY	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RuTEC	0.00%	0.00%	0.00%	0.19%	0.00%	99.81%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%
SaNUF	0.01%	0.01%	0.00%	0.00%	0.01%	0.00%	96.43%	0.19%	0.04%	0.00%	0.05%	0.07%	3.19%
SaPAL	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	33.33%	53.01%	0.00%	0.00%	0.00%	13.62%	0.01%
SaROS	3.06%	2.17%	14.11%	0.02%	0.00%	0.00%	0.02%	0.00%	32.42%	1.77%	14.53%	0.01%	31.89%
UsATL	1.20%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	4.33%	91.68%	2.70%	0.01%	0.08%
UsFAP	0.58%	1.70%	2.38%	0.04%	0.00%	0.00%	0.14%	0.00%	18.03%	0.57%	73.41%	0.00%	3.15%
UsPET	2.41%	0.00%	0.00%	0.00%	0.00%	0.00%	0.34%	3.55%	0.68%	0.11%	25.83%	67.01%	0.07%
YuSPB	1.86%	0.97%	4.51%	0.03%	0.00%	0.00%	0.00%	0.00%	9.80%	0.01%	0.38%	0.00%	82.45%

Table 25: SD of the assigned probabilities across the three hold-out test sets on the concatenated CC1 dataset, for each combination of true and predicted class.

True \ Pred.	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
AuMAK	10.14%	0.01%	0.04%	0.00%	0.00%	0.00%	0.01%	0.01%	8.29%	0.33%	0.03%	7.60%	9.78%
AuOLD	0.01%	44.60%	12.91%	17.35%	0.00%	41.60%	0.01%	0.00%	0.26%	0.00%	1.23%	0.00%	0.00%
AuQUE	0.06%	12.10%	14.94%	2.06%	0.00%	0.00%	0.00%	0.00%	8.47%	0.00%	9.95%	0.00%	4.55%
CaKEL	0.00%	3.06%	21.66%	26.79%	0.00%	0.35%	0.00%	0.00%	0.56%	0.00%	0.65%	0.00%	0.51%
ChHEY	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RuTEC	0.00%	0.00%	0.00%	0.26%	0.00%	0.27%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%
SaNUF	0.02%	0.02%	0.00%	0.00%	0.01%	0.00%	5.03%	0.27%	0.05%	0.00%	0.07%	0.10%	4.52%
SaPAL	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%	47.12%	39.86%	0.00%	0.00%	0.00%	16.59%	0.01%
SaROS	2.38%	2.81%	19.79%	0.03%	0.00%	0.00%	0.02%	0.00%	24.33%	2.09%	12.40%	0.01%	43.62%
UsATL	1.68%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	5.99%	6.80%	3.74%	0.00%	0.11%
UsFAP	0.81%	2.40%	2.30%	0.05%	0.00%	0.00%	0.20%	0.00%	24.84%	0.51%	34.49%	0.00%	4.45%
UsPET	1.60%	0.00%	0.00%	0.00%	0.00%	0.00%	0.48%	2.00%	0.94%	0.15%	36.53%	36.84%	0.10%
YuSPB	1.82%	0.96%	3.42%	0.04%	0.00%	0.00%	0.00%	0.00%	7.81%	0.01%	0.37%	0.00%	13.22%

Table 26: Overview of the distribution of the most important features among the feature groups and magnifications for each run. For example, in the third feature selection run, of the 30 most important features the only feature groups selected were FOS and LBP. Only 2 were FOS features, the 28 remaining were LBP features. The 2 FOS features were extracted at 100x magnification. Among the LBP features, 14 were also extracted at 100x magnification, and the remaining 9 and 5 were extracted from the images acquired at respectively 250x and 1000x magnifications.

Which test	Magnification	FOS	LBP	GLSZM	GLRLM	GLCM 1	GLCM 3	GLCM 10	GLCM 15	GLCM 20	AMT
First	100x	1	11	0	0	0	0	0	0	0	0
	250x	0	6	0	0	0	0	0	0	0	1
	1000x	0	11	0	0	0	0	0	0	0	0
Second	100x	0	10	0	0	0	0	0	0	0	0
	250x	0	10	0	0	1	1	0	0	0	0
	1000x	0	8	0	0	0	0	0	0	0	0
Third	100x	2	14	0	0	0	0	0	0	0	0
	250x	0	9	0	0	0	0	0	0	0	0
	1000x	0	5	0	0	0	0	0	0	0	0

Table 27: Overview of the features that occurred more than once among the three magnifications for each run on the concatenated CC1 dataset.

Which test	L16_4_3	L36_9_1	L40_10_1	L44_11_1	L12_3_4
First	2	2	2	2	
Second			2	2	
Third			2	2	2

4.2.2. General results for SEM data

Table 28 summarizes the performance on the datasets at all magnifications originating from the unpressed SEM images. The accuracy for predicting all samples in each test are presented along with their mean accuracy. In addition, the averaged assigned class probabilities for the three tests are also shown. To clarify, prediction in Table 18 gives the mean assigned probability accuracy for colour category 1 at 100x magnification only.

The highest achieved mean prediction accuracy averaged over all colour categories was obtained when the datasets for different magnifications were concatenated. But, the achieved accuracy of 90 % was only 2 % higher than the performance achieved at 100x magnification alone. The mean assigned probability accuracy averaged over all colour categories was above 97 % for 100x, 250x, and the concatenation of all magnifications. The difference in performance between using only a magnification of 100x and the concatenation of all was small.

Colour category 1 had both the poorest mean prediction and assigned probability accuracies for 100x, 250x, and all magnifications datasets. Also, this category and along with colour category 3 contained 13 classes, whereas colour category 4, 5, and 6 contained respectively 9, 4, and 4 classes. This must be considered when comparing the overall performance of each of the categories. A trained machine learning model on each of the black-dark brown and light brown-dark yellow samples was trained for classifying 13 classes, whereas the other categories had fewer classes to discriminate.

Furthermore, the accuracies averaged over the colour categories were not weighted according to the number of classes they each contained. As more classes were included in category 1, the

average accuracy for all classes would be smaller. A final note, the average number of features selected for the three runs might indicate the difficulty of classifying the datasets.

Table 28: Summary of performance on SEM datasets for unpressed samples for all colour categories. The first column denotes the magnification, the second gives a description of the summarized performance, the 3rd to 7th columns denote the colour category and the last column shows the average over all colour classes of mean prediction accuracy, mean assignment accuracy, and the average number of features.

Magnification	Description	CC1	CC3	CC4	CC5	CC6	Mean over all colour categories	
100x	First test acc.:	62%	92%	78%	100%	100%		
	Second test acc.:	69%	77%	100%	100%	100%		
	Third test acc.:	69%	100%	100%	75%	100%		
	Mean prediction acc.:	**67%**	**90%**	**93%**	**92%**	**100%**	Mean prediction acc.:	88%
	Mean assigned prob. acc.:	**85%**	**100%**	**100%**	**100%**	**100%**	Mean assigned prob. acc.:	97%
	Mean assigned prob. SD:	29%	14%	11%	14%	0%	Avg. # of feats. in each CC	5.5
	Avg. # of feats.:	**6**	**4.7**	**5**	**6**	**6**		
250x	First test acc.:	85%	69%	100%	100%	100%		
	Second test acc.:	54%	77%	89%	100%	100%		
	Third test acc.:	62%	92%	89%	75%	100%		
	Mean prediction acc.:	**67%**	**79%**	**93%**	**92%**	**100%**	Mean prediction acc.:	86%
	Mean assigned prob. acc.:	**92%**	**100%**	**100%**	**100%**	**100%**	Mean assigned prob. acc.:	98%
	Mean assigned prob. SD:	23%	19%	12%	13%	3%	Avg. # of feats. in each CC	5.4
	Avg. # of feats.:	**7**	**6**	**5.7**	**4.7**	**3.7**		
500x	First test acc.:		85%	78%	50%	75%		
	Second test acc.:		62%	89%	50%	50%		
	Third test acc.:		92%	78%	100%	100%		
	Mean prediction acc.:		**79%**	**81%**	**67%**	**75%**	Mean prediction acc.:	76%
	Mean assigned prob. acc.:		**92%**	**100%**	**50%**	**100%**	Mean assigned prob. acc.:	86%
	Mean assigned prob. SD:		23%	18%	26%	19%	Avg. # of feats. in each CC	5.8
	Avg. # of feats.:		**6.7**	**5.3**	**6.3**	**5**		
1000x	First test acc.:	46%						
	Second test acc.:	77%						
	Third test acc.:	85%						
	Mean prediction acc.:	**69%**					Mean prediction acc.:	69%
	Mean assigned prob. acc.:	**85%**					Mean assigned prob. acc.:	85%
	Mean assigned prob. SD:	25%					Avg. # of feats. in each CC	4.0
	Avg. # of feats.:	**4**						
all magn.	First test acc.:	69%	62%	100%	100%	100%		
	Second test acc.:	77%	92%	89%	75%	100%		
	Third test acc.:	92%	100%	100%	100%	100%		
	Mean prediction acc.:	**79%**	**85%**	**96%**	**92%**	**100%**	Mean prediction acc.:	90%
	Mean assigned prob. acc.:	**92%**	**100%**	**100%**	**100%**	**100%**	Mean assigned prob. acc.:	98%
	Mean assigned prob. SD:	22%	16%	6%	11%	8%	Avg. # of feats. in each CC	4.3
	Avg. # of feats.:	**5**	**4.3**	**5.7**	**3.3**	**3**		

Two tests were performed to assess if the optimised features gave higher performance than randomly features. These tests were done by randomly selecting features for performance estimation (instead of finding optimised features) and permuting the class labels but still using optimised features. The test results are shown in Figure 34 and Table 29 . Both tests were performed on colour category 1 at 100x magnification and for each run using different hold-out test sets. The resulting performance can be compared with the obtained hold-out test accuracies given in Table 28 for optimised features.

As Table 29 reports, the selection of random features for the first run compared to the optimised test result gives a p-value of 0.215, which means that we cannot reject a null-hypotheses that the performance using optimized features was significantly better than the average accuracy obtained from prediction using randomly selected features on a 95 % confidence interval (CI). For the second and third run, the results using optimized features were significantly better than randomly selecting features from all original features on the same CI.
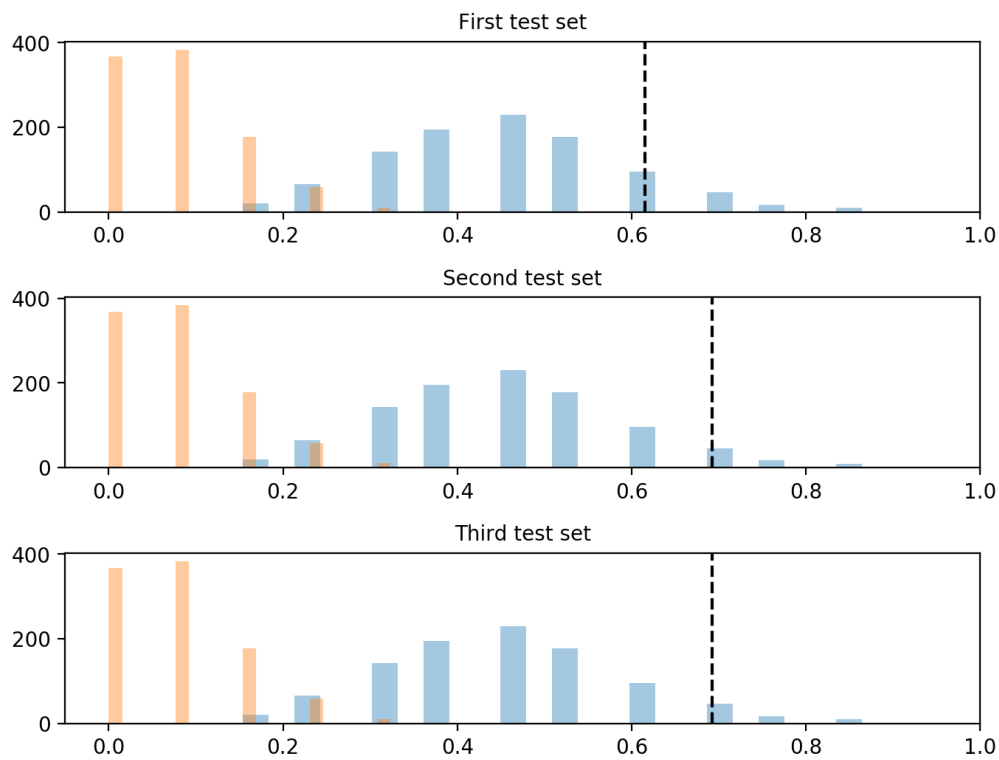
Figure 34: Histogram of obtained accuracies for randomly selected features (blue) and permuted class labels (orange) on the CC1 (at 100x) dataset, a 1000 times (each time with different random selection or permutation). The vertical line denotes the obtained performance using optimized features.

Table 29: Comparison between test run results obtained by using the optimized feature set (column 3), features selected at random (columns 4 and 5) and permuted classed labels (columns 6 and 7), on the CC1 (at 100x) dataset. The selection of random features and assessment of performance were repeated 1000 times (each time with different random selection or permutation), the same applies for the permutation test. The number of random features matches the number of optimized features on each test set.

Which test run	Number of features	Test acc.	Random features		Permuted test labels	
			Avg. acc.	p-value	Avg. acc.	p-value
first	6	62%	56%	0.215	7.2%	0.0
second	6	69%	54%	0.027	7.4%	0.0
third	6	69%	45%	0.027	7.4%	0.0

As one of the hold-out test results was not significantly better than performance achieved at random feature selection, it was possible that most of the features in the dataset contained meaningful information. Therefore, an in-depth investigation was conducted by manipulating the dataset. The original features were reduced randomly to only five features, and then 95 features of pure noise were added. Table 30 reports the classification results achieved on this new noisy dataset. The accuracy for each of the three test sets using optimized features (selection on the now 100 features) decreased but not as much as when randomly chosen features were used. Thus, the difference in performance obtained between models trained on random and optimized feature was greater in this case compared with the original dataset tested in

Table 29. In Table 30, the performance for all three optimized features was significantly better that averaged performance obtained by randomly selecting features, on a 95 % CI.
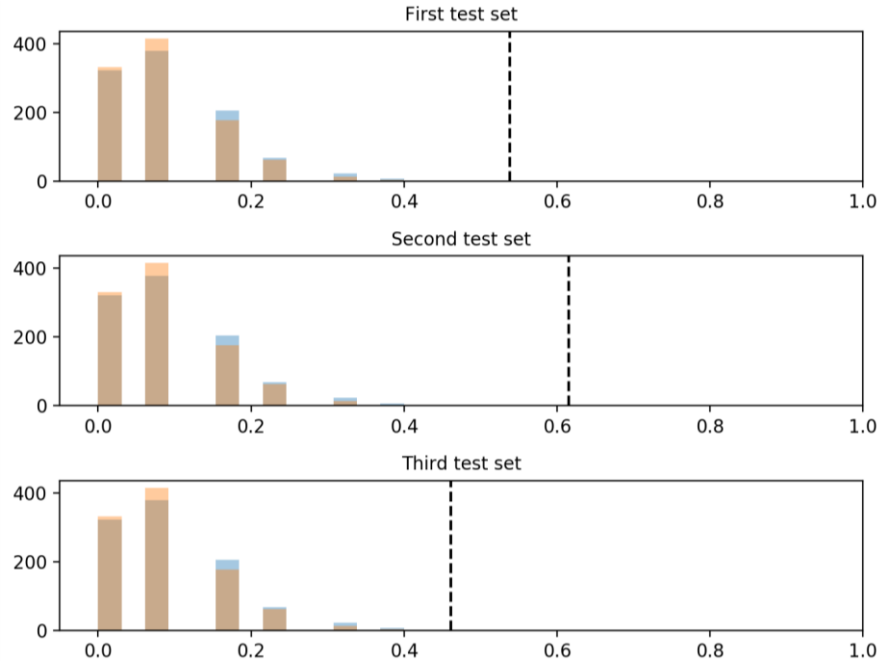
*Figure 35: Histogram of obtained accuracies for randomly selected features (blue) and permuted class labels (orange) on the **modified (noisy)** CC1 (at 100x) dataset, a 1000 times (each time with different random selection or permutation). The horizontal axis denotes accuracy, the vertical axis denotes the counts. The vertical line denotes the obtained performance using optimized features.*

Table 30: Comparison between test run results obtained using the optimized feature set (column 3) on noisy data (5 random original features and 95 noisy features) of the CC1 (at 100x) dataset, features selected at random (columns 4 and 5) and permuted classed labels (columns 6 and 7). The selection of random features and assessment of performance was repeated 1000 times, the same applies for the permutation test. The number of random features matches the number of optimized features on each test set.

Which test run	Number of features	Test acc.	Random features		Permuted test labels	
			Avg. acc.	p-value	Avg. acc.	p-value
first	2	54%	10%	0.0	7.2%	0.0
second	2	62%	8.0%	0.0	7.5%	0.0
third	2	46%	8.6%	0.0	7.9%	0.0

4.3. Model building for hyperspectral images

The LDA classifier was used for model development on the datasets originating from the hyperspectral images. As the hyperspectral images were acquired at only one magnification, the number of models developed equals the number of colour categories in the datasets. In this section, results are shown for colour category 3. In addition, an overview of the overall results for all colour categories is given in section 4.3.2. The remaining results can be found in appendix II.

4.3.1. Colour category 3

Feature selection was performed three times on the dataset for colour category 3. Different test sets were held out from the dataset in each of the three times. Table 31 shows the validation curves arising from the three feature selections. In addition, the selected features are shown to the right of their respective validation curves. In all three cases, only two features were selected. The validation accuracies were constant from 30 features down to only two features at above 90 % accuracy, before dropping down to around 60 % for the last feature. The union of the features indicates that the selected features (wavelengths) appear in two bands, as visualized in the three plots in

Table 32. These plots show the spectra for all samples, the mean spectra from each sample holder, and the average spectra for each class.

The first, second and third test results are shown in Table 33, Table 34 and 28. All samples were correctly classified. On average, as Table 36 reports, eight out of eleven classes were correctly classified with assigned probabilities above 90 %. The variation of assigned probabilities is given in Table 37. The class Aus_Run was on average assigned to USA_Fal with a 32 % probability, at a SD of only 3.5 %.

Table 31: Validation curves obtained during feature selection on the hyperspectral image dataset where each curve was obtained with a different hold-out test set. The first column denotes which run of the feature selection the curve on the corresponding row belongs to, the second column contains validation graphs for the respective runs, and the third column contains the selected features for these runs. The bottom row gives the union of features of the third column. The horizontal and vertical axes of the validation graphs are respectively the number of features and accuracy. The graphs should be read from right to left since there was a reduction of features, the number of features were reduced from 30 to 1. There are three curves in each of the three validation graphs. The blue curve is the training accuracy, green is validation accuracy, and the red is their difference. The light blue and green shaded bands are their corresponding SD.

Which run	Validation curves	Features selected
First		2 features in total: 1478.01 1563.20
Second		2 features in total: 1478.01 1556.09
Third		2 features in total: 1485.09 1566.75
Union features	1478.01, 1485.09, 1556.09, 1563.20, 1566.75	

Table 32: The spectra for all samples (top) and the mean for each sample holder (middle) and each class (bottom). The top figure plots the spectra for all samples, the middle figure plots the average spectra for each sample holder, the bottom figure plots the average spectra for each class. All spectra are colour coded by class, shown in the top left corner in each plot. The black vertical dotted lines indicate all selected wavelengths (i.e. union of features, Table 24) given in the union feature set. The wavelengths (nm) along the horizontal axis are given in [nm] and the vertical axis is absorbance in arbitrary units.
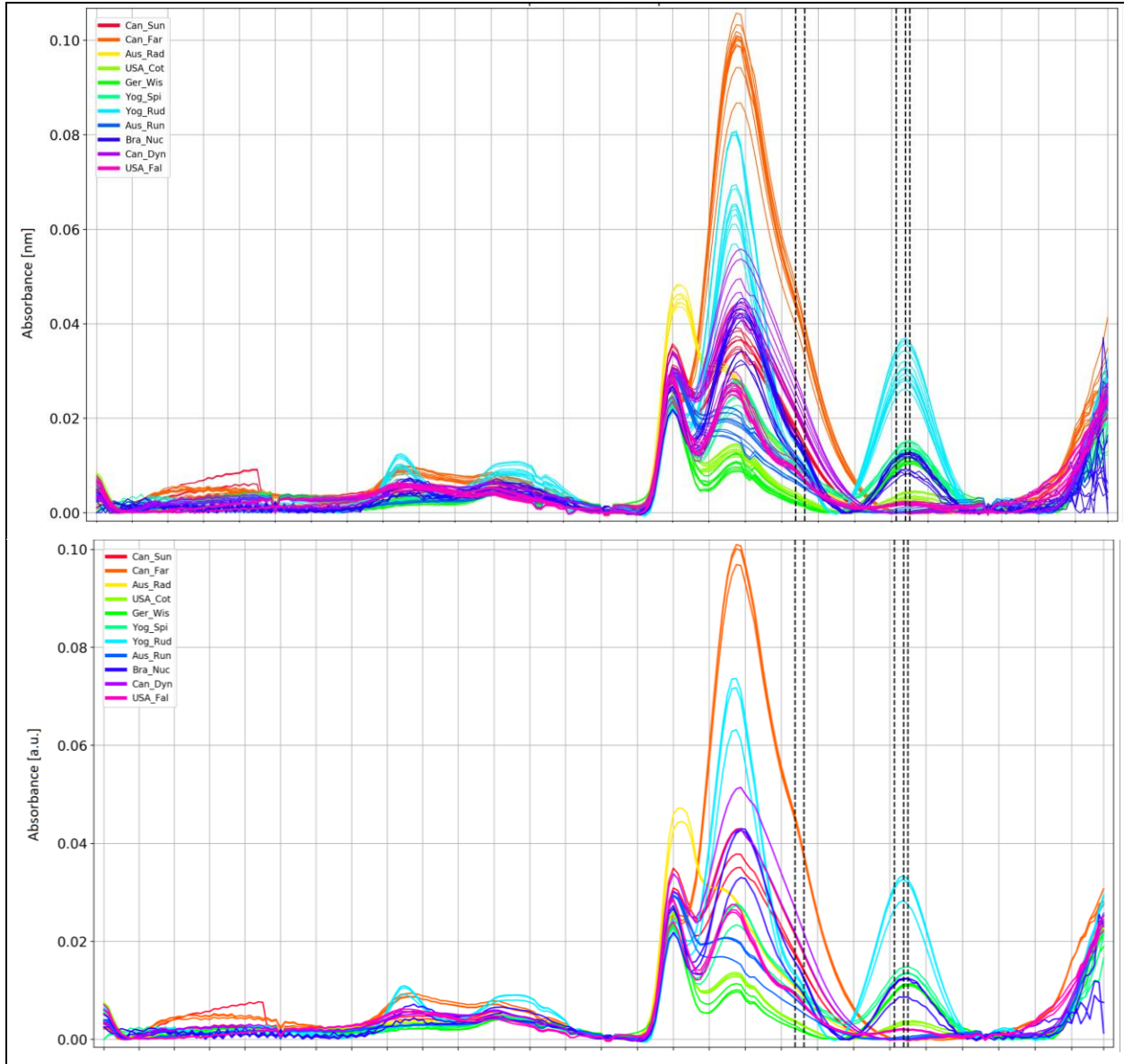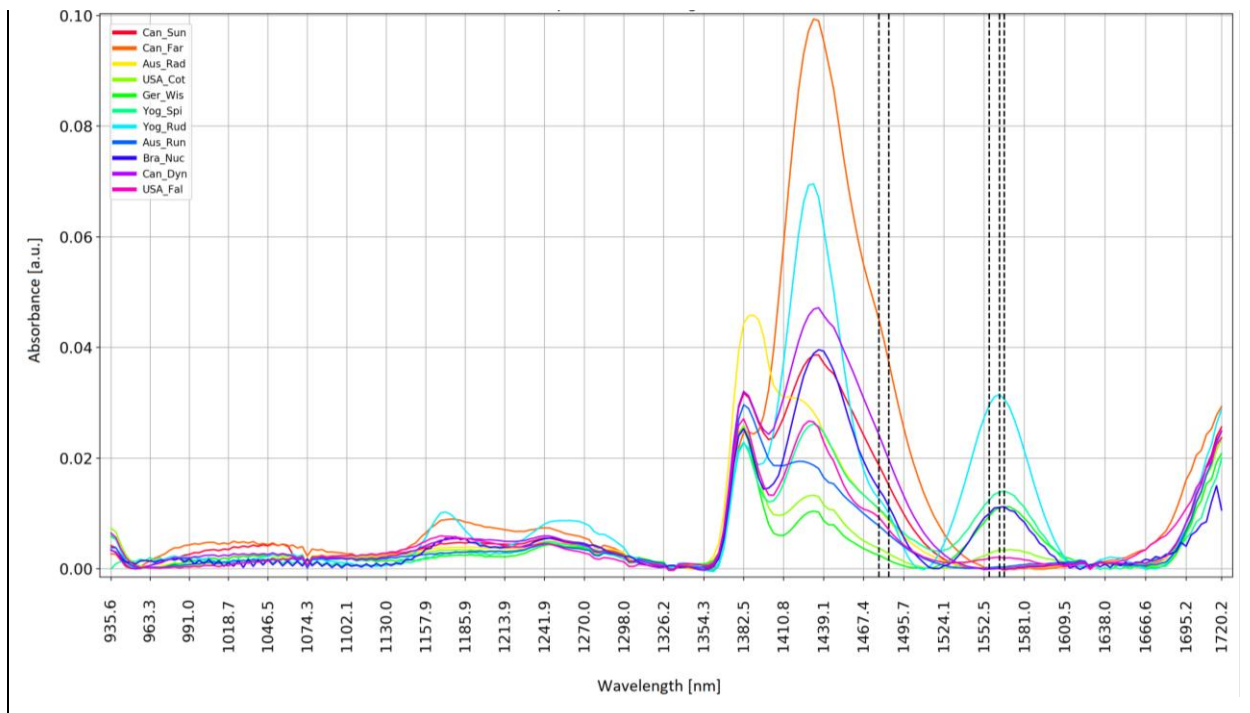
Table 33: Prediction matrix of assigned probabilities for the first hold-out test set of the hyperspectral image dataset for colour category 3. The first column denotes the names of the original images that have been classified, and the second column denotes the true class of these samples. The rest of the columns on the first row denote the predicted classes.

filenames	Pred. / True	Aus_Rad	Aus_Run	Bra_Nuc	Can_Dyn	Can_Far	Can_Sun	Ger_Wis	USA_Cot	USA_Fal	Yog_Rud	Yog_Spi
QD_40-02_0201BL	Aus_Rad	88.46%	6.63%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	4.91%	0.00%	0.00%
QD_65-02_0144BL	Aus_Run	4.41%	69.78%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	25.81%	0.00%	0.00%
QD_15-02_0066TR	Bra_Nuc	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_20-02_0094BL	Can_Dyn	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_27-03_0255BR	Can_Far	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_25-02_0242TL	Can_Sun	0.00%	0.00%	0.00%	0.01%	0.00%	99.99%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_18-02_0186BR	Ger_Wis	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
QD_07-01_0235BL	USA_Cot	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%
QD_31-1_0045BR	USA_Fal	3.52%	23.91%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	72.57%	0.00%	0.00%
QD_72-01_0098BL	Yog_Rud	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
QD_29-02_0080TR	Yog_Spi	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.99%

Table 34: Prediction matrix of assigned probabilities for the second hold-out test set of the hyperspectral image dataset for colour category 3. The first column denotes the names of the original images that have been classified, and the second column denotes the true class of these samples. The rest of the columns on the first row denote the predicted classes.

filenames	Pred. / True	Aus_Rad	Aus_Run	Bra_Nuc	Can_Dyn	Can_Far	Can_Sun	Ger_Wis	USA_Cot	USA_Fal	Yog_Rud	Yog_Spi
QD_40-01_0198BR	Aus_Rad	73.56%	19.18%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	7.26%	0.00%	0.00%
QD_65-01_0143TL	Aus_Run	0.18%	58.63%	0.00%	0.00%	0.00%	0.00%	0.00%	2.51%	38.68%	0.00%	0.00%
QD_15-03_0068BR	Bra_Nuc	0.00%	0.00%	99.98%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%
QD_20-01_0093BL	Can_Dyn	0.00%	0.00%	0.00%	93.64%	0.00%	6.36%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_27-02_0251BL	Can_Far	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_25-01_0240TR	Can_Sun	0.00%	0.00%	0.00%	13.37%	0.00%	86.63%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_18-03_0187TR	Ger_Wis	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
QD_07-03_0238BL	USA_Cot	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%
QD_31-3_0048TR	USA_Fal	12.65%	12.47%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	74.88%	0.00%	0.00%
QD_72-02_0102TR	Yog_Rud	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
QD_29-03_0081BL	Yog_Spi	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.99%

Table 35: Prediction matrix of assigned probabilities for the third hold-out test set of the hyperspectral image dataset for colour category 3. The first column denotes the names of the original images that have been classified, and the second column denotes the true class of these samples. The rest of the columns on the first row denote the predicted classes.

filenames	Pred. / True	Aus_Rad	Aus_Run	Bra_Nuc	Can_Dyn	Can_Far	Can_Sun	Ger_Wis	USA_Cot	USA_Fal	Yog_Rud	Yog_Spi
QD_40-02_0201BR	Aus_Rad	78.64%	15.22%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	6.14%	0.00%	0.00%
QD_65-03_0145BL	Aus_Run	3.22%	65.91%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	30.87%	0.00%	0.00%
QD_15-01_0065BR	Bra_Nuc	0.00%	0.00%	99.73%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.27%
QD_20-01_0093TL	Can_Dyn	0.00%	0.00%	0.00%	88.28%	0.00%	11.72%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_27-01_0247BL	Can_Far	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_25-03_0245BL	Can_Sun	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_18-01_0189BR	Ger_Wis	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
QD_07-02_0239TL	USA_Cot	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%
QD_31-2_0046TR	USA_Fal	2.14%	35.75%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	62.11%	0.00%	0.00%
QD_72-03_0105TR	Yog_Rud	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
QD_29-01_0077TR	Yog_Spi	0.00%	0.00%	1.29%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	98.71%

Table 36: The average assigned probabilities for the three hold-out test sets, for each combination of true and predicted class on the hyperspectral image dataset for colour category 3.

Pred. / True	Aus_Rad	Aus_Run	Bra_Nuc	Can_Dyn	Can_Far	Can_Sun	Ger_Wis	USA_Cot	USA_Fal	Yog_Rud	Yog_Spi
Aus_Rad	80.22%	13.68%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	6.10%	0.00%	0.00%
Aus_Run	2.60%	64.77%	0.00%	0.00%	0.00%	0.00%	0.00%	0.84%	31.78%	0.00%	0.00%
Bra_Nuc	0.00%	0.00%	99.90%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%
Can_Dyn	0.00%	0.00%	0.00%	93.97%	0.00%	6.03%	0.00%	0.00%	0.00%	0.00%	0.00%
Can_Far	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Can_Sun	0.00%	0.00%	0.00%	4.46%	0.00%	95.54%	0.00%	0.00%	0.00%	0.00%	0.00%
Ger_Wis	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
USA_Cot	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%
USA_Fal	6.10%	24.04%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	69.86%	0.00%	0.00%
Yog_Rud	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
Yog_Spi	0.00%	0.00%	0.44%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.56%

Table 37: SD of the assigned probabilities across the three hold-out test sets, for each combination of true and predicted class on the hyperspectral image dataset for colour category 3.

Pred. / True	Aus_Rad	Aus_Run	Bra_Nuc	Can_Dyn	Can_Far	Can_Sun	Ger_Wis	USA_Cot	USA_Fal	Yog_Rud	Yog_Spi
Aus_Rad	**2.84%**	**2.32%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	**0.54%**	0.00%	0.00%
Aus_Run	**1.31%**	**3.20%**	0.00%	0.00%	0.00%	0.00%	0.00%	**1.04%**	**3.49%**	0.00%	0.00%
Bra_Nuc	0.00%	0.00%	**0.10%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	**0.10%**
Can_Dyn	0.00%	0.00%	0.00%	**2.61%**	0.00%	**2.61%**	0.00%	0.00%	0.00%	0.00%	0.00%
Can_Far	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Can_Sun	0.00%	0.00%	0.00%	**5.56%**	0.00%	**5.56%**	0.00%	0.00%	0.00%	0.00%	0.00%
Ger_Wis	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
USA_Cot	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
USA_Fal	**4.33%**	**9.50%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	**5.25%**	0.00%	0.00%
Yog_Rud	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Yog_Spi	0.00%	0.00%	**0.54%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	**0.54%**

4.3.2. General results for hyperspectral imaging data

Table 38 reports the overall results obtained for the classification of each colour category based on hyperspectral imaging. The classification accuracy of the three test runs was 100 %, except for the first test run on colour category 6, despite using about twice as many features. However, if classifications were based on the average assigned probabilities, then the average accuracy of the three tests would have been 100 %.

Table 38: Summary of classification performance on hyperspectral image datasets. The first column gives a description of the summarized performance, the 2nd to 5th column denotes the colour category and the last column shows the average over all colour classes of mean prediction accuracy, mean attribution accuracy, and average number of features used in the classification.

Description	CC1	CC3	CC4	CC6	Mean over all colour categories	
First test acc.:	100%	100%	100%	67%		
Second test acc.:	100%	100%	100%	100%		
Third test acc.:	100%	100%	100%	100%		
Mean prediction acc.:	**100%**	**100%**	**100%**	**89%**	Mean prediction acc.:	97%
Mean assigned prob. acc.:	**100%**	**100%**	**100%**	**100%**	Mean assigned prob. acc.:	100%
Mean assigned prob. SD:	0%	13%	3%	13%	Avg. # of feats. in each CC	2.9
Avg. # of feats.:	**2**	**2**	**2.7**	**4.7**		

As was done on the colour category 1 dataset originating from unpressed SEM images at 100x magnification, two tests were performed on the colour category 3 dataset originating from hyperspectral images. These tests were to (1) assess the added benefit of selecting features with the used approach compared to randomly selecting features, and (2) checking the performance of

optimised features on permuted class labels. The results from these tests are shown in Figure 36 and Table 39.
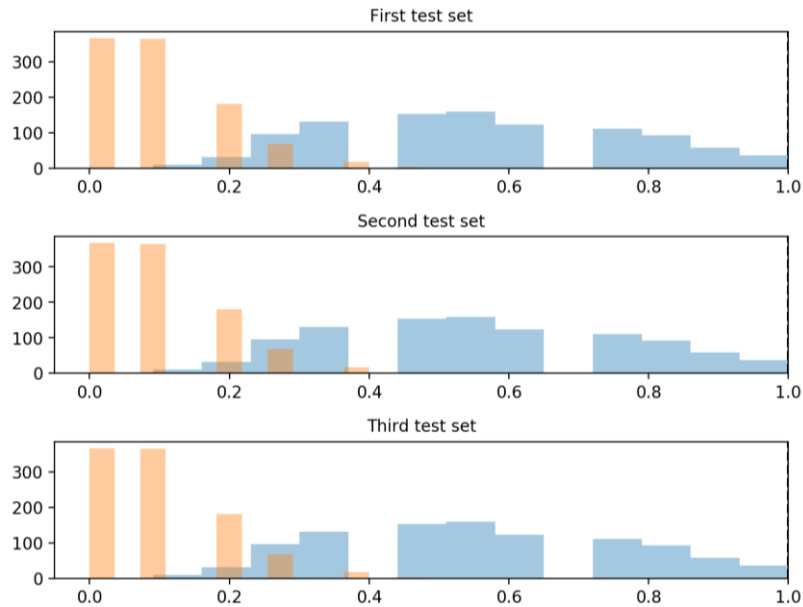
Figure 36: Histogram of obtained accuracies for randomly selected features (blue) and permuted class labels (orange) on the CC3 dataset of spectra, a 1000 times (each time with different random selection or permutation). The horizontal axis denotes accuracy, the vertical axis denotes the counts. The vertical line (at 1.0) denotes the obtained performance using optimized features.

Table 39: Comparison between test run results obtained by using the optimized feature set (column 3), features selected at random (columns 4 and 5) and permuted classed labels (columns 6 and 7) for the hyperspectral image dataset CC3. The selection of random features and assessment of performance were repeated 1000 times, the same applies for the permutation test. The number of random features matches the number of optimized features on each test set.

Which test run	Number of features	Test acc.	Random features		Permuted test labels	
			Avg. acc.	p-value	Avg. acc.	p-value
first	2	100%	52%	0.0	9.2%	0.0
second	2	100%	50%	0.0	9.2%	0.0
third	2	100%	56%	0.0	9.2%	0.0

There was a difference in performance between random and optimized feature trained models, but not as extensive as the difference using SEM dataset, shown in Table 22 and 23. Each of the models trained on optimized features gave results that were significantly better (on 95 % CI) than average performance given by randomly selecting features.

To repeat the same tests as was done on the SEM dataset, the hyperspectral dataset was manipulated to contain only five random original features and 95 other features of pure noise. Table 33 reports the results achieved on this new dataset. The performance from randomly selecting features fell, and the p-values were unchanged at zero.
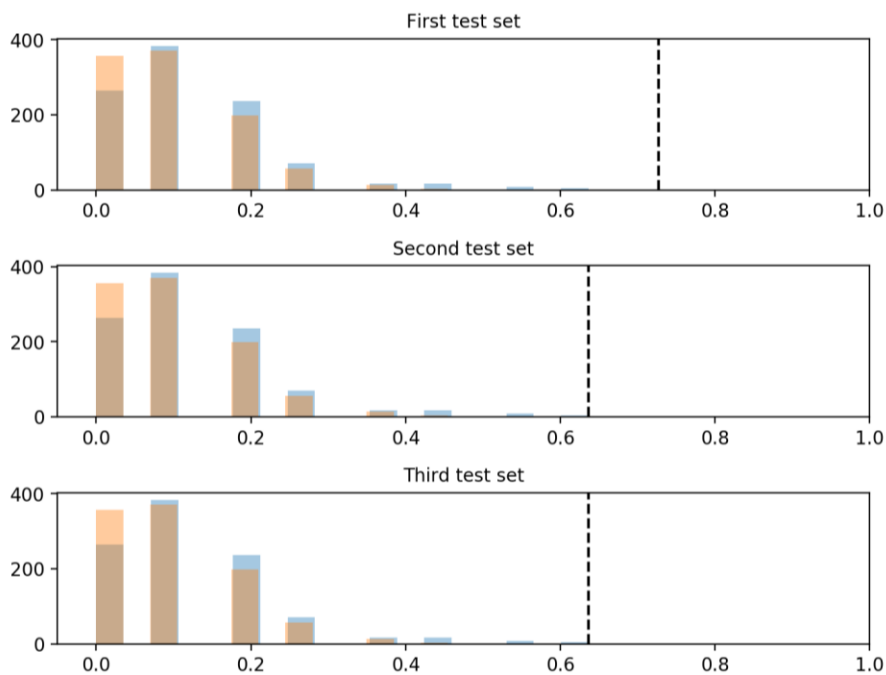
Figure 37: Histogram of obtained accuracies for randomly selected features (blue) and permuted class labels (orange) on the modified (noisy) CC3 dataset of spectra, a 1000 times (each time with different random selection or permutation). The horizontal axis denotes accuracy, the vertical axis denotes the counts. The vertical line denotes the obtained performance using optimized features.

Table 40: Comparison between hold-out test results obtained using optimized features (column 3) on noisy data (5 randomly original and 95 noisy features), features selected at random (columns 4 and 5) and permuted classed labels (columns 6 and 7) for the hyperspectral image dataset CC3. The selection of random features and assessment of performance were repeated 1000 times, the same applies for the permutation test. The number of random features matches the number of optimized features on each test set.

Which test run	Number of features	Test acc.	Random features		Permuted test labels	
			Avg. acc.	p-value	Avg. acc.	p-value
first	3	73%	11%	0.0	9.2%	0.0
second	3	64%	12%	0.0	9.2%	0.0
third	2	64%	12%	0.0	9.2%	0.0

The wavelengths selected for all colour categories are shown in Figure 38 along with the union of wavelengths found for all colour categories. Two pairs of wavelengths from different colour categories occurred at the same wavelength. This is not visible in the figure. Wavelengths 1478.01 *nm* and 1252.41 *nm* were selected for colour categories ['CC1', 'CC3'] and ['CC4', 'CC6'], respectively. The absorbance of all the spectra change of the most within the range between 1354 *nm* and 1638 *nm*. A local absorbance peak at 1382.5 *nm* seems to be present for all spectra. Many of the spectra in colour category 1 differed from the others by having relatively high absorbance between 1439 *nm* and 1552 *nm*. Some spectra of colour category 1 did not following this trend, instead resembling some spectra of the colour category 4 (blue).
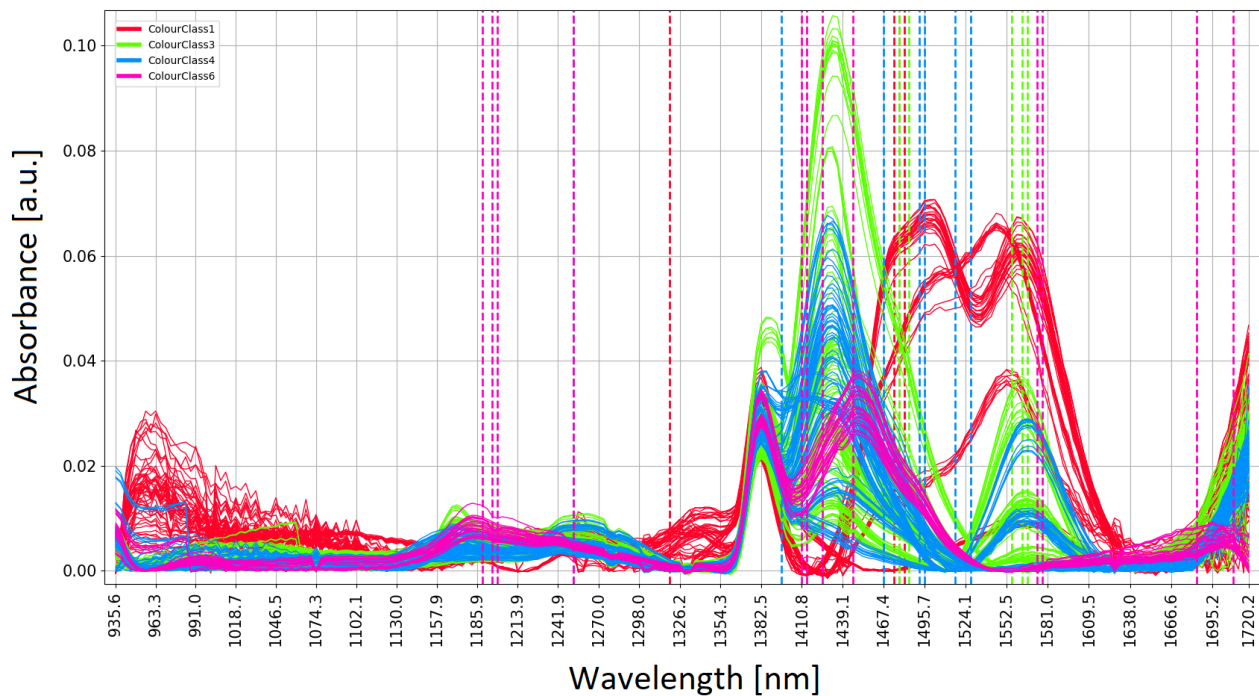
Figure 38: The spectra dataset with wavelengths in nm (abscissa), relative absorbance (ordinate) and stipulate lines denoting the union feature set.

4.3.3. The effect of removing correlated wavelengths from colour category 3

As observed in chapter 4.3.1, the selected features grouped into two wavelength bands. In the two bands, two and three wavelengths were selected. After inspecting

Table 32, it appeared that the wavelengths within each band, which were a maximum of 11 *nm* wide, were closely correlated. Therefore, it was of interest to investigate if the method of feature selection could select a smaller union set of features if the features of the dataset were less correlated. Hence, the pre-processing method was developed to discard features given a threshold of correlation (chapter 3.4.5).

For the colour category 3 dataset, the pre-processing method discarded 189 features, thereby retaining only 35 features. Feature selection on this reduced dataset (of 35 features) resulted again in two bands, but now with one feature less in each band. The bands were separated by 35 nm and 4 nm from the two bands in chapter 4.3.1. Even though the number of features was reduced to 35, the accuracies remained unchanged at 100 % for each test. This is reported in Table 41. Furthermore, as a result of starting with fewer features less computational time was needed to find and test the optimal feature sets.

Table 41: Accuracies, selected features and the union feature set when the dataset of colour category 3 was pre-processed by removing correlating features.

Which test	Acc.	Selected feature set		
first	100%	1442.6		1552.53
second	100%	1442.6	1545.42	
third	100%	1442.6	1545.42	
Union features		1442.6	1545.42	1552.53

5. Discussion

The aim of this thesis was to develop machine learning (ML) models for rapidly classifying the origin of UOCs. Satisfactory performance was achieved for all colour categories of UOCs, although the black-dark brown coloured UOCs were challenging to classify. This chapter first considers the methodological influence on the results along with possible errors. The obtained results are discussed in its own right before they are viewed the results in light of other literature findings. The chapter finishes with recommendations for future work.

5.1. Methodological considerations

Feature selection

The SEM images were not corrected for their number of pixels. This correction was deemed unnecessary as the images used for each dataset had the same number of pixels. However, it did differ across magnifications and colour categories. Only SEM images of category 1 and 4 UOCs had a consistent number of pixels in the image dimensions across all magnifications. The variation in pixel dimensions between the other categories were small; in fact, the maximum difference in pixel dimensions between categories 3, 5 and 6 was four pixels. Large difference in dimensions make the interpretation of optimized features between datasets difficult. Let us say that some features were selected for one category, but not for another; the cause might be the difference in image texture between the two categories or that the number of pixels in each dimension of the category images were not the same.

Mainly LBP features were selected, and hence they seemed to contain more meaningful information for discrimination by the LDA classifier. The reason why other feature groups were not selected could be that the classifier itself was not very good at using them, or that the selected parameters for the feature extraction algorithms were poorly selected.

The final features were chosen by visually examining the training and validation curve for each of the three runs on each dataset. The decision was based on the combination between the number of

features, and the trade-off between training and validation accuracy (bias-variance trade-off). This decision approach affected the model performance as classification accuracy was observed to be dependent on the number of features (when relatively few features were used). The average number of selected features for each dataset among all datasets differed by about a handful. A possibility would have been to devise an approach for automatically selecting the number of features from validation curves to avoid human interception. This would have speeded up the analyses.

The chosen metric to measure model performance throughout this work was "accuracy". This is a common metric for classification purposes (Raschka & Mirjalili, Python Machine Learning, 2019). However, other performance metrics can be used. Receiver Operating Characteristic Area Under the Curve (ROC AUC) is one technique to assess model performance given their false positive rate, and true positive rate (i.e. the metric recall) (Raschka & Mirjalili, Python Machine Learning, 2019).

Sample limitation

There was a limited amount of data available for each UOC category. This affects the models' ability to generalise the learning of known data onto new and unknown data. In addition, it is challenging to estimate the final model performance on new unknown data. A fixed amount of UOC was available for each origin (class). The UOC powder was distributed into a maximum of three sample holders for each class. Even though it was assumed that sample holders containing the same class of UOC could be treated as independent samples, there was still a limited number of available independent samples as the images acquired for each sample holder were considered to be dependent.

Per sample holder, it was acquired five non-overlapping images at each of the different magnifications (for the SEM images), as explained in chapter 3.2.2.1. Images originating from the same sample holder were treated as dependent images and labelled with group id. The group identification ensured that the dependency of images was considered during dataset splitting, by prohibiting dependent images being used in both training and validation folds at the same time. Therefore, assuming that all sample holders were independent of each other, all results are unbiased regarding images used, except for the hold-out test results. For some random sample holders, one image was preserved as a hold-out test sample; the other images from the same sample

holder were used to develop and train the model. This strategy can be discussed, but when comparing the hold-out test results with the corresponding unbiased validation curves, the hold-out test results were within the standard deviation of the validation result.

For the hyperspectral images, one image was acquired for each sample holder, and a square area within the image was cropped and then divided into four sub-images. These four sub-images were treated as one group for each sample holder and, just like for the SEM images, the dataset splitting considered the dependent image groups. Here also, the hold-out test results scored within the SD of unbiased validation performance.

A random set of unique test images was held out three times for each dataset to limit the chance of dataset split dependency on performance; this gives a more robust final model performance estimate.

The reported standard deviation (SD) values in the heatmaps, validation curves and variation of assigned probabilities of predictions were population standard deviation. It can be argued that sample standard deviation should have been used instead. However, these values were only included to illustrate the spread of values; no decisions were made based on any values of SD.

Image acquisition technique and consistency

The image focus varies between SEM images. The level of sharpness, or lack thereof, and how it affects the resulting model performance is uncertain. As unfocusing an image blurs out the pixels, it could be that relatively small-scale texture that is informative is removed. The same applies to the hyperspectral images. For some sample holders, several images were of poor image quality. The best image was visually determined and chosen for each sample; hence some images were discarded.

5.2. Findings

Classifiers and feature groups influence on performance

The heatmap of performance as a function of classifiers and feature groups on pressed black-dark brown coloured UOCs samples imaged with 250x SEM was presented in Figure 27. The heatmap indicates that LR, SVM or LDA were the best performing classifiers. LDA had the highest average accuracy, but the difference between classifiers are small. On the unpressed black-dark brown coloured UOCs acquired with SEM at 250x (Figure 30), LDA still performed best on average, but now with a greater difference of average accuracy to LR and SVM. To get a better foundation for picking out the best classifier among these, the frequency of hyperparameter sets used in the n-CV for the three classifiers with highest performance was investigated. There was no clear consistently chosen set for LR and SVM. This indicates that the hyperparameter decision was ambiguous and dependent on the data split. LDA does not have issues concerning hyperparameters and was hence the preferred classifier.

The LBP features seem to perform better than the other feature groups (excluding all feature groups together) in the first heatmap of pressed UOC sample, Figure 27. On the unpressed UOCs, the performance of LBP features was superior (Figure 30). In spite of these trends, LBP cannot be conclusively declared as the most informative features as the LBP feature group contained somewhere between two and eighteen times more features than the other groups.

Pressed VS unpressed samples performance

Some remarks on the difference in performance between pressed and unpressed datasets must be made. The performance on the pressed samples was much better than on the unpressed; this is most likely due to the different UOC sample preparation methods used for these two datasets. As the sample size for each dataset is small, it is difficult to say whether the pressed sample preparation is better than unpressed. A hypothesis is that the unpressed preparation was better at retaining agglomerates and other macro textures in the UOCs. These textures might contain useful information that is removed by pressing the sample. Since the pressed UOCs were more uniform than the unpressed, these images would not have that large texture variation within classes. At the

same time, the unpressed UOCs that retained macro structures showed larger texture variation within classes. The result is that a random image from an unpressed class of UOC differed more from its class than what an image from a pressed class would. The unpressed samples did not perform better than the pressed samples because the dataset was too small. A larger dataset would allow the machine learning algorithm to learn the variation amongst samples and generalize this to new data. The only way to test this hypothesis is to attain a larger dataset and test the model; however, the possibilities of attaining such a large dataset are slim as UOCs of different known origins are difficult to legally obtain.

Performance of SEM datasets

The performance on the SEM datasets, originating from different colour categories and magnifications were all satisfactory, especially categories 3-6 which achieved mean prediction accuracies at 90 % and above using only 100x magnification. Colour category 1 was more challenging to classify but still managed to get decent prediction accuracy. If the prediction were based on the average assigned class probabilities for the three test sets (mean assigned accuracy), then a 92 % accuracy was achieved at 250x magnification. All the other colour categories achieved 100 % accuracy when the mean assigned probabilities were used for predictions. Hence, there was no need for any colour categories to concatenate measurements from all magnifications to achieve that high accuracy. Also, the average number of optimal features was at only a maximum of 7 for all datasets. Each of the average numbers of features was often lower than the number of classes for their corresponding dataset, except for category 5 and 6, where a few features more were occasionally selected.

On CC1 at 100x, the calculated p-values confirmed that in two out of the three runs the developed model performed better than the average performance of a model using randomly selected features for model training. It was also seen indication that the method is robust for finding informational features in noisy data.

Performance of hyperspectral dataset

The performance on the hyperspectral datasets, one originating from each colour category, was excellent. The average predictive accuracy over the three hold-out test sets was 100 % for colour categories 1, 3, and 4. Only about two wavelengths were needed to classify these samples correctly. Category 6, however, was more challenging to classify, but still achieved nearly 90 % accuracy. This category also achieved 100 % by using the averaged assigned probabilities for predictions. One can also see that this category was more challenging to classify as the average number of features used among the three tests in this category was about twice as many as the others, but still performed worse than the others. Looking in Table 5 in chapter 3.2.1, it can be seen that there were only four classes in colour category 6, and their chemical compositions were all peroxides. The sample spectra can be found in appendix II, Figure 43. A visual inspection of the spectra reveals that they are all quite similar.

On CC3, the calculated p-values confirmed that in three out of the three runs the developed model performed better than the average performance of a model using randomly selected features for model training. It was also seen indication that the method is robust for finding informational features in noisy data.

Bias-variance trade-off

The appropriate levels of bias-variance trade-off were not always established while maintaining high performance for the selected features. Tendencies of overfitting models appear on the SEM images of colour category 1 at 100x magnification where six features were selected for all three validation curves (Table 13).

Feature sets

As expected, the three optimal feature sets obtained for each dataset were not all equal, which raises the question of which features should be used for unknown sample prediction. It seems that the feature selection was dependent on the split of the dataset, but the dependency might be due to randomness because features might have been correlated. As showed for colour category 3 in the

hyperspectral dataset, more features were equal across the three optimised feature sets after removing correlated features among the original features. The algorithm for removing correlated features should be tested more extensively, and on SEM data as well. Even though correlated features were removed, the three sets of optimized features were not identical. As the selection of features was dependent on the splitting of datasets, it could be advisable to include all selected features for final model use. Therefore, it is recommended to use the union feature set for predicting unknown samples. In regards to the importance of these features, those that were selected more than in one set is probably more important than the ones appearing only once. The many-times occurring features indicate that they give meaningful discriminative information independently of dataset split.

No clear wavelength bands appeared to be selected overall for the colour categories among the hyperspectral datasets. Given that all classes achieved high accuracy, it would be interesting to see if an LDA model could predict all samples from all colour categories together, and then check the number of features selected. An investigation like this might result in only a handful of wavelengths, which would be useful knowledge in the field. Furthermore, removing correlated features at the beginning and then determine specific discriminating wavelength bands can give worthful knowledge for application purposes. Both hyperspectral cameras and SEMs are expensive tools, but a camera that can detect only a few wavelengths is cheaper to build and more accessible in the field. If selected wavelengths can discriminate UOCs by machine learning, we can build specialised multispectral cameras operating at these wavelengths. Such cameras would be relatively inexpensive and portable, enabling on-site classification of unknown UOC and speeding up forensic investigation.

5.3. Other related findings

Results obtained with a SEM dataset

In a previous Master thesis on the same topic, Smit and Sogn (Smit & Sogn, 2018) divided each acquired SEM image into four sub-images and assumed that they were independent. This study assumed that such sub-images were dependent, and therefore image division was avoided. Each

individually acquired SEM image from the same sample holder was also considered dependent. Therefore, images belonging to the same sample holder were assigned a unique identification number to handle this dependency. Hence, both in the n-CV and feature selection, the images belonging to the same sample holder were not split across both training and validation sets at the same time. However, the grouping was not performed for hold-out tests handling. This potentially added bias might be comparable to Smit and Sogn hold-out test results. As this study did not replicate the study of Smit and Sogn in the combination of grouping data during splitting, it is uncertain how large the bias Smit and Sogn might have gotten in their n-CV. However, the n-CVs used in this thesis were first assessed without considering grouping. A decrease in performance was observed after the grouping of data was implemented.

In Smit and Sogn (Smit & Sogn, 2018) study, n-CV was used for checking model performance. In comparison, n-CV was used to find the best classifier in this study. No feature selection was performed prior to n-CV in this study. Therefore, some of the groups of features that were used in the n-CV contained a great number of features. In Smit and Sogn's study, only 6-9 features were used in the n-CV giving the highest achieved accuracy. Their n-CV accuracies were 88 % - 94 %, and final model performance on hold-out test sets was 92 % – 100 %. As mentioned in the last paragraph, they assumed that sub-images were independent, which could potentially have added bias into their n-CV performance estimate. In this study, the highest average n-CV performance achieved on the pressed UOC SEMs at 250x was at 95 %. Also, the dataset used in Smit and Sogn contained six classes, whereas this study's dataset contained 14. Also, Smit and Sogn performed their n-CV only once, while this study ran it 33 times with different splits of the data. Thus, by averaging the performance for these runs, the results are less prone to be dependent on the splitting of data, i.e. randomness, compared to Smit and Sogn's results. Their study's method consisted of holding out the selected test data from the start, meaning that the n-CV did not contain the hold-out test data. In this study, the initial screening of classifiers by n-CV used all data, including the hold-out test data. The test data should be held out in further work as that is a better practice to ensure no bias affecting the results. In this study, however, it is believed that the decision of the best classifier would not have been changed if the hold-out test samples were removed, because the difference in performance between LDA and the other classifiers (both regarding accuracy and stability) was thought to be so large that it would not change the outcome.

Sample preparation and the number of classes differ between the studies, making the results less comparable. However, this study achieved 67 % accuracy (on average for the three test sets) on 13 classes in the unpressed dataset of colour category 1 at 250x magnification, in contrast to the six classes Smit and Sogn classified at 100 % accuracy dataset (Smit & Sogn, 2018). It should be noted that, in contrast to Smit and Sogn's analysis, this study repeated the entire analysis for three different hold-out test sets to get three final test results in order to give more robust estimates less prone to dataset split dependency.

Results obtained with a spectral dataset

Different absorption characteristics were observed for the UOCs in this study. (Plaue, Klunder, Czerwinski, & Hutcheon, 2012) and (Plaue J. , 2013) found that the informative wavelengths in the near infra-red (NIR) spectra are mainly due to the O-H, C-H, and N-H overtone and combination bands. Example of absorption peaks of the O-H bond is 1450 nm, and 1480 nm for N-H.

Looking into the spectra of the acquired hyperspectral images (Figure 38), we see something interesting. The spectra for all classes in colour category 6 had a peak appearing around 1450 nm, found to be associated with the O-H bond (Plaue, Klunder, Czerwinski, & Hutcheon, 2012). Looking into which classes this colour category consisted of, one sees that there were four classes, all of which were peroxides. In (Plaue J. , 2013) the major compositions of the UOCs (except EverYe) were found. Irigaray had $UO_4 \cdot 2H_2O$ and $UO_4 \cdot 4H_2O$, UMobil had $UO_4 \cdot 2H_2O$, and Rabbla had $UO_4 \cdot 2H_2O$, $UO_3 \cdot 0.8H_2O$. These compositions might explain why an absorption peak was observed around 1450 nm. The classes Aus_Oly and Can_Key (see Figure 44 in appendix II) in colour category 1, which according to (Plaue J. , 2013) were classes where the major component was U_3O_8, had a peak appearing around 1505 and 1565 nm. These absorption bands were also observed in (Plaue, Klunder, Czerwinski, & Hutcheon, 2012) for the component U_3O_8.

It would have been interesting to find known absorption peaks for different components of UOCs and try using them as discriminating wavelengths instead of using a feature selection algorithm. However, there is limited literature on NIR spectroscopy of uranium compounds (Plaue, Klunder, Czerwinski, & Hutcheon, 2012).

AMT features did not show to be that informative compared to LBP features. The difference in performance could be a result of not optimal settings for extracting those features. Lorenzo Fongaro (Fongaro, Ho, Kvaal, Mayer, & Rondinella, 2016) has previously classified SEM images of UOCs with PLS-DA on AMT features.

5.4. Further work

Image acquisition and pre-processing guidelines

Inconsistent naming conventions for the images produced much unnecessary and time-consuming work. It is highly advisable to use a consistent naming convention on all images.

The procedure for preparing raw SEM images for feature extraction (such as cropping and rescaling) was not fully known. However, it was not observed difference in image dimensions (number of pixels) within datasets, which means that the resolution of the images within each dataset was equal. This was important since LBP features were the most chosen features for SEM data. Smit and Sogn (Smit & Sogn, 2018) point out that the LBP algorithm for extracting features is sensitive to the number of pixels in an image. Therefore, corrections should be made if images of unequal pixel-size appear within the same dataset. Even though this was not a problem in this thesis, it is thereby advisable to establish a standard processing method to handle the raw images before feature selection. A standard method would ensure that all images available for analysis are comparable. It might be interesting to investigate the effect of pixel dimensions on performance. Such an investigation could provide useful insights into whether higher pixel dimensions give better discrimination features for classification, or if there is no difference when using lower dimensions. When lower pixel dimensions are used, it takes a shorter computational time to extract features.

Evaluation of misclassified samples

A method to track the classification of each validation sample was implemented in the sequential backward feature selection code. This information should be investigated further. An investigation

strategy would be to find all misclassified samples and visually compare them to other correctly classified samples. Investigation of misclassified samples could reveal the cause of misclassification. For example, image acquisition faults such as an unfocused image. This information is valuable as it could be used to improve the classification of unknown samples; Images of faulty acquisition could be treated as outliers and excluded from model development, probably increasing the performance. Moreover, because the error of image acquisition was found, that knowledge should then be used to prevent faulty image acquisition of unknown samples. If the bad practice of image acquisition cannot be determined, then all known images should be included for model training as the same kind of "outliers" might appear in the acquisition of images of new unknown samples.

Other strategies for feature selection

The baseline standard was to input all feature into the feature selection at the start of LDA model development, without any feature reduction. As the n-CVs indicated that using LBP features alone were almost as good as using all features, it would be interesting to see how good the models would become if feature selection were applied only on LBP features. In addition, computational time would decrease since a fewer number of features would have been used.

As satisfactory results were obtained by using first feature importance by permutation, then sequential backward feature selection, and in the end, the selection of optimal features on validation curves, no further investigation of feature selection was conducted. However, the implementation of a genetic algorithm (GA), a simple evolutionary algorithm (EA), was briefly tested. It is an optimization technique based on randomness and is inspired by Charles Darwin's theory of natural evolution. More information can be found in (Gad, 2018). To further evaluate the goodness of the analysis´ method for optimizing features, it is recommended to try other feature selection methods, such as the mentioned GA or other EAs, or for example L1 regularisation (Raschka & Mirjalili, Python Machine Learning, 2019). An interesting approach for selecting features, proposed by (Jenul, et al., 2020), was to use an ensemble of unique elastic net trained models, called RENT (repeated elastic net technique for feature selection). The method is aimed at binary classification but can be extended to this study's multiclass problem. *Sklearn* in (1.12. Multiclass and multilabel algorithms, u.d.) explains different scheme approaches that can extend

RENT. One example could be the scheme approach one-vs-one (OVO) that was used by the support vector machine (SVM) algorithm C-Support Vector Classification (SVC)) in the initial assessment of classifiers in this study. The *Biorad* project used for feature extraction in this study also includes a tool for assessing the different combinations of feature selectors and classifiers (Albunni, u.d.). Discarding features with low variance is a method of feature reduction (Smit & Sogn, 2018). This method was tested but not included in the methodology.

Including spatial information in hyperspectral dataset

An advantage of hyperspectral imaging (HSI) is that continuous spectra are acquired for each pixel in the spatial dimension. However, in this study, only the mean spectra over the spatial dimensions were used. Now, the accuracy obtained using this straightforward approach achieved 100 % accuracy for classification for all colour category datasets but one. Nevertheless, fewer classes were included in each colour category compared to the classes included in SEM imaging. Hence, the accuracy could decrease if more classes were included. If so, and also to improve the accuracy on colour category 6, utilizing the spatial information in the hyperspectral images could increase performance. One way to include spatial information in datasets extracted from hyperspectral images is to augment a centre pixel spectrum with its neighbouring pixels spectra, for each pixel in an image. Such a dataset would contain rows equal to the number of pixels in the image, and columns equal to the centre pixel spectrum and each of the spectrum of the neighbouring pixels. Then, one could do pixel-wise classification of all pixels in an image and thereafter majority voting (e.g. the class with the highest total assignment, probability, or coherent zone of a cluster ensemble) to classify the entire image. The author implemented a code using this feature extraction method, and the idea was inspired by (Jamme & Duponchel, 2017). The reason why datasets extracted in this way were not used in this study, was again it was not necessary for getting satisfactory results. Also, this feature extraction method introduces a noteworthy challenge of dataset size, probably not a *big data* issue, but still demanding. Put in numbers, considering the eight neighbouring pixels of one-pixel distance for each centre pixel in a 250 by 250 pixel hyperspectral image of 224 wavelength bands; we get $(8+1) \times 224$ features and 250×250 rows. Also, for colour category 3 that consists of 124 samples (images), this dataset would be extracted 124 times. Hence, $9 \times 224 \times 250 \times 250 \times 124 = 15,624,000,000$ values. This is 54,394 times larger

than the biggest SEM dataset used in this study (CC1, concatenated magnifications) of 195 samples and 1473 features resulting in 287,235 values. The RAM needed to load the augmented pixel datasets into memory would be larger than the 16 Gb RAM used for conducting this study (perhaps *online learning* could be a solution (Raschka & Mirjalili, Python Machine Learning, 2019)). Different ways to tackle this size challenge include converting the numerical type of the values to lower bit-precise values, binning the wavelengths, and using only a selection(s) of the spectra. Partial least squares discriminant analysis (PLS-DA) is recommended for handling a large number of features. It is a popular algorithm for handling correlated features such as wavelengths (Pelliccia, PLS Discriminant Analysis for binary classification in Python, 2020). (Barker & Rayens, 2002) gives a statistical explanation for why PLS should be used instead of principal component analysis (PCA) for reducing dimensions with the goal of classification given data of known classes. For the interested reader, (Chevallier, Bertrand, Kohler, & Courcoux, 2006) can be read.

Assessing feature importance

In this study, feature importance was assessed for all feature groups together. It could be interesting to see if more features than those belonging to the LBP group were selected as optimal features - if one selected several features based on feature importance for each group first, before putting them together in the sequential backward selection. However, other ways for assessing the importance of features might need to be explored in further work. An article that argues against obtaining feature importance by permutation when using correlated features was discovered after the analysis was completed (Hooker & Mentch, 2019).

Additional evaluation of classifiers

An improvement of the n-CV can also be made to investigate further if other classifiers perform better than LDA, by using different groups of features and more hyperparameters to tune. In colour category 6, there was one hold-out test result that did not achieve 100 % accuracy. One could check if some other classifiers could get better performance using the same number of features,

but whether it is worthwhile is unsure. The results are good enough, and one could quickly improve that result by including one or more features.

Model performance VS information loss

Smit and Sogn (Smit & Sogn, 2018) used wavelet transformations on their SEM images; this could also have been tested in this study to improve the performance. However, an effort was made to keep the images as un-altered as possible to get more intuitive and understandable final models. In the field of nuclear forensics, it is vital to understand "why" images are assigned to a class. One could try the transformation to improve the performance, as well as feature engineering such as log transformations (Raschka, Machine Learning FAQ, u.d.), but at the cost of less understandable features.

Additional ideas for investigation

There are yet new ideas that could be interesting to investigate in further research. Examples of such are to extract textural features on score images obtained from PCA on hyperspectral images. Another alternative would be to treat the spectral dimension in hyperspectral images as spatial and use conventional feature extractors used for 3D-spatial images. These ideas have not been investigated in this thesis.

Applying graphical user interface to create a more user-friendly tool

Obvious improvements can be made to the written codes, to make them easier and more robust for use by others. Such rewritings include, among other things: remove unnecessary parts/streamlining, include instructions, more comments, and functionality such that wrong inputs show understandable errors. Also, one could add progress bars with the estimated time for finishing because some things take time to run. In addition, it would be nice to pack the codes within a graphical user interface (GUI) in such a way that non-programmers can develop their own models or predict unknown samples based on the results obtained in this study. It seems that this could be easily done with the Python package *PySimpleGUI* (PySimpleGUI, u.d.).

6. Conclusions

This thesis constructed a method that successfully classified SEM and hyperspectral images of UOCs with satisfying performance. Machine learning can discriminate the origin of UOCs based on physical appearance and absorbance spectra captured with appropriate imaging. This method is a new tool that can be used in nuclear forensic investigations to assess interdicted UOCs. Not only could this method give an indication of the origin of an unknown sample (or at least exclude some possibilities), it will do so rapidly; the only significant time usage lies in the image acquisition.

The initial assessment of the classifiers showed that the classifiers LR, SVM, and LDA were the best-performing ones on the samples of pressed black coloured UOCs acquired with SEM at 250x. These were also the highest performing classifiers on the unpressed samples of the same colour category and SEM magnification. However, the performance suggested that LDA was superior. LDA was chosen for model development based on the results and the examination of parameter frequency. The initial assessment suggested that LBP features were the most informative for discriminating classes. This observation was confirmed again in the final model assessment, where primarily LBP features were selected.

Features were selected and performance estimated for the final models. For the SEM datasets the colour categories 3, 4, 5, and 6 achieved on average at least 90 % accuracy across hold-out test data at 100x magnification. The black-coloured category was more difficult to classify but still achieved an average classification accuracy of 79 % when features of 100x, 250x, and 1000x magnification were used together. Averaging the assigned probabilities for each class across the hold-out test sets and then predicting achieved an accuracy of 92 % at 250 x magnification. The rest of the colour categories at 100x magnification achieved 100 % accuracy by following the same procedure. The models for hyperspectral data achieved 100 % accuracy on average across the hold-out test sets for the colour categories 1, 3, and 4. Colour category 6 averaged at 89 %. However, all categories achieved 100 % accuracy when predictions were made on the averaged assigned probabilities.

Three feature sets were obtained for each dataset. The final model performance probably depends on the dataset split. Therefore, the recommendation is to use the union feature set when implementing the final model. Removing all correlated features from the start might make the three sets more similar to each other. This strategy should be considered for future work and method implementation.

The thesis highlights the importance of nuclear forensics, puts UOCs classification in context. The investigation of UOCs origin might take months to complete. This makes any additional early phase analyses on evidence material highly appreciated as it may reduce the duration of the investigation. Further work should first exclusively focus on creating non-programmer-friendly application software for accessing the final models and using them for predicting interdicted UOCs. It is the authors believes that the achieved model performance is good enough for implementation, and thus should provide value in forensics as soon as possible. Further methodological and model improvements should be prioritised after implementation.

References

(n.d.). Retrieved December 14, 2020, from PySimpleGUI:
    https://pysimplegui.readthedocs.io/en/latest/readme/

1.12. Multiclass and multilabel algorithms. (n.d.). Retrieved December 14, 2020, from scikit-
    learn: https://scikit-learn.org/stable/modules/multiclass.html#multiclass

Albunni, A. (n.d.). *ahmedalbuni/biorad*. Retrieved December 14, 2020, from Git Hub:
    https://github.com/ahmedalbuni/biorad

Amigo, M. J., Babamoradi, H., & Elcoroaristizabal, S. (2015). Hyperspectral image analysis. A
    tutorial. *Analytica Chimica Acta*, 34-51.

Barker, M., & Rayens, W. (2002). Partial least squares for discrimination. *Journal of
    Chemometrics*, 166-173. doi:10.1002/cem.785

Basantia, N., Nollet, L. M., & Kamruzzaman, M. (2019). *Hyperspectral Imaging Analysis and
    Applications for Food Quality.* Taylor & Francis Group, LLC.

Budinger, P. A., Drenski, T. L., Varnes, A. W., & Mooney, J. R. (1980). The Case of the Great
    Yellow Cake Caper. *Analytical Chemistry*, 942A-948A.

Burger, W., & Burge, M. J. (2016). *Digital Image Processing* (2nd ed.). London: Springer.

Chevallier, S., Bertrand, D., Kohler, A., & Courcoux, P. (2006). Application of PLS-DA in
    multivariate image analysis. *Journal of Chemometrics*, 221-229. doi:10.1002/cem.994

Edelman, G. J., Gaston, E., van Leeuwen, T. G., Cullen, P. J., & Aalders, M. C. (2012).
    Hyperspectral imaging for non-contact analysis of forensic traces. *Forensic Science
    International*, 28-39.

Eilers, P. H., & Boelens, H. F. (2005). *Baseline Correction with Asymmetric Least Squares
    Smoothing.* Amsterdam: Biosystems Data Analysis Group, Swammerdam Institute for
    Life Sciences (SILS), University of Amsterdam.

EU Commission. (2020). *The JRC in Karlsruhe (Germany)*. Retrieved December 14, 2020, from
    https://ec.europa.eu/jrc/en/about/jrc-site/karlsruhe

Feature Importance Permutation. (n.d.). Retrieved December 14, 2020, from mlxtend: https://rasbt.github.io/mlxtend/user_guide/evaluate/feature_importance_permutation/

Fongaro, L., Ho, M. L., Kvaal, K., Mayer, K., & Rondinella, V. V. (2016). Application of the angle measure technique as image texture analysis method for the identification of uranium ore concentrate samples: New perspective in nuclear forensics. *Talanta*, 463-474.

Gad, A. F. (2018). *Practical Computer Vision Applications Using Deep Learning with CNNs.* Egypt: Apress. doi:https://doi.org/10.1007/978-1-4842-4167-7

Halstensen, M., Kvaal, K., & Esbensen, K. H. (2019). Image analytical sandstone plug poro-perm prediction using angle measure technique (AMT) and chemometrics – A feasibility study. *Chemometrics and Intelligent Laboratory Systems*, 9.

Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The Elements of Statistical Learning* (2nd ed.). Stanford: Springer.

Hooker, G., & Mentch, L. (2019). *Please Stop Permuting Features An Explanation and Alternatives.* Cornell University. Retrieved from https://arxiv.org/abs/1905.03151

IAEA. (2001). *COUNTRY NUCLEAR FUEL CYCLE PROFILES.* Vienna: IAEA. Retrieved from https://www-pub.iaea.org/MTCD/Publications/PDF/TRS404_scr.pdf

IAEA. (2014). *ApplIcAtIon of nuclEAr forEnsIcs In combAtIng IllIcIt trAffIckIng of nuclEAr And othEr rAdIoActIvE mAtErIAl.* Vienna: IAEA. Retrieved from https://www-pub.iaea.org/MTCD/Publications/PDF/TE-1730_web.pdf

IAEA. (2015). *NUCLEAR FORENSICS IN SUPPORT OF INVESTIGATIONS.* Vienna: IAEA. Retrieved from https://www-pub.iaea.org/MTCD/Publications/PDF/Pub1687web-74206224.pdf

IAEA. (2019). *NUCLEAR FUEL CYCLE SIMULATION SYSTEM: IMPROVEMENTS AND APPLICATIONS.* Vienna: IAEA. Retrieved from https://www-pub.iaea.org/MTCD/Publications/PDF/TE-1864web.pdf

IAEA. (2020, December 14). Retrieved from https://www.iaea.org/topics/nuclear-power-and-climate-change

Idaho National Laboratory. (n.d.). Retrieved December 14, 2020, from https://factsheets.inl.gov/FactSheets/ExperimentalBreederReactorI.pdf

INTERNATIONAL ATOMIC ENERGY AGENCY. (2015). ADVANCES IN NUCLEAR FORENSICS: COUNTERING THE EVOLVING THREAT OF NUCLEAR AND OTHER RADIOACTIVE MATERIAL OUT OF REGULATORY CONTROL. *SUMMARY OF AN INTERNATIONAL CONFERENCE ORGANIZED BY THE INTERNATIONAL ATOMIC ENERGY AGENCY AND HELD IN VIENNA, 7–10 JULY 2014.* Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY.

INTERNATIONAL ATOMIC ENERGY AGENCY. (2020). *NUCLEAR POWER REACTORS IN THE WORLD.* Vienna: IAEA.

Jamme, F., & Duponchel, L. (2017). Neighbouring pixel data augmentation: a simple way to fuse spectral and spatial information for hyperspectral imaging data analysis. *Journal of Chemometrics*, 8. doi:https://doi.org/10.1002/cem.2882

Jenul, A., Schrunner, S., Liland, K. H., Indahl, U. G., Futsæther, C. M., & Tomic, O. (2020). *RENT - REPEATED ELASTIC NET TECHNIQUE FOR FEATURE SELECTION.* Ås: Norwegian University of Life Sciences. Retrieved from https://arxiv.org/pdf/2009.12780.pdf

Joint Working Group of the APS and the AAAS. (n.d.). *Nuclear Forensics Role, State of the Art, and Program Needs.* AAAS Publication Services.

Keegan, E., & al., e. (2014). Nuclear forensic analysis of an unknown uranium ore concentrate sample seized in a criminal investigation in Australia. *Forensic Science International*, 111-121.

Khursheed, A. (2011). Scanning Electron Microscope Optics And Spectrometers. *World Scientific*.

Klunder, G. L., Plaue, J. W., Spackman, P. E., Grant, P., Lindvall, R. E., & Hutcheon, I. (2012). Application of Visible-Near Infrared Reflectance Spectroscopy to Uranium Ore Concentrates for Nuclear Forensic Analysis and Attribution. *Applied Spectroscopy*, 20.

Kristo, J. M., & Tumey, J. S. (2012). The state of nuclear forensics. *Nuclear Instruments and Methods in Physics Research B*, 656-661.

Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling.* New York: Springer.

Kvaal, K. (n.d.). jAMT Explorer (ImageJ plugin). Norwegian University of Life Sciences (NMBU).

Lande, I. B. (n.d.). *iLAN-Git/masterthesis*. Retrieved December 15, 2020, from GitHub: https://github.com/iLAN-Git/masterthesis

Lin, D. H. (2015). *Study on the applicability of structural and morphological parameters of selected uranium compounds for nuclear forensic purposes.* Singapore: Combined Faculties for the Natural Sciences and Mathematics of the Ruperto-Carola University of Heidelberg, Germany.

Local Binary Pattern for texture classification. (n.d.). Retrieved December 14, 2020, from scikit-image: https://scikit-image.org/docs/dev/auto_examples/features_detection/plot_local_binary_pattern.html

Manolakis, D., Lockwood, R., & Cooley, T. (2016). *Hyperspectral Imaging Remote Sensing*. Cornwall: Cambridge University Press.

Marchetti, M., Mayer, K., Wallenius, M., Antonio, B., Wiss, T., Lutzenkirchen, K., & Fongaro, L. (2019). *Image texture analysis and colorimetry for the classification of uranium ore concentrate powders.* Karlsruhe: European Commission, Joint Research Centre.

Mayer, K., Wallenius, M., & Fanghänel, T. (2007). Nuclear forensic science-From cradle to maturity. *Journal of Alloys and Compounds*, 444-445(0).

Mayer, K., Wallenius, M., & Ray, I. (2005). Nuclear forensics—a methodology providing clues on the origin of illicitly trafficked nuclear materials. *The Royal Society of Chemistry*, 433-441. doi:10.1039/b412922a

nature research CUSTOM MEDIA and Hitachi High-Technologies. (n.d.). *In pictures: details revealed with advanced SEM*. Retrieved October 31, 2020, from nature research: https://www.nature.com/articles/d42473-019-00127-2

nevernervous78. (n.d.). Retrieved December 14, 2020, from GitHub: https://github.com/nevernervous78/nirpyresearch/blob/master/snippets/Scatter_corrections_techniques.ipynb

Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002). *Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns* . Oulu: Machine Vision and Media Processing Unit Infotech Oulu, University of Oulu.

Pelliccia, D. (2018, July 21). *Two scatter correction techniques for NIR spectroscopy in Python*. Retrieved December 14, 2020, from NIRPY RESEARCH: https://nirpyresearch.com/two-scatter-correction-techniques-nir-spectroscopy-python/

Pelliccia, D. (2020, March 29). *PLS Discriminant Analysis for binary classification in Python*. Retrieved December 14, 2020, from NIRPY RESEARCH: https://nirpyresearch.com/pls-discriminant-analysis-binary-classification-python/

Pietikäinen, M., Hadid, A., Zhao, G., & Ahonen, T. (2011). *Computer Vision Using Local Binary Patterns.* London: Springer.

Plaue, J. (2013). *Forensic Signatures of Chemical Process History in Uranium Oxides.* Las Vegas: University of Nevada.

Plaue, J. W., Klunder, G. L., Czerwinski, K. R., & Hutcheon, I. D. (2012). NEAR INFRARED REFLECTANCE SPECTROSCOPY AS A PROCESS SIGNATURE IN URANIUM OXIDES. *Methods & Applications of Radianalytical Chemistry*, 10.

Python baseline correction library. (n.d.). Retrieved December 14, 2020, from Stack Overflow: https://stackoverflow.com/questions/29156532/python-baseline-correction-library

Radiomic Features. (n.d.). Retrieved December 14, 2020, from Pyradiomics: https://pyradiomics.readthedocs.io/en/v3.0/features.html

Raschka, S. (n.d.). *Machine Learning FAQ*. Retrieved December 14, 2020, from
 https://sebastianraschka.com/faq/docs/dataprep-vs-dataengin.html

Raschka, S., & Mirjalili, V. (2019). *Python Machine Learning* (3rd ed.). Birmingham: Packt
 Publishing Ltd.

Schutt, R., & O'Neil, C. (2014). *Doing Data Science.* O'Reilly Media, Inc.

scikit-learn Stratified GroupKFold #13621. (n.d.). Retrieved December 14, 2020, from Git Hub:
 https://github.com/scikit-learn/scikit-learn/issues/13621

scikit-learn/sklearn/model_selection/_split.py. (n.d.). Retrieved December 14, 2020, from Git
 Hub: https://github.com/scikit-learn/scikit-
 learn/blob/23657353d3c78bcd0bba0710be2781c898926944/sklearn/model_selection/_spl
 it.py

scipy.signal.savgol_filter. (n.d.). Retrieved December 15, 2020, from SciPy.org:
 https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.savgol_filter.html

Sequential Feature Selector. (n.d.). Retrieved December 15, 2020, from mlxtend:
 https://rasbt.github.io/mlxtend/user_guide/feature_selection/SequentialFeatureSelector/

sklearn.discriminant_analysis.LinearDiscriminantAnalysis. (n.d.). Retrieved December 15,
 2020, from scikit-learn: https://scikit-
 learn.org/stable/modules/generated/sklearn.discriminant_analysis.LinearDiscriminantAna
 lysis.html

sklearn.ensemble.AdaBoostClassifier. (n.d.). Retrieved December 14, 2020, from scikit-learn:
 https://scikit-
 learn.org/stable/modules/generated/sklearn.ensemble.AdaBoostClassifier.html

sklearn.ensemble.RandomForestClassifier. (n.d.). Retrieved December 14, 2020, from scikit-
 learn: https://scikit-
 learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html

sklearn.linear_model.LogisticRegression. (n.d.). Retrieved December 14, 2020, from scikit-
learn: https://scikit-
learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html

sklearn.naive_bayes.GaussianNB. (n.d.). Retrieved December 14, 2020, from scikit-learn:
https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.GaussianNB.html

sklearn.neighbors.KNeighborsClassifier. (n.d.). Retrieved December 14, 2020, from scikit-learn:
https://scikit-
learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html

sklearn.preprocessing.StandardScaler. (n.d.). Retrieved December 14, 2020, from scikit-learn:
https://scikit-
learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html

sklearn.svm.SVC. (n.d.). Retrieved December 14, 2020, from scikit-learn: https://scikit-
learn.org/stable/modules/generated/sklearn.svm.SVC.html

Smit, A. K., & Sogn, L. E. (2018). *Utvikling av analyseprogram for identifikasjon av
pulvermateriale basert på bildetekstur.* Ås: Fakultet for realfag og teknologi, NMBU.

van Griethuysen, J. J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., . . . Aerts,
H. J. (2017). Computational Radiomics System to Decode the Radiographic Phenotype.
Cancer Research, e104-e107. doi:10.1158/0008-5472.CAN-17-0339

Varga, Z., Wallenius, M., Mayer, K., & Meppen, M. (2011). Analysis of uranium ore
concentrates for origin assessment. *Proc. Radiochim. Acta*, 1-4.

Appendix

I. Materials and Methods

Table 42: The first order statistics (FOS) features extracted from SEM images.

FOS
10Percentile
90Percentile
Energy
Entropy
InterquartileRange
Kurtosis
Maximum
MeanAbsoluteDeviation
Mean
Median
Minimum
Range
RobustMeanAbsoluteDeviation
RootMeanSquared
Skewness
TotalEnergy
Uniformity
Variance

Table 43: The angle measure technique (AMT) features extracted from SEM images.

AMT						
0	10	20	30	40	50	60
2	12	22	32	42	52	62
4	14	24	34	44	54	64
6	16	26	36	46	56	66
8	18	28	38	48	58	68

Table 44: The Gray level Co-occurence Matrix (GLCM), Gray Level Size Zone Matrix (GLSZM), and Gray Level Run Length Matrix (GLRLM) features extracted from SEM images.

GLCM	GLSZM	GLRLM
Autocorrelation	GrayLevelNonUniformity	GrayLevelNonUniformity
ClusterProminence	GrayLevelNonUniformityNormalized	GrayLevelNonUniformityNormalized
ClusterShade	GrayLevelVariance	GrayLevelVariance
ClusterTendency	HighGrayLevelZoneEmphasis	HighGrayLevelRunEmphasis
Contrast	LargeAreaEmphasis	LongRunEmphasis
Correlation	LargeAreaHighGrayLevelEmphasis	LongRunHighGrayLevelEmphasis
DifferenceAverage	LargeAreaLowGrayLevelEmphasis	LongRunLowGrayLevelEmphasis
DifferenceEntropy	LowGrayLevelZoneEmphasis	LowGrayLevelRunEmphasis
DifferenceVariance	SizeZoneNonUniformity	RunEntropy
Id	SizeZoneNonUniformityNormalized	RunLengthNonUniformity
Idm	SmallAreaEmphasis	RunLengthNonUniformityNormalized
Idmn	SmallAreaHighGrayLevelEmphasis	RunPercentage
Idn	SmallAreaLowGrayLevelEmphasis	RunVariance
Imc1	ZoneEntropy	ShortRunEmphasis
Imc2	ZonePercentage	ShortRunHighGrayLevelEmphasis
InverseVariance	ZoneVariance	ShortRunLowGrayLevelEmphasis
JointAverage		
JointEnergy		
JointEntropy		
MCC		
MaximumProbability		
SumAverage		
SumEntropy		
SumSquares		

Table 45: The local binary pattern (LBP) features extracted from SEM images.

LBP						
L4_1_1	L16_4_12	L24_6_13	L28_7_28	L36_9_5	L40_10_8	L44_11_7
L4_1_2	L16_4_13	L24_6_14	L28_7_29	L36_9_6	L40_10_9	L44_11_8
L4_1_3	L16_4_14	L24_6_15	L28_7_30	L36_9_7	L40_10_10	L44_11_9
L4_1_4	L16_4_15	L24_6_16	L32_8_1	L36_9_8	L40_10_11	L44_11_10
L4_1_5	L16_4_16	L24_6_17	L32_8_2	L36_9_9	L40_10_12	L44_11_11
L4_1_6	L16_4_17	L24_6_18	L32_8_3	L36_9_10	L40_10_13	L44_11_12
L8_2_1	L16_4_18	L24_6_19	L32_8_4	L36_9_11	L40_10_14	L44_11_13
L8_2_2	L20_5_1	L24_6_20	L32_8_5	L36_9_12	L40_10_15	L44_11_14
L8_2_3	L20_5_2	L24_6_21	L32_8_6	L36_9_13	L40_10_16	L44_11_15
L8_2_4	L20_5_3	L24_6_22	L32_8_7	L36_9_14	L40_10_17	L44_11_16
L8_2_5	L20_5_4	L24_6_23	L32_8_8	L36_9_15	L40_10_18	L44_11_17
L8_2_6	L20_5_5	L24_6_24	L32_8_9	L36_9_16	L40_10_19	L44_11_18
L8_2_7	L20_5_6	L24_6_25	L32_8_10	L36_9_17	L40_10_20	L44_11_19
L8_2_8	L20_5_7	L24_6_26	L32_8_11	L36_9_18	L40_10_21	L44_11_20
L8_2_9	L20_5_8	L28_7_1	L32_8_12	L36_9_19	L40_10_22	L44_11_21
L8_2_10	L20_5_9	L28_7_2	L32_8_13	L36_9_20	L40_10_23	L44_11_22
L12_3_1	L20_5_10	L28_7_3	L32_8_14	L36_9_21	L40_10_24	L44_11_23
L12_3_2	L20_5_11	L28_7_4	L32_8_15	L36_9_22	L40_10_25	L44_11_24
L12_3_3	L20_5_12	L28_7_5	L32_8_16	L36_9_23	L40_10_26	L44_11_25
L12_3_4	L20_5_13	L28_7_6	L32_8_17	L36_9_24	L40_10_27	L44_11_26
L12_3_5	L20_5_14	L28_7_7	L32_8_18	L36_9_25	L40_10_28	L44_11_27
L12_3_6	L20_5_15	L28_7_8	L32_8_19	L36_9_26	L40_10_29	L44_11_28
L12_3_7	L20_5_16	L28_7_9	L32_8_20	L36_9_27	L40_10_30	L44_11_29
L12_3_8	L20_5_17	L28_7_10	L32_8_21	L36_9_28	L40_10_31	L44_11_30
L12_3_9	L20_5_18	L28_7_11	L32_8_22	L36_9_29	L40_10_32	L44_11_31
L12_3_10	L20_5_19	L28_7_12	L32_8_23	L36_9_30	L40_10_33	L44_11_32
L12_3_11	L20_5_20	L28_7_13	L32_8_24	L36_9_31	L40_10_34	L44_11_33
L12_3_12	L20_5_21	L28_7_14	L32_8_25	L36_9_32	L40_10_35	L44_11_34
L12_3_13	L20_5_22	L28_7_15	L32_8_26	L36_9_33	L40_10_36	L44_11_35
L12_3_14	L24_6_1	L28_7_16	L32_8_27	L36_9_34	L40_10_37	L44_11_36
L16_4_1	L24_6_2	L28_7_17	L32_8_28	L36_9_35	L40_10_38	L44_11_37
L16_4_2	L24_6_3	L28_7_18	L32_8_29	L36_9_36	L40_10_39	L44_11_38
L16_4_3	L24_6_4	L28_7_19	L32_8_30	L36_9_37	L40_10_40	L44_11_39
L16_4_4	L24_6_5	L28_7_20	L32_8_31	L36_9_38	L40_10_41	L44_11_40
L16_4_5	L24_6_6	L28_7_21	L32_8_32	L40_10_1	L40_10_42	L44_11_41
L16_4_6	L24_6_7	L28_7_22	L32_8_33	L40_10_2	L44_11_1	L44_11_42
L16_4_7	L24_6_8	L28_7_23	L32_8_34	L40_10_3	L44_11_2	L44_11_43
L16_4_8	L24_6_9	L28_7_24	L36_9_1	L40_10_4	L44_11_3	L44_11_44
L16_4_9	L24_6_10	L28_7_25	L36_9_2	L40_10_5	L44_11_4	L44_11_45
L16_4_10	L24_6_11	L28_7_26	L36_9_3	L40_10_6	L44_11_5	L44_11_46
L16_4_11	L24_6_12	L28_7_27	L36_9_4	L40_10_7	L44_11_6	
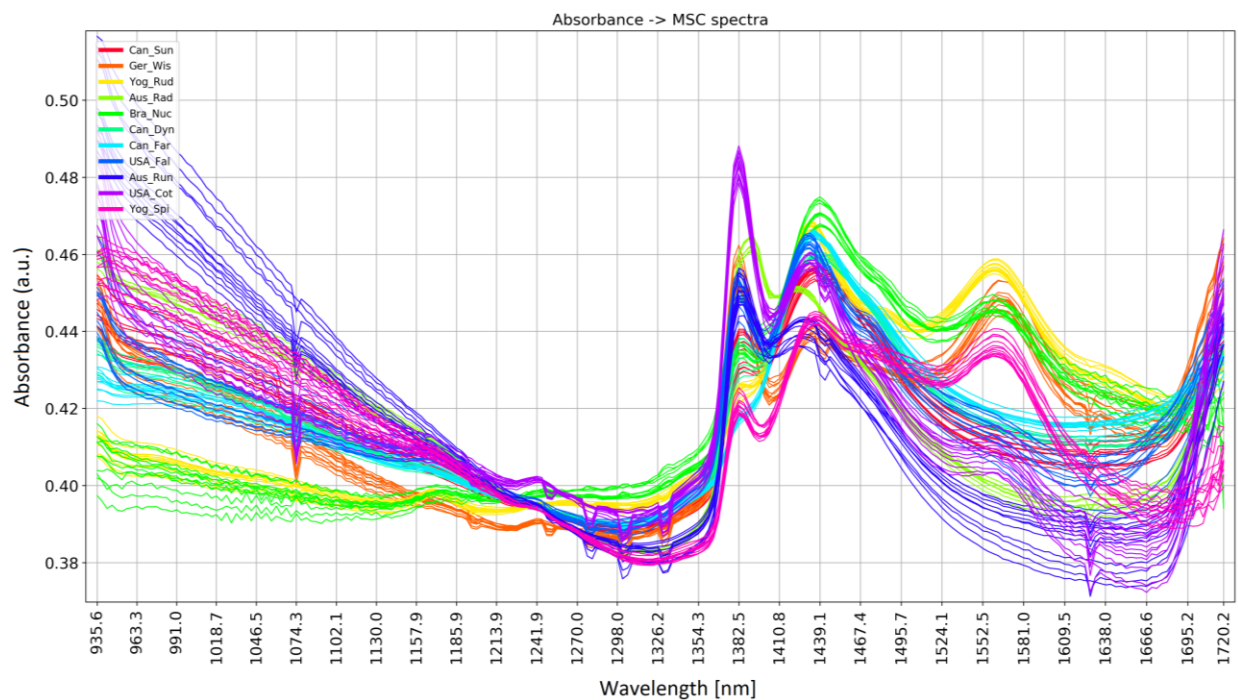
Figure 39: : MSC correction applied on the absorbance converted spectra of samples in the colour category 3 dataset. The vertical axis denotes absorbance in arbitrary units and the horizontal axis denotes wavelengths in nm.
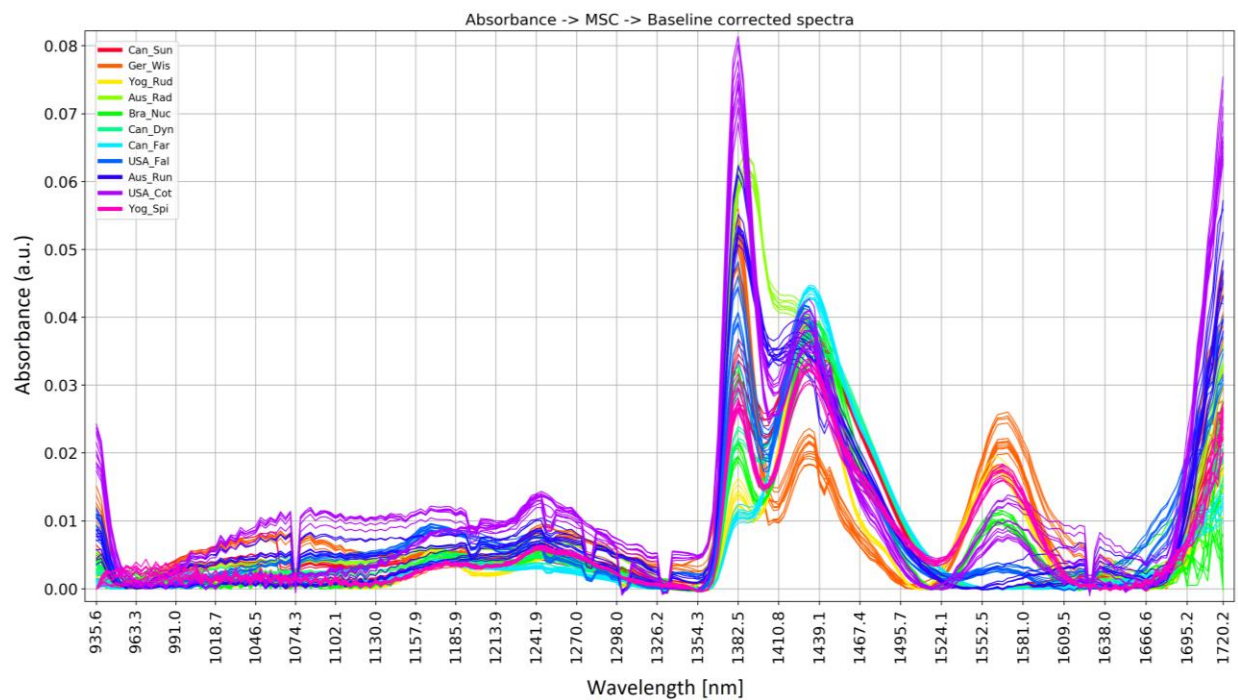
Figure 40: : Baseline correction applied on MSC corrected absorbance spectra of samples in the colour category 3 dataset. The vertical axis denotes absorbance in arbitrary units and the horizontal axis denotes wavelengths in nm.
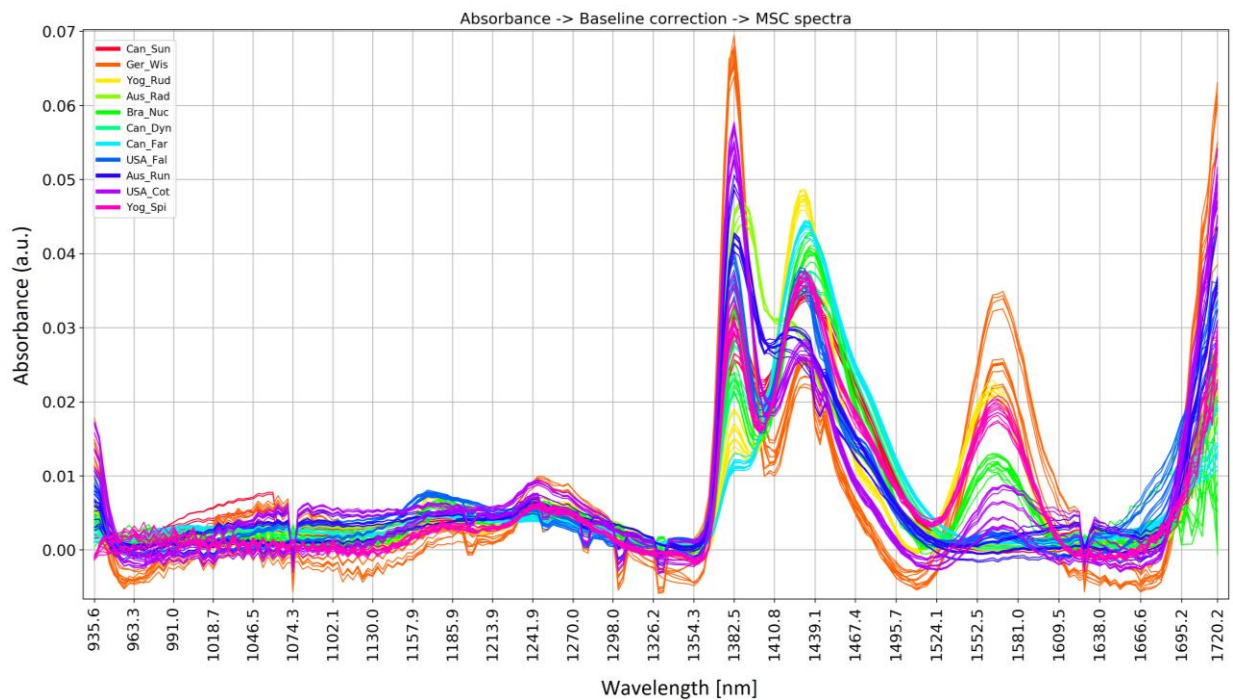
Figure 41: : MSC correction applied on baseline corrected absorbance spectra of samples in the colour category 3 dataset. The vertical axis denotes absorbance in arbitrary units and the horizontal axis denotes wavelengths in nm.
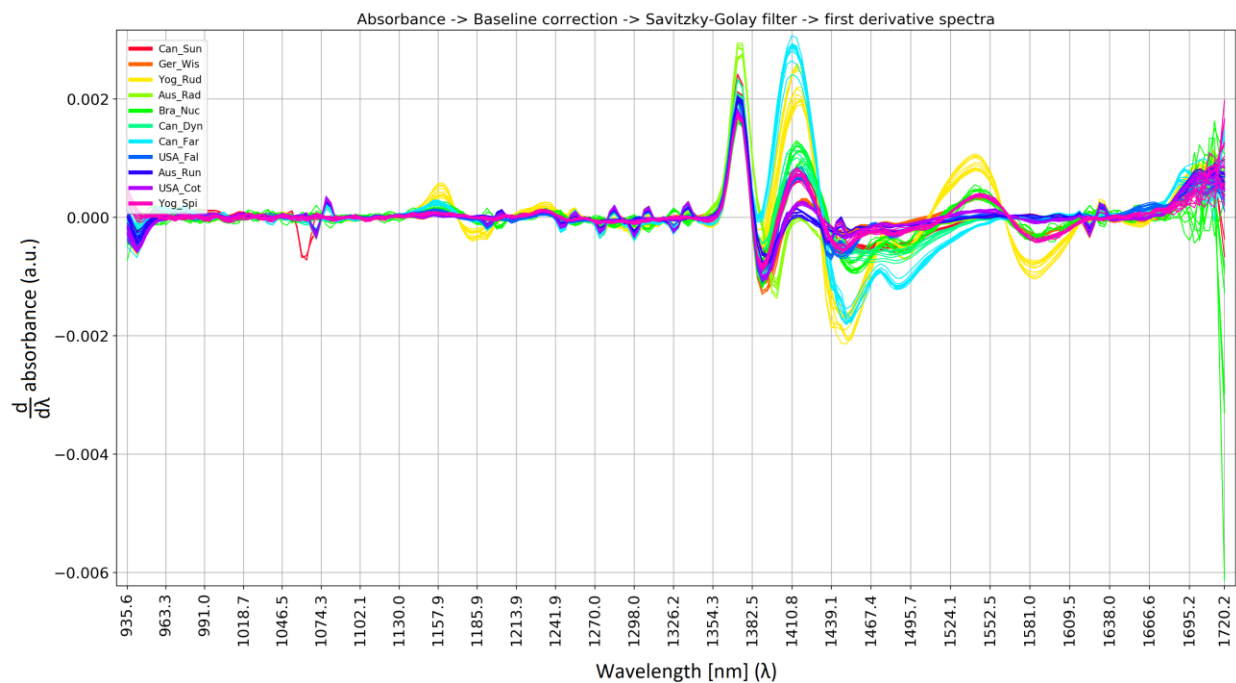
Figure 42: : First derivative transformation applied on Savitzky-Golay smoothed baseline corrected absorbance spectra of samples in the colour category 3 dataset. The vertical axis denotes the first derivative of absorbance in arbitrary units and the horizontal axis denotes wavelengths in nm.

II. Results
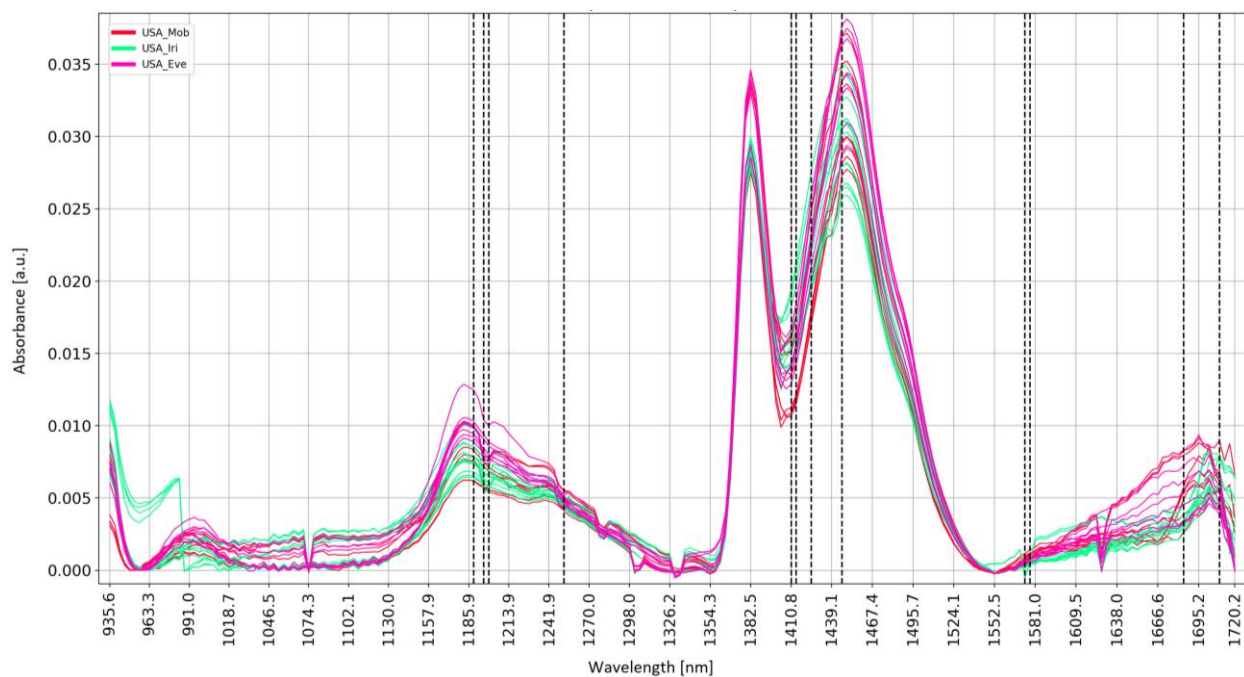
Figure 43: The spectra for all samples in CC6. The spectra are colour coded by class, shown in the top left corner in each plot. The black vertical dotted lines indicate all selected wavelengths (i.e. union of features) given in the union feature set. The wavelengths (nm) along the horizontal axis are given in [nm] and the vertical axis is absorbance in arbitrary units. The composition of all samples are peroxides.
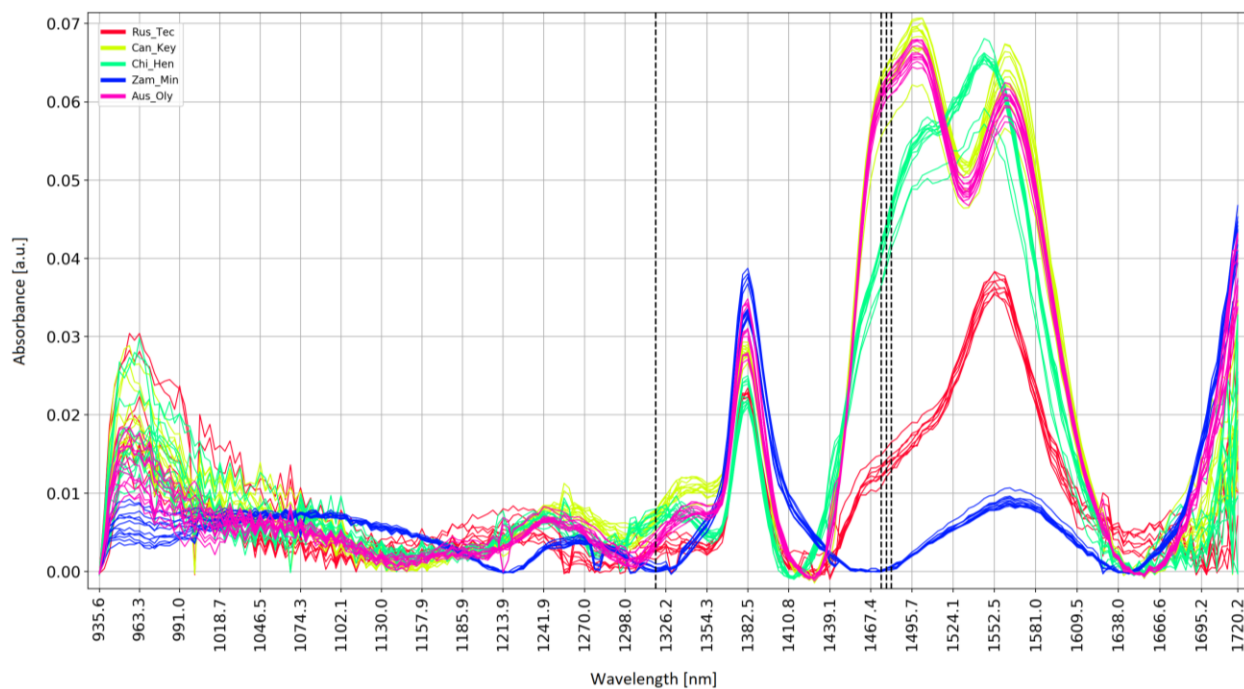
Figure 44: The spectra for all samples in CC1. The spectra are colour coded by class, shown in the top left corner in each plot. The black vertical dotted lines indicate all selected wavelengths (i.e. union of features) given in the union feature set. The wavelengths (nm) along the horizontal axis are given in [nm] and the vertical axis is absorbance in arbitrary units. The composition of all samples are peroxides.

IMAGING TECHNIQUE	COLOUR CATEGORY	MAGN.	VALIDATION CURVES		HOLD-OUT TEST RESULTS		SUMMARIZED ASSIGNED PROBABILITIES	
			Table nr.	Page nr.	Table nr.	Page nr.	Table nr.	Page nr.
SEM	1	250x	Table 46	116	Table 47	117	Table 48	118
		1000x	Table 49	119	Table 50	120	Table 51	121
	3	100x	Table 52	122	Table 53	123	Table 54	124
		250x	Table 55	125	Table 56	126	Table 57	127
		500x	Table 58	128	Table 59	129	Table 60	130
		All magn.	Table 61	131	Table 62	132	Table 63	133
	4	100x	Table 64	134	Table 65	135	Table 66	136
		250x	Table 67	137	Table 68	138	Table 69	139
		500x	Table 70	140	Table 71	141	Table 72	142
		All magn.	Table 73	143	Table 74	144	Table 75	145
	5	100x	Table 76	146	Table 77	147	Table 78	147
		250x	Table 79	148	Table 80	149	Table 81	149
		500x	Table 82	150	Table 83	151	Table 84	151
		All magn.	Table 85	152	Table 86	153	Table 87	153
	6	100x	Table 88	154	Table 89	155	Table 90	155
		250x	Table 91	156	Table 92	157	Table 93	157
		500x	Table 94	158	Table 95	159	Table 96	159
		All magn.	Table 97	160	Table 98	161	Table 99	161
HYPER-SPECTRAL CAMERA	1	-	Table 100	162	Table 101	163	Table 102	163
	3 (extra)	-	Table 103	164	Table 104	165	Table 105	166
	4	-	Table 106	167	Table 107	168	Table 108	169
	6	-	Table 109	170	Table 110	171	Table 111	171

Texture

Colour class 1

250x

Table 46
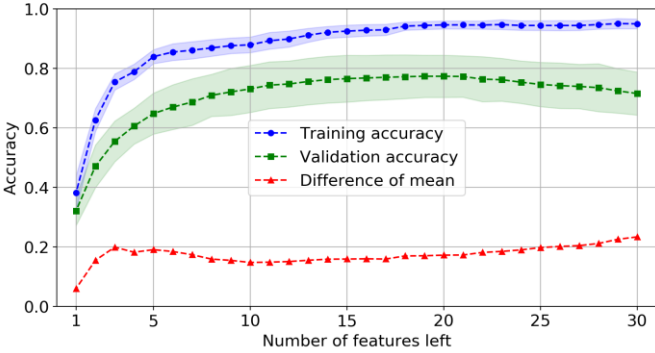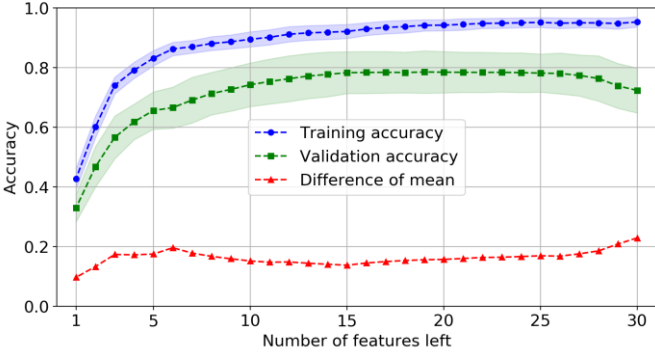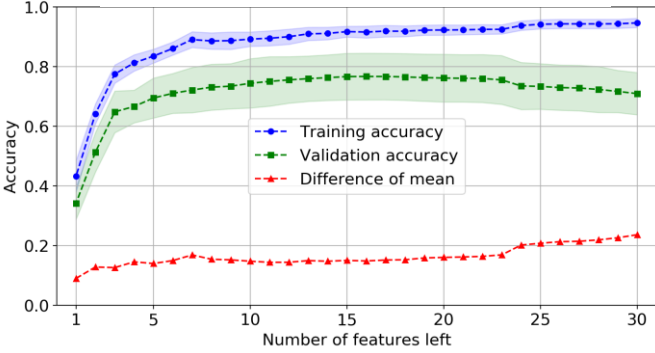
Which run	Validation curves	Features selected
First		8 features in total: L4_1_4 L24_6_1 L16_4_3 L8_2_3 L20_5_3 glcm_InverseVariance_d_3 L16_4_1 glcm_ClusterProminence_d_10
Second		8 features in total: L24_6_1 L8_2_3 L4_1_4 glcm_InverseVariance_d_3 L20_5_3 L16_4_1 L12_3_5 glcm_ClusterProminence_d_10
Third		5 features in total: L20_5_1 L8_2_3 L20_5_3 L12_3_4 glcm_ClusterProminence_d_10
Union features	glcm_ClusterProminence_d_10, L12_3_5, L16_4_3, L4_1_4, L12_3_4, L16_4_1, glcm_InverseVariance_d_3, L20_5_3, L24_6_1, L8_2_3, L20_5_1	

Table 47

First test files	Pred. / True	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
C1AuMAKU3O8250x.tif	AuMAK	80.62%	0.00%	0.00%	0.00%	0.00%	0.00%	0.50%	0.08%	7.07%	0.60%	0.05%	11.06%	0.02%
A3AuOLDU3O8250x.tif	AuOLD	22.70%	76.49%	0.32%	0.01%	0.00%	0.00%	0.14%	0.00%	0.26%	0.00%	0.01%	0.01%	0.06%
C1AuQUEU3O8250x.tif	AuQUE	0.00%	0.03%	73.25%	24.95%	0.00%	0.00%	0.00%	0.00%	0.08%	0.00%	0.03%	0.00%	1.66%
A2CaKELU3O8250x.tif	CaKEL	0.18%	0.13%	92.87%	3.80%	0.00%	0.00%	0.01%	0.00%	0.18%	0.00%	0.00%	0.00%	2.82%
A5ChHEYUO2_250x.tif	ChHEY	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A2RuTECMix_250x.tif	RuTEC	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C1SaNUFU3O8250x.tif	SaNUF	0.13%	0.00%	0.00%	0.00%	0.00%	0.00%	68.67%	24.82%	0.00%	0.00%	0.00%	6.38%	0.00%
C4SaPALU3O8250x.tif	SaPAL	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	1.31%	87.20%	0.00%	0.00%	0.00%	11.47%	0.00%
C3SaROSU3O8250x.tif	SaROS	2.88%	0.03%	0.09%	0.02%	0.00%	0.00%	0.02%	0.00%	52.22%	0.19%	44.25%	0.00%	0.30%
A3UsATLU3O8250x.tif	UsATL	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	99.33%	0.65%	0.00%	0.00%
C2UsFAPU3O8250x.tif	UsFAP	0.01%	0.01%	0.03%	0.32%	0.00%	0.00%	0.00%	0.00%	1.51%	0.00%	98.02%	0.00%	0.10%
A1UsPETMix_250x.tif	UsPET	0.08%	0.00%	0.00%	0.00%	0.00%	0.00%	0.76%	0.00%	0.22%	0.06%	98.85%	0.02%	0.00%
A4YuSPBUH__250x.tif	YuSPB	0.64%	0.04%	4.92%	1.64%	0.00%	0.00%	0.00%	0.00%	33.72%	0.39%	3.76%	0.00%	54.89%

Second test files	Pred. / True	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
A4AuMAKU3O8250x.tif	AuMAK	41.88%	0.00%	0.07%	0.00%	0.00%	0.00%	44.25%	0.25%	2.53%	0.00%	0.05%	10.69%	0.28%
B5AuOLDU3O8250x.tif	AuOLD	0.00%	19.13%	0.02%	0.07%	0.00%	80.79%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A3AuQUEU3O8250x.tif	AuQUE	0.47%	0.01%	2.56%	1.83%	0.00%	0.00%	0.00%	0.00%	89.59%	0.00%	3.06%	0.00%	2.48%
B4CaKELU3O8250x.tif	CaKEL	0.00%	0.06%	2.21%	97.38%	0.00%	0.00%	0.00%	0.00%	0.04%	0.00%	0.01%	0.00%	0.30%
B3ChHEYUO2_250x.tif	ChHEY	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C5RuTECMix_250x.tif	RuTEC	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B2SaNUFU3O8250x.tif	SaNUF	0.10%	0.00%	90.07%	3.88%	0.00%	0.00%	1.04%	0.00%	0.09%	0.00%	0.01%	0.00%	4.81%
A5SaPALU3O8250x.tif	SaPAL	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	98.47%	1.50%	0.00%	0.00%	0.00%	0.03%	0.00%
B3SaROSU3O8250x.tif	SaROS	2.24%	0.00%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	96.86%	0.27%	0.52%	0.00%	0.09%
B4UsATLU3O8250x.tif	UsATL	0.05%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	7.97%	91.70%	0.26%	0.00%	0.02%
B5UsFAPU3O8250x.tif	UsFAP	0.32%	0.00%	0.52%	0.48%	0.00%	0.00%	0.00%	0.00%	69.66%	0.02%	22.27%	0.00%	6.74%
C3UsPETMix_250x.tif	UsPET	0.22%	0.00%	0.00%	0.00%	0.00%	0.00%	0.03%	9.68%	0.00%	0.00%	0.00%	90.08%	0.00%
B4YuSPBUH__250x.tif	YuSPB	0.08%	0.00%	4.83%	2.86%	0.00%	0.00%	0.00%	0.00%	10.28%	0.02%	0.77%	0.00%	81.15%

Third test files	Pred. / True	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
B5AuMAKU3O8250.tif	AuMAK	22.36%	0.05%	0.42%	0.00%	0.00%	0.00%	0.00%	0.00%	69.32%	7.07%	0.30%	0.30%	0.20%
C4AuOLDU3O8250.tif	AuOLD	0.53%	53.50%	16.34%	3.18%	0.00%	0.00%	0.00%	0.00%	25.56%	0.00%	0.61%	0.00%	0.28%
B1AuQUEU3O8250.tif	AuQUE	1.35%	73.36%	23.21%	0.77%	0.00%	0.00%	0.30%	0.00%	0.01%	0.00%	0.00%	0.00%	1.01%
C5CaKELU3O8250.tif	CaKEL	1.84%	13.21%	33.06%	5.88%	0.00%	0.00%	0.00%	0.00%	12.75%	0.03%	0.30%	0.00%	32.94%
C4ChHEYUO2_250x.tif	ChHEY	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B5RuTECMix_250x.tif	RuTEC	0.00%	0.00%	0.00%	0.01%	0.00%	99.99%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A2SaNUFU3O8250x.tif	SaNUF	22.95%	0.00%	0.61%	0.00%	0.00%	0.00%	66.93%	2.01%	0.07%	0.00%	0.00%	6.76%	0.66%
B3SaPALU3O8250x.tif	SaPAL	0.19%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	12.74%	0.00%	0.00%	0.00%	87.07%	0.00%
A4SaROSU3O8250x.tif	SaROS	50.45%	2.51%	8.99%	0.02%	0.00%	0.00%	0.00%	0.00%	36.51%	0.31%	0.02%	0.03%	1.17%
C2UsATLU3O8250x.tif	UsATL	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.04%	99.95%	0.00%	0.00%	0.00%
A3UsFAPU3O8250x.tif	UsFAP	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	2.96%	2.29%	94.68%	0.00%	0.06%
B2UsPETMix_250x.tif	UsPET	6.57%	0.00%	0.00%	0.00%	0.00%	0.00%	0.03%	8.65%	0.00%	0.00%	0.00%	84.75%	0.00%
C4YuSPBUH__250x.tif	YuSPB	0.40%	1.53%	27.13%	11.17%	0.00%	0.00%	0.03%	0.00%	0.98%	0.00%	0.62%	0.00%	58.14%

Table 48

Average assigned probabilities

True \ Pred.	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
AuMAK	48.29%	0.02%	0.16%	0.00%	0.00%	0.00%	14.92%	0.11%	26.30%	2.56%	0.13%	7.35%	0.17%
AuOLD	7.75%	49.71%	5.56%	1.09%	0.00%	26.93%	0.05%	0.00%	8.61%	0.00%	0.20%	0.00%	0.11%
AuQUE	0.61%	24.47%	33.00%	9.18%	0.00%	0.00%	0.10%	0.00%	29.89%	0.00%	1.03%	0.00%	1.72%
CaKEL	0.67%	4.47%	42.71%	35.69%	0.00%	0.00%	0.00%	0.00%	4.32%	0.01%	0.10%	0.00%	12.02%
ChHEY	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RuTEC	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
SaNUF	7.73%	0.00%	30.23%	1.29%	0.00%	0.00%	45.55%	8.94%	0.05%	0.00%	0.01%	4.38%	1.82%
SaPAL	0.07%	0.00%	0.00%	0.00%	0.00%	0.00%	33.26%	33.81%	0.00%	0.00%	0.00%	32.85%	0.00%
SaROS	18.52%	0.85%	3.03%	0.02%	0.00%	0.00%	0.01%	0.00%	61.86%	0.25%	14.93%	0.01%	0.52%
UsATL	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	2.67%	97.00%	0.30%	0.00%	0.01%
UsFAP	0.11%	0.00%	0.18%	0.27%	0.00%	0.00%	0.00%	0.00%	24.71%	0.77%	71.66%	0.00%	2.30%
UsPET	2.29%	0.00%	0.00%	0.00%	0.00%	0.00%	0.27%	6.11%	0.07%	0.02%	32.95%	58.28%	0.00%
YuSPB	0.37%	0.52%	12.29%	5.22%	0.00%	0.00%	0.01%	0.00%	15.00%	0.14%	1.72%	0.00%	64.73%

SD of assigned probabilities

True \ Pred.	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
AuMAK	24.21%	0.02%	0.18%	0.00%	0.00%	0.00%	20.74%	0.10%	30.47%	3.20%	0.12%	4.99%	0.11%
AuOLD	10.58%	23.57%	7.63%	1.48%	0.00%	38.08%	0.07%	0.00%	11.99%	0.00%	0.28%	0.00%	0.12%
AuQUE	0.56%	34.57%	29.68%	11.16%	0.00%	0.00%	0.14%	0.00%	42.21%	0.00%	1.44%	0.00%	0.60%
CaKEL	0.83%	6.18%	37.63%	43.63%	0.00%	0.00%	0.00%	0.00%	5.96%	0.01%	0.14%	0.00%	14.83%
ChHEY	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RuTEC	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
SaNUF	10.77%	0.00%	42.32%	1.83%	0.00%	0.00%	31.48%	11.26%	0.04%	0.00%	0.01%	3.10%	2.13%
SaPAL	0.08%	0.00%	0.00%	0.00%	0.00%	0.00%	46.11%	38.03%	0.00%	0.00%	0.00%	38.62%	0.00%
SaROS	22.57%	1.17%	4.21%	0.01%	0.00%	0.00%	0.01%	0.00%	25.56%	0.05%	20.73%	0.02%	0.47%
UsATL	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	3.74%	3.75%	0.27%	0.00%	0.01%
UsFAP	0.15%	0.00%	0.24%	0.20%	0.00%	0.00%	0.00%	0.00%	31.79%	1.08%	34.95%	0.00%	3.14%
UsPET	3.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.35%	4.34%	0.10%	0.03%	46.60%	41.26%	0.00%
YuSPB	0.23%	0.71%	10.49%	4.24%	0.00%	0.00%	0.01%	0.00%	13.78%	0.18%	1.45%	0.00%	11.69%

1000x

Table 49

Which run	Validation curves	Features selected
First		**4 features in total:** L16_4_7 L12_3_9 L44_11_7 L12_3_6
Second		**4 features in total:** L16_4_7 L12_3_9 L12_3_6 L40_10_6
Third		**4 features in total:** L16_4_7 L12_3_9 L16_4_6 L36_9_6
Union features	L16_4_6', 'L16_4_7', 'L44_11_7', 'L12_3_9', 'L12_3_6', 'L36_9_6', 'L40_10_6'	

Table 50

First test files	Pred. / True	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
C1AuMAKU3O81000.tif	AuMAK	26.80%	0.02%	30.03%	0.00%	0.00%	0.00%	0.00%	0.00%	9.04%	5.08%	0.31%	2.74%	25.99%
A3AuOLDU3O81000.tif	AuOLD	0.00%	24.72%	0.00%	60.96%	1.07%	0.00%	6.47%	6.42%	0.02%	0.00%	0.29%	0.05%	0.00%
C1AuQUEU3O81000.tif	AuQUE	9.98%	0.19%	49.18%	0.00%	0.00%	0.00%	0.00%	0.00%	21.42%	0.05%	5.41%	10.13%	3.61%
A2CaKELU3O81000.tif	CaKEL	0.46%	80.07%	0.35%	12.30%	0.64%	0.00%	0.00%	0.09%	1.68%	0.00%	3.33%	1.08%	0.00%
A5ChHEYUO2_1000.tif	ChHEY	0.00%	0.04%	0.00%	0.44%	87.87%	11.48%	0.00%	0.00%	0.00%	0.00%	0.17%	0.00%	0.00%
A2RuTECMix_1000.tif	RuTEC	0.03%	0.00%	0.03%	0.00%	3.09%	96.26%	0.00%	0.00%	0.01%	0.37%	0.05%	0.00%	0.17%
C1SaNUFU3O81000.tif	SaNUF	0.00%	0.02%	0.00%	0.30%	0.00%	0.00%	59.73%	39.91%	0.00%	0.00%	0.03%	0.01%	0.00%
C4SaPALU3O81000.tif	SaPAL	0.01%	0.24%	0.12%	0.01%	0.00%	0.00%	0.35%	94.96%	0.53%	0.00%	0.18%	3.59%	0.00%
C3SaROSU3O81000.tif	SaROS	12.85%	0.44%	29.35%	0.00%	0.00%	0.00%	0.00%	0.04%	24.99%	0.00%	1.70%	30.33%	0.28%
A3UsATLU3O81000.tif	UsATL	0.09%	0.01%	0.06%	0.00%	0.16%	69.59%	0.00%	0.00%	0.03%	29.93%	0.07%	0.00%	0.05%
C2UsFAPU3O81000.tif	UsFAP	0.77%	38.44%	3.44%	13.41%	0.06%	0.00%	0.09%	1.61%	9.59%	0.00%	25.15%	7.42%	0.00%
A1UsPETMix_1000.tif	UsPET	0.00%	0.88%	0.00%	36.04%	0.01%	0.00%	57.16%	5.37%	0.01%	0.00%	0.52%	0.01%	0.00%
A4YuSPBUH__1000.tif	YuSPB	1.46%	0.00%	1.39%	0.00%	0.00%	0.02%	0.00%	0.00%	0.18%	0.70%	0.02%	0.10%	96.14%

Second test files	Pred. / True	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
A4AuMAKU3O81000.tif	AuMAK	22.41%	0.38%	14.44%	0.02%	0.00%	0.00%	0.00%	0.25%	24.88%	0.00%	1.01%	36.53%	0.08%
B5AuOLDU3O81000.tif	AuOLD	0.27%	95.11%	0.07%	2.73%	0.07%	0.00%	0.00%	0.01%	0.80%	0.00%	0.31%	0.63%	0.00%
A3AuQUEU3O81000.tif	AuQUE	5.22%	0.04%	41.87%	0.13%	0.00%	0.00%	0.00%	0.10%	22.80%	0.01%	14.86%	14.29%	0.69%
B4CaKELU3O81000.tif	CaKEL	0.01%	9.57%	0.08%	79.54%	0.75%	0.00%	0.52%	0.03%	0.30%	0.00%	9.00%	0.20%	0.00%
B3ChHEYUO2_1000.tif	ChHEY	0.00%	0.07%	0.00%	0.96%	80.08%	18.76%	0.00%	0.00%	0.00%	0.00%	0.12%	0.00%	0.00%
C5RuTECMix_1000.tif	RuTEC	0.00%	0.00%	0.00%	0.00%	6.47%	93.49%	0.00%	0.00%	0.00%	0.03%	0.00%	0.00%	0.00%
B2SaNUFU3O81000.tif	SaNUF	0.00%	0.00%	0.00%	1.11%	0.00%	0.00%	98.48%	0.30%	0.00%	0.00%	0.11%	0.00%	0.00%
A5SaPALU3O81000.tif	SaPAL	0.00%	0.00%	0.00%	0.10%	0.00%	0.00%	53.73%	45.89%	0.01%	0.00%	0.13%	0.14%	0.00%
B3SaROSU3O81000.tif	SaROS	4.91%	6.45%	14.57%	4.52%	0.01%	0.00%	0.00%	0.10%	30.40%	0.00%	26.04%	13.00%	0.01%
B4UsATLU3O81000.tif	UsATL	4.46%	0.00%	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.03%	93.17%	0.00%	0.00%	2.25%
B5UsFAPU3O81000.tif	UsFAP	1.77%	0.17%	25.28%	1.16%	0.00%	0.00%	0.02%	0.54%	18.77%	0.00%	35.54%	16.64%	0.11%
C3UsPETMix_1000.tif	UsPET	19.55%	0.96%	19.45%	0.11%	0.00%	0.00%	0.00%	0.06%	36.58%	0.01%	3.35%	19.91%	0.03%
B4YuSPBUH__1000.tif	YuSPB	0.84%	0.00%	1.39%	0.00%	0.00%	0.01%	0.00%	0.00%	0.12%	1.19%	0.04%	0.03%	96.39%

Third test files	Pred. / True	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
B5AuMAKU3O81000.tif	AuMAK	46.06%	0.01%	21.27%	0.00%	0.00%	0.00%	0.00%	0.00%	10.65%	19.55%	0.42%	1.76%	0.28%
C4AuOLDU3O81000.tif	AuOLD	0.20%	60.96%	0.21%	31.20%	0.14%	0.01%	0.00%	0.03%	2.11%	0.00%	4.02%	1.13%	0.00%
B1AuQUEU3O81000.tif	AuQUE	6.23%	0.01%	46.64%	0.03%	0.00%	0.00%	0.00%	0.11%	18.14%	0.03%	10.10%	16.69%	2.01%
C5CaKELU3O81000.tif	CaKEL	0.00%	0.63%	0.02%	78.08%	14.35%	0.97%	0.85%	0.03%	0.07%	0.00%	4.97%	0.04%	0.00%
C4ChHEYUO2_1000.tif	ChHEY	0.00%	0.59%	0.00%	0.93%	94.11%	4.36%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B5RuTECMix_1000.tif	RuTEC	0.00%	0.00%	0.00%	0.03%	2.69%	97.28%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
A2SaNUFU3O81000.tif	SaNUF	0.00%	0.00%	0.00%	0.31%	0.91%	0.00%	98.77%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
B3SaPALU3O81000.tif	SaPAL	0.19%	0.41%	1.45%	1.48%	0.00%	0.00%	0.31%	66.98%	2.85%	0.00%	5.58%	20.75%	0.00%
A4SaROSU3O81000.tif	SaROS	89.82%	0.35%	2.36%	0.00%	0.00%	0.00%	0.00%	0.00%	3.80%	0.28%	0.03%	3.35%	0.01%
C2UsATLU3O81000.tif	UsATL	20.12%	0.00%	0.59%	0.00%	0.00%	0.00%	0.00%	0.00%	0.26%	78.88%	0.00%	0.05%	0.10%
A3UsFAPU3O81000.tif	UsFAP	0.11%	0.25%	4.13%	19.53%	0.44%	0.72%	0.01%	0.07%	4.26%	0.00%	68.72%	1.71%	0.04%
B2UsPETMix_1000.tif	UsPET	31.21%	0.04%	21.74%	0.00%	0.00%	0.00%	0.00%	0.11%	15.70%	0.01%	0.90%	30.14%	0.16%
C4YuSPBUH__1000.tif	YuSPB	7.48%	0.04%	13.86%	0.02%	0.00%	0.01%	0.00%	0.08%	3.36%	0.04%	1.32%	11.80%	61.99%

Table 51

Average assigned probabilities

True \ Pred.	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
AuMAK	31.76%	0.14%	21.91%	0.01%	0.00%	0.00%	0.00%	0.08%	14.86%	8.21%	0.58%	13.68%	8.78%
AuOLD	0.16%	60.26%	0.09%	31.63%	0.43%	0.00%	2.16%	2.15%	0.98%	0.00%	1.54%	0.60%	0.00%
AuQUE	7.14%	0.08%	45.90%	0.06%	0.00%	0.00%	0.00%	0.07%	20.79%	0.03%	10.13%	13.70%	2.10%
CaKEL	0.15%	30.09%	0.15%	56.64%	5.25%	0.33%	0.46%	0.05%	0.68%	0.00%	5.77%	0.44%	0.00%
ChHEY	0.00%	0.23%	0.00%	0.78%	87.35%	11.53%	0.00%	0.00%	0.00%	0.00%	0.10%	0.00%	0.00%
RuTEC	0.01%	0.00%	0.01%	0.01%	4.08%	95.68%	0.00%	0.00%	0.00%	0.13%	0.02%	0.00%	0.06%
SaNUF	0.00%	0.01%	0.00%	0.57%	0.30%	0.00%	85.66%	13.41%	0.00%	0.00%	0.04%	0.00%	0.00%
SaPAL	0.07%	0.22%	0.52%	0.53%	0.00%	0.00%	18.13%	69.28%	1.13%	0.00%	1.96%	8.16%	0.00%
SaROS	35.86%	2.41%	15.43%	1.51%	0.00%	0.00%	0.00%	0.05%	19.73%	0.09%	9.25%	15.56%	0.10%
UsATL	8.22%	0.01%	0.25%	0.00%	0.05%	23.20%	0.00%	0.00%	0.10%	67.33%	0.02%	0.02%	0.80%
UsFAP	0.88%	12.95%	10.95%	11.37%	0.17%	0.24%	0.04%	0.74%	10.87%	0.00%	43.14%	8.59%	0.05%
UsPET	16.92%	0.63%	13.73%	12.05%	0.00%	0.00%	19.05%	1.85%	17.43%	0.01%	1.59%	16.69%	0.06%
YuSPB	3.26%	0.01%	5.55%	0.01%	0.00%	0.01%	0.00%	0.03%	1.22%	0.64%	0.46%	3.98%	84.84%

SD of assigned probabilities

True \ Pred.	AuMAK	AuOLD	AuQUE	CaKEL	ChHEY	RuTEC	SaNUF	SaPAL	SaROS	UsATL	UsFAP	UsPET	YuSPB
AuMAK	10.27%	0.17%	6.38%	0.01%	0.00%	0.00%	0.00%	0.12%	7.12%	8.28%	0.31%	16.17%	12.17%
AuOLD	0.11%	28.74%	0.09%	23.78%	0.46%	0.00%	3.05%	3.02%	0.86%	0.00%	1.75%	0.44%	0.00%
AuQUE	2.05%	0.08%	3.03%	0.05%	0.00%	0.00%	0.00%	0.05%	1.95%	0.02%	3.86%	2.71%	1.20%
CaKEL	0.21%	35.53%	0.14%	31.36%	6.44%	0.46%	0.35%	0.03%	0.71%	0.00%	2.39%	0.46%	0.00%
ChHEY	0.00%	0.25%	0.00%	0.24%	5.74%	5.88%	0.00%	0.00%	0.00%	0.00%	0.07%	0.00%	0.00%
RuTEC	0.01%	0.00%	0.01%	0.01%	1.69%	1.60%	0.00%	0.00%	0.00%	0.17%	0.02%	0.00%	0.08%
SaNUF	0.00%	0.01%	0.00%	0.38%	0.43%	0.00%	18.33%	18.74%	0.00%	0.00%	0.05%	0.00%	0.00%
SaPAL	0.09%	0.17%	0.65%	0.67%	0.00%	0.00%	25.17%	20.10%	1.23%	0.00%	2.56%	9.01%	0.00%
SaROS	38.30%	2.86%	11.04%	2.13%	0.00%	0.00%	0.00%	0.04%	11.48%	0.13%	11.89%	11.16%	0.13%
UsATL	8.60%	0.01%	0.24%	0.00%	0.08%	32.81%	0.00%	0.00%	0.11%	27.08%	0.03%	0.02%	1.02%
UsFAP	0.69%	18.02%	10.14%	7.64%	0.19%	0.34%	0.03%	0.65%	5.99%	0.00%	18.58%	6.15%	0.05%
UsPET	12.88%	0.42%	9.75%	16.97%	0.00%	0.00%	26.95%	2.49%	14.98%	0.00%	1.25%	12.51%	0.07%
YuSPB	2.99%	0.02%	5.88%	0.01%	0.00%	0.00%	0.00%	0.04%	1.51%	0.47%	0.61%	5.53%	16.16%

Colour class 3

100x

Table 52

Which run	Validation curves	Features selected
First		**5 features in total:** L16_4_4 L20_5_8 L8_2_2 L4_1_2 L24_6_23
Second		**4 features in total:** L8_2_3 L16_4_4 L16_4_3 L4_1_6
Third		**5 features in total:** L16_4_3 L40_10_6 L12_3_5 L8_2_2 first_order_InterquartileRange
Union features	L4_1_6, L20_5_8, L16_4_4, L8_2_3, L16_4_3, L40_10_6, first_order_InterquartileRange, L8_2_2, L4_1_2, L12_3_5, L24_6_23	

Table 53

filenames	True \ Pred.	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
C100Anacon-ADU_005.tif	Anacon	81.61%	0.00%	0.00%	15.69%	0.05%	0.05%	0.00%	0.00%	0.06%	0.00%	0.00%	0.00%	2.55%
C100BeCong-Hyd_001.tif	BeCong	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A100CaDyno-Hyd_002.tif	CaDyno	0.00%	0.00%	99.99%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B100Cotter-ADU_003.tif	Cotter	19.55%	0.00%	0.00%	80.15%	0.00%	0.29%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%
C100Fallsc-ADU_003.tif	Fallsc	0.00%	0.00%	0.00%	0.00%	59.02%	0.00%	0.00%	0.00%	39.26%	0.06%	0.00%	0.45%	1.21%
E100Nucleb-ADU_002.tif	Nucleb	0.00%	0.00%	0.00%	0.01%	0.00%	99.96%	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A100RadiHi-ADU_002.tif	RadiHi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.86%	0.13%	0.00%	0.00%	0.00%	0.00%	0.00%
A100RumJun-ADU_003.tif	RumJun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	6.94%	93.06%	0.00%	0.00%	0.00%	0.00%	0.00%
B100SpisYe-ADU_003.tif	SpisYe	0.00%	0.00%	0.00%	0.00%	18.87%	0.00%	0.00%	0.00%	68.00%	0.04%	0.00%	5.88%	7.20%
A100USAESI-ADU_003.tif	USAESI	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.99%	0.00%	0.01%	0.00%
A100Wismut-ADU_003.tif	Wismut	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.14%	0.00%	99.83%	0.02%	0.01%
C100Yeelir-NTYC_005.tif	Yeelir	0.00%	0.00%	0.16%	0.00%	4.80%	0.00%	0.00%	0.00%	19.32%	0.05%	0.00%	75.43%	0.24%
D100macass-XXX_004.tif	macass	4.41%	0.00%	0.00%	55.13%	0.23%	0.22%	0.00%	0.00%	2.25%	0.00%	0.00%	0.00%	37.75%

filenames	True \ Pred.	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
A100Anacon-ADU_005.tif	Anacon	90.00%	0.00%	0.00%	8.63%	0.00%	0.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.03%
B100BeCong-Hyd_004.tif	BeCong	0.00%	39.26%	0.00%	0.00%	0.00%	0.00%	0.20%	60.54%	0.00%	0.00%	0.00%	0.00%	0.00%
B100CaDyno-Hyd_001.tif	CaDyno	0.00%	0.00%	79.21%	0.00%	0.09%	0.00%	0.00%	0.00%	0.05%	0.00%	0.00%	20.65%	0.00%
C100Cotter-ADU_004.tif	Cotter	13.65%	0.00%	0.00%	86.05%	0.00%	0.30%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B100Fallsc-ADU_002.tif	Fallsc	0.00%	0.00%	0.02%	0.00%	71.45%	0.00%	0.00%	0.00%	18.65%	2.14%	0.06%	2.97%	4.71%
D100Nucleb-ADU_003.tif	Nucleb	1.17%	0.00%	0.00%	1.45%	0.00%	97.38%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C100RadiHi-ADU_005.tif	RadiHi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.69%	0.31%	0.00%	0.00%	0.00%	0.00%	0.00%
B100RumJun-ADU_005.tif	RumJun	0.00%	19.42%	0.00%	0.00%	0.00%	0.00%	0.03%	80.55%	0.00%	0.00%	0.00%	0.00%	0.00%
A100SpisYe-ADU_003.tif	SpisYe	0.00%	0.00%	10.88%	0.00%	20.33%	0.00%	0.00%	0.00%	25.80%	0.00%	0.10%	42.82%	0.08%
B100USAESI-ADU_002.tif	USAESI	0.00%	0.00%	0.00%	0.01%	1.32%	0.00%	0.00%	0.00%	1.27%	5.14%	0.17%	0.04%	92.05%
B100Wismut-ADU_002.tif	Wismut	0.00%	0.00%	0.00%	0.00%	0.04%	0.00%	0.00%	0.00%	0.00%	0.00%	99.95%	0.00%	0.00%
A100Yeelir-NTYC_003.tif	Yeelir	0.00%	0.00%	3.37%	0.00%	1.21%	0.00%	0.00%	0.00%	0.44%	0.09%	0.00%	94.89%	0.00%
E100macass-XXX_002.tif	macass	1.29%	0.00%	0.00%	0.41%	0.29%	0.00%	0.00%	0.00%	9.30%	0.03%	0.00%	0.01%	88.68%

filenames	True \ Pred.	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
B100Anacon-ADU_001.tif	Anacon	74.21%	0.00%	0.00%	25.79%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A100BeCong-Hyd_001.tif	BeCong	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A100CaDyno-Hyd_007.tif	CaDyno	0.00%	0.00%	98.71%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.29%	0.00%
A100Cotter-ADU_002.tif	Cotter	1.00%	0.00%	0.00%	98.89%	0.00%	0.11%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A100Fallsc-ADU_005.tif	Fallsc	0.00%	0.00%	0.00%	0.00%	58.77%	0.00%	0.00%	0.00%	40.63%	0.00%	0.18%	0.39%	0.03%
F100Nucleb-ADU_004.tif	Nucleb	0.00%	0.00%	0.00%	1.56%	0.00%	98.44%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B100RadiHi-ADU_004.tif	RadiHi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.91%	0.09%	0.00%	0.00%	0.00%	0.00%	0.00%
C100RumJun-ADU_005.tif	RumJun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	99.99%	0.00%	0.00%	0.00%	0.00%	0.00%
C100SpisYe-ADU_004.tif	SpisYe	0.00%	0.00%	0.00%	0.00%	29.83%	0.00%	0.00%	0.00%	69.80%	0.00%	0.05%	0.32%	0.00%
A100USAESI-ADU_006.tif	USAESI	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.75%	0.00%	0.00%	0.25%
C100Wismut-ADU_001.tif	Wismut	0.00%	0.00%	0.00%	0.00%	0.23%	0.00%	0.00%	0.00%	0.45%	0.00%	99.31%	0.00%	0.00%
B100Yeelir-NTYC_001.tif	Yeelir	0.00%	0.00%	0.25%	0.00%	0.62%	0.00%	0.00%	0.00%	1.49%	0.01%	0.53%	97.11%	0.00%
F100macass-XXX_001.tif	macass	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%	0.06%	0.06%	0.33%	0.00%	99.52%

Table 54

Average assigned probabilities

True \ Pred.	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
Anacon	81.94%	0.00%	0.00%	16.70%	0.02%	0.13%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%	1.19%
BeCong	0.00%	79.75%	0.00%	0.00%	0.00%	0.00%	0.07%	20.18%	0.00%	0.00%	0.00%	0.00%	0.00%
CaDyno	0.00%	0.00%	92.64%	0.00%	0.03%	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%	7.32%	0.00%
Cotter	11.40%	0.00%	0.00%	88.36%	0.00%	0.24%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Fallsc	0.00%	0.00%	0.01%	0.00%	63.08%	0.00%	0.00%	0.00%	32.85%	0.73%	0.08%	1.27%	1.98%
Nucleb	0.39%	0.00%	0.00%	1.00%	0.00%	98.60%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RadiHi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.82%	0.18%	0.00%	0.00%	0.00%	0.00%	0.00%
RumJun	0.00%	6.47%	0.00%	0.00%	0.00%	0.00%	2.33%	91.20%	0.00%	0.00%	0.00%	0.00%	0.00%
SpisYe	0.00%	0.00%	3.63%	0.00%	23.01%	0.00%	0.00%	0.00%	54.53%	0.01%	0.05%	16.34%	2.43%
USAESI	0.00%	0.00%	0.00%	0.00%	0.44%	0.00%	0.00%	0.00%	0.42%	68.29%	0.06%	0.02%	30.77%
Wismut	0.00%	0.00%	0.00%	0.00%	0.09%	0.00%	0.00%	0.00%	0.20%	0.00%	99.70%	0.01%	0.00%
Yeelir	0.00%	0.00%	1.26%	0.00%	2.21%	0.00%	0.00%	0.00%	7.09%	0.05%	0.18%	89.14%	0.08%
macass	1.90%	0.00%	0.00%	18.51%	0.18%	0.07%	0.00%	0.00%	3.87%	0.03%	0.11%	0.00%	75.32%

SD of assigned probabilities

True \ Pred.	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
Anacon	6.45%	0.00%	0.00%	7.04%	0.02%	0.15%	0.00%	0.00%	0.03%	0.00%	0.00%	0.00%	1.05%
BeCong	0.00%	28.63%	0.00%	0.00%	0.00%	0.00%	0.09%	28.54%	0.00%	0.00%	0.00%	0.00%	0.00%
CaDyno	0.00%	0.00%	9.51%	0.00%	0.04%	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%	9.45%	0.00%
Cotter	7.74%	0.00%	0.00%	7.82%	0.00%	0.09%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Fallsc	0.00%	0.00%	0.01%	0.00%	5.92%	0.00%	0.00%	0.00%	10.05%	0.99%	0.07%	1.20%	1.99%
Nucleb	0.55%	0.00%	0.00%	0.71%	0.00%	1.06%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RadiHi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%
RumJun	0.00%	9.15%	0.00%	0.00%	0.00%	0.00%	3.26%	8.04%	0.00%	0.00%	0.00%	0.00%	0.00%
SpisYe	0.00%	0.00%	5.13%	0.00%	4.86%	0.00%	0.00%	0.00%	20.33%	0.02%	0.04%	18.86%	3.38%
USAESI	0.00%	0.00%	0.00%	0.00%	0.62%	0.00%	0.00%	0.00%	0.60%	44.66%	0.08%	0.02%	43.33%
Wismut	0.00%	0.00%	0.00%	0.00%	0.10%	0.00%	0.00%	0.00%	0.19%	0.00%	0.28%	0.01%	0.00%
Yeelir	0.00%	0.00%	1.49%	0.00%	1.85%	0.00%	0.00%	0.00%	8.66%	0.04%	0.25%	9.74%	0.11%
macass	1.85%	0.00%	0.00%	25.90%	0.11%	0.10%	0.00%	0.00%	3.94%	0.03%	0.16%	0.00%	26.93%

250x

Table 55

Which run	Validation curves	Features selected
First		**6 features in total:** glrlm_LongRunLowGrayLevelEmphasis glcm_ClusterProminence_d_20 L44_11_22 glszm_SmallAreaHighGrayLevelEmphasis L20_5_6 first_order_InterquartileRange
Second		**5 features in total:** glrlm_LongRunLowGrayLevelEmphasis L16_4_5 L8_2_3 glcm_InverseVariance_d_20 first_order_Minimum
Third		**7 features in total:** glrlm_LongRunLowGrayLevelEmphasis L44_11_42 first_order_InterquartileRange L12_3_4 L16_4_7 L12_3_1 glrlm_ShortRunEmphasis
Union features	L12_3_4, glszm_SmallAreaHighGrayLevelEmphasis, first_order_Minimum, L16_4_5, L44_11_42, L44_11_22, L8_2_3, L12_3_1, first_order_InterquartileRange, glcm_InverseVariance_d_20, L16_4_7, glrlm_LongRunLowGrayLevelEmphasis, glrlm_ShortRunEmphasis, L20_5_6, glcm_ClusterProminence_d_20	

Table 56

filenames	Pred. / True	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
C250Anacon-ADU_005.tif	Anacon	30.09%	0.02%	0.00%	2.05%	31.46%	0.00%	0.00%	0.00%	1.32%	34.39%	0.66%	0.00%	0.00%
C250BeCong-Hyd_001.tif	BeCong	0.00%	99.96%	0.00%	0.00%	0.00%	0.00%	0.00%	0.03%	0.00%	0.01%	0.00%	0.00%	0.00%
A250CaDyno-Hyd_002.tif	CaDyno	96.83%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.53%	0.00%	2.63%	0.00%	0.00%
B250Cotter-ADU_003.tif	Cotter	0.00%	0.00%	0.00%	99.98%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C250Fallsc-ADU_003.tif	Fallsc	0.02%	0.00%	0.00%	0.19%	84.80%	0.00%	0.00%	0.00%	0.01%	0.00%	14.97%	0.00%	0.00%
E250Nucleb-ADU_002.tif	Nucleb	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A250RadiHi-ADU_002.tif	RadiHi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A250RumJun-ADU_003.tif	RumJun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B250SpisYe-ADU_003.tif	SpisYe	10.20%	0.00%	0.00%	0.01%	0.01%	0.00%	0.00%	0.00%	82.15%	0.00%	0.08%	6.57%	0.99%
A250USAESI-ADU_003.tif	USAESI	0.00%	16.22%	0.00%	0.26%	80.41%	0.00%	0.06%	0.00%	0.01%	2.93%	0.10%	0.00%	0.00%
A250Wismut-ADU_003.tif	Wismut	0.94%	0.00%	0.00%	0.03%	2.54%	0.00%	0.00%	0.00%	0.03%	0.00%	96.47%	0.00%	0.00%
C250Yeelir-NTYC_005.tif	Yeelir	1.24%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	4.22%	0.00%	0.00%	93.64%	0.90%
D250macass-XXX_004.tif	macass	0.00%	0.00%	0.00%	1.12%	0.01%	0.00%	0.01%	0.00%	50.20%	0.00%	0.72%	0.40%	47.55%

filenames	Pred. / True	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
A250Anacon-ADU_005.tif	Anacon	99.87%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.11%	0.00%	0.00%
B250BeCong-Hyd_004.tif	BeCong	0.00%	30.61%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	69.39%	0.00%	0.00%	0.00%
B250CaDyno-Hyd_001.tif	CaDyno	0.00%	0.00%	99.98%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%
C250Cotter-ADU_004.tif	Cotter	0.06%	0.00%	0.00%	2.46%	0.01%	0.00%	4.23%	0.00%	19.50%	0.00%	3.22%	0.59%	69.94%
B250Fallsc-ADU_002.tif	Fallsc	0.00%	0.00%	0.00%	0.00%	99.38%	0.00%	0.00%	0.00%	0.17%	0.00%	0.45%	0.00%	0.00%
D250Nucleb-ADU_003.tif	Nucleb	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C250RadiHi-ADU_005.tif	RadiHi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B250RumJun-ADU_005.tif	RumJun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A250SpisYe-ADU_003.tif	SpisYe	0.04%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	63.82%	0.00%	4.14%	27.22%	4.77%
B250USAESI-ADU_002.tif	USAESI	0.00%	0.00%	0.00%	0.01%	11.73%	0.00%	0.00%	0.00%	6.17%	0.00%	82.07%	0.01%	0.00%
B250Wismut-ADU_002.tif	Wismut	0.00%	0.00%	0.00%	0.00%	3.76%	0.00%	0.00%	0.00%	0.44%	0.00%	95.79%	0.00%	0.00%
A250Yeelir-NTYC_003.tif	Yeelir	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	13.04%	0.00%	0.00%	72.18%	14.78%
E250macass-XXX_002.tif	macass	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.83%	0.00%	0.00%	13.62%	85.55%

filenames	Pred. / True	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
B250Anacon-ADU_001.tif	Anacon	99.99%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A250BeCong-Hyd_001.tif	BeCong	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A250CaDyno-Hyd_007.tif	CaDyno	0.00%	0.00%	99.99%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%
A250Cotter-ADU_002.tif	Cotter	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A250Fallsc-ADU_005.tif	Fallsc	0.00%	0.00%	0.00%	0.00%	99.93%	0.00%	0.00%	0.00%	0.06%	0.01%	0.00%	0.00%	0.00%
F250Nucleb-ADU_004.tif	Nucleb	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B250RadiHi-ADU_004.tif	RadiHi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C250RumJun-ADU_005.tif	RumJun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C250SpisYe-ADU_004.tif	SpisYe	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	94.57%	0.00%	0.06%	3.74%	1.62%
A250USAESI-ADU_006.tif	USAESI	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	99.98%	0.00%	0.00%	0.00%
C250Wismut-ADU_001.tif	Wismut	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.96%	0.00%	0.00%
B250Yeelir-NTYC_001.tif	Yeelir	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	6.26%	0.00%	0.00%	85.30%	8.44%
F250macass-XXX_001.tif	macass	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	12.16%	0.00%	0.00%	48.00%	39.85%

Table 57

Average assigned probabilities

Pred. \ True	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
Anacon	76.65%	0.01%	0.00%	0.69%	10.49%	0.00%	0.00%	0.00%	0.44%	11.46%	0.26%	0.00%	0.00%
BeCong	0.00%	76.86%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	23.13%	0.00%	0.00%	0.00%
CaDyno	32.28%	0.00%	66.66%	0.00%	0.00%	0.00%	0.00%	0.00%	0.18%	0.00%	0.88%	0.01%	0.00%
Cotter	0.02%	0.00%	0.00%	67.48%	0.01%	0.00%	1.41%	0.00%	6.50%	0.00%	1.07%	0.20%	23.31%
Fallsc	0.01%	0.00%	0.00%	0.06%	94.70%	0.00%	0.00%	0.00%	0.08%	0.00%	5.14%	0.00%	0.00%
Nucleb	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RadiHi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RumJun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
SpisYe	3.41%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	80.18%	0.00%	1.42%	12.51%	2.46%
USAESI	0.00%	5.41%	0.00%	0.09%	30.72%	0.00%	0.02%	0.00%	2.06%	34.30%	27.39%	0.00%	0.00%
Wismut	0.32%	0.00%	0.00%	0.01%	2.10%	0.00%	0.00%	0.00%	0.16%	0.00%	97.41%	0.00%	0.00%
Yeelir	0.41%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	7.84%	0.00%	0.00%	83.71%	8.04%
macass	0.00%	0.00%	0.00%	0.37%	0.00%	0.00%	0.00%	0.00%	21.06%	0.00%	0.24%	20.67%	57.65%

SD of assigned probabilities

Pred. \ True	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
Anacon	32.92%	0.01%	0.00%	0.96%	14.83%	0.00%	0.00%	0.00%	0.62%	16.21%	0.29%	0.00%	0.00%
BeCong	0.00%	32.70%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	32.71%	0.00%	0.00%	0.00%
CaDyno	45.64%	0.00%	47.13%	0.00%	0.01%	0.00%	0.00%	0.00%	0.25%	0.00%	1.24%	0.01%	0.00%
Cotter	0.03%	0.00%	0.00%	45.98%	0.01%	0.00%	1.99%	0.00%	9.19%	0.00%	1.52%	0.28%	32.97%
Fallsc	0.01%	0.00%	0.00%	0.09%	7.00%	0.00%	0.00%	0.00%	0.07%	0.00%	6.95%	0.00%	0.00%
Nucleb	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RadiHi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RumJun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
SpisYe	4.80%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	12.63%	0.00%	1.92%	10.47%	1.66%
USAESI	0.00%	7.65%	0.00%	0.12%	35.46%	0.00%	0.03%	0.00%	2.91%	46.46%	38.67%	0.00%	0.00%
Wismut	0.43%	0.00%	0.00%	0.01%	1.57%	0.00%	0.00%	0.00%	0.20%	0.00%	1.83%	0.00%	0.00%
Yeelir	0.58%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	3.77%	0.00%	0.00%	8.83%	5.68%
macass	0.00%	0.00%	0.00%	0.53%	0.00%	0.00%	0.01%	0.00%	21.12%	0.00%	0.34%	20.06%	19.98%

500x

Table 58

Which run	Validation curves	Features selected
First		7 features in total: glszm_HighGrayLevelZoneEmphasis L44_11_26 L40_10_40 L8_2_4 L44_11_22 L28_7_6 L28_7_4
Second		6 features in total: glszm_SmallAreaHighGrayLevelEmphasis L44_11_26 L28_7_4 first_order_MeanAbsoluteDeviation L44_11_44 glcm_JointEntropy_d_1
Third		7 features in total: L44_11_26 L40_10_40 L32_8_9 L8_2_4 glszm_HighGrayLevelZoneEmphasis L40_10_20 glrlm_RunLengthNonUniformity
Union features	glszm_HighGrayLevelZoneEmphasis, L44_11_26, glszm_SmallAreaHighGrayLevelEmphasis, L32_8_9, L44_11_22, L28_7_4, glcm_JointEntropy_d_1, glrlm_RunLengthNonUniformity, L40_10_20, first_order_MeanAbsoluteDeviation, L44_11_44, L28_7_6, L8_2_4, L40_10_40	

Table 59

filenames	True \ Pred.	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
C500Anacon-ADU_005.tif	Anacon	94.13%	0.00%	0.00%	0.51%	0.00%	0.00%	1.55%	0.00%	0.00%	0.00%	0.63%	3.18%	0.00%
C500BeCong-Hyd_001.tif	BeCong	0.00%	99.99%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A500CaDyno-Hyd_002.tif	CaDyno	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B500Cotter-ADU_003.tif	Cotter	0.00%	0.00%	0.00%	99.95%	0.00%	0.00%	0.05%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C500Fallsc-ADU_003.tif	Fallsc	0.00%	0.00%	0.00%	0.00%	99.37%	0.00%	0.00%	0.00%	0.00%	0.13%	0.48%	0.00%	0.01%
E500Nucleb-ADU_002.tif	Nucleb	0.00%	0.00%	0.00%	0.00%	0.00%	100.0%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A500RadiHi-ADU_002.tif	RadiHi	0.00%	0.00%	0.00%	0.24%	0.00%	0.00%	99.76%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A500RumJun-ADU_003.tif	RumJun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.97%	0.00%	0.03%	0.00%	0.00%	0.00%
B500SpisYe-ADU_003.tif	SpisYe	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
A500USAESI-ADU_003.tif	USAESI	0.00%	0.00%	0.00%	0.00%	0.40%	0.00%	0.19%	22.23%	0.00%	28.44%	48.74%	0.00%	0.00%
A500Wismut-ADU_003.tif	Wismut	0.00%	0.00%	0.00%	0.00%	1.24%	0.00%	0.00%	0.00%	0.00%	0.00%	96.62%	2.14%	0.00%
C500Yeelir-NTYC_005.tif	Yeelir	0.00%	0.00%	0.00%	0.00%	0.03%	0.00%	0.00%	0.00%	0.00%	0.01%	1.99%	97.94%	0.03%
D500macass-XXX_004.tif	macass	0.00%	0.00%	0.00%	0.00%	10.18%	0.00%	0.00%	0.00%	0.00%	3.61%	1.51%	63.74%	20.96%

filenames	True \ Pred.	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
A500Anacon-ADU_005.tif	Anacon	99.98%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%
B500BeCong-Hyd_004.tif	BeCong	75.42%	24.36%	0.00%	0.00%	0.00%	0.00%	0.20%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%
B500CaDyno-Hyd_001.tif	CaDyno	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C500Cotter-ADU_004.tif	Cotter	0.01%	0.00%	0.00%	0.57%	0.00%	0.00%	3.66%	0.03%	0.00%	0.12%	95.59%	0.02%	0.00%
B500Fallsc-ADU_002.tif	Fallsc	0.00%	0.00%	0.00%	0.00%	90.83%	0.00%	0.00%	0.00%	0.00%	0.18%	0.00%	0.03%	8.96%
D500Nucleb-ADU_003.tif	Nucleb	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C500RadiHi-ADU_005.tif	RadiHi	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	99.96%	0.00%	0.00%	0.00%	0.03%	0.00%	0.00%
B500RumJun-ADU_005.tif	RumJun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A500SpisYe-ADU_003.tif	SpisYe	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
B500USAESI-ADU_002.tif	USAESI	0.00%	0.00%	0.00%	0.00%	0.31%	0.00%	0.00%	0.00%	0.00%	1.39%	94.52%	3.07%	0.72%
B500Wismut-ADU_002.tif	Wismut	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.99%	0.01%	0.00%
A500Yeelir-NTYC_003.tif	Yeelir	0.00%	0.00%	0.00%	0.00%	1.82%	0.00%	0.00%	0.00%	0.00%	0.03%	0.00%	34.58%	63.57%
E500macass-XXX_002.tif	macass	0.00%	0.00%	0.00%	0.00%	56.45%	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	1.26%	42.27%

filenames	True \ Pred.	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
B500Anacon-ADU_001.tif	Anacon	99.30%	0.04%	0.00%	0.01%	0.00%	0.00%	0.62%	0.00%	0.00%	0.00%	0.01%	0.02%	0.00%
A500BeCong-Hyd_001.tif	BeCong	0.00%	0.97%	0.00%	0.00%	0.00%	0.00%	0.65%	0.00%	0.00%	98.23%	0.02%	0.12%	0.01%
A500CaDyno-Hyd_007.tif	CaDyno	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A500Cotter-ADU_002.tif	Cotter	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
A500Fallsc-ADU_005.tif	Fallsc	0.00%	0.00%	0.00%	0.00%	98.87%	0.00%	0.00%	0.00%	0.00%	0.05%	0.07%	0.02%	0.99%
F500Nucleb-ADU_004.tif	Nucleb	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
B500RadiHi-ADU_004.tif	RadiHi	0.00%	0.84%	0.00%	0.00%	0.00%	0.00%	98.83%	0.00%	0.00%	0.00%	0.32%	0.00%	0.00%
C500RumJun-ADU_005.tif	RumJun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C500SpisYe-ADU_004.tif	SpisYe	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
A500USAESI-ADU_006.tif	USAESI	0.00%	0.00%	0.00%	0.00%	1.11%	0.00%	0.00%	0.00%	0.00%	93.42%	5.46%	0.01%	0.00%
C500Wismut-ADU_001.tif	Wismut	0.00%	0.00%	0.00%	0.00%	0.07%	0.00%	0.00%	0.00%	0.00%	0.00%	99.16%	0.78%	0.00%
B500Yeelir-NTYC_001.tif	Yeelir	22.86%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	77.14%	0.00%
F500macass-XXX_001.tif	macass	0.00%	0.00%	0.00%	0.00%	0.07%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.05%	98.88%

Table 60

Average assigned probabilities

True \ Pred.	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
Anacon	97.80%	0.01%	0.00%	0.17%	0.00%	0.00%	0.73%	0.00%	0.00%	0.00%	0.21%	1.07%	0.00%
BeCong	25.14%	41.78%	0.00%	0.00%	0.00%	0.00%	0.29%	0.00%	0.00%	32.74%	0.01%	0.04%	0.00%
CaDyno	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Cotter	0.00%	0.00%	0.00%	66.84%	0.00%	0.00%	1.24%	0.01%	0.00%	0.04%	31.86%	0.01%	0.00%
Fallsc	0.00%	0.00%	0.00%	0.00%	96.36%	0.00%	0.00%	0.00%	0.00%	0.12%	0.18%	0.02%	3.32%
Nucleb	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RadiHi	0.00%	0.28%	0.00%	0.08%	0.00%	0.00%	99.52%	0.00%	0.00%	0.00%	0.12%	0.00%	0.00%
RumJun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.99%	0.00%	0.01%	0.00%	0.00%	0.00%
SpisYe	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
USAESI	0.00%	0.00%	0.00%	0.00%	0.60%	0.00%	0.06%	7.41%	0.00%	41.08%	49.57%	1.03%	0.24%
Wismut	0.00%	0.00%	0.00%	0.00%	0.43%	0.00%	0.00%	0.00%	0.00%	0.00%	98.59%	0.98%	0.00%
Yeelir	7.62%	0.00%	0.00%	0.00%	0.61%	0.00%	0.00%	0.00%	0.00%	0.01%	0.66%	69.89%	21.20%
macass	0.00%	0.00%	0.00%	0.00%	22.23%	0.00%	0.00%	0.00%	0.00%	1.21%	0.50%	22.01%	54.04%

SD of assigned probabilities

True \ Pred.	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
Anacon	2.61%	0.02%	0.00%	0.24%	0.00%	0.00%	0.63%	0.00%	0.00%	0.00%	0.29%	1.49%	0.00%
BeCong	35.55%	42.26%	0.00%	0.00%	0.00%	0.00%	0.27%	0.00%	0.00%	46.31%	0.01%	0.05%	0.00%
CaDyno	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Cotter	0.00%	0.00%	0.00%	46.86%	0.00%	0.00%	1.72%	0.01%	0.00%	0.06%	45.06%	0.01%	0.00%
Fallsc	0.00%	0.00%	0.00%	0.00%	3.92%	0.00%	0.00%	0.00%	0.00%	0.05%	0.21%	0.01%	4.01%
Nucleb	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
RadiHi	0.00%	0.40%	0.00%	0.11%	0.00%	0.00%	0.49%	0.00%	0.00%	0.00%	0.14%	0.00%	0.00%
RumJun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	0.02%	0.00%	0.00%	0.00%
SpisYe	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
USAESI	0.00%	0.00%	0.00%	0.00%	0.36%	0.00%	0.09%	10.48%	0.00%	38.62%	36.36%	1.45%	0.34%
Wismut	0.00%	0.00%	0.00%	0.00%	0.57%	0.00%	0.00%	0.00%	0.00%	0.00%	1.43%	0.88%	0.00%
Yeelir	10.78%	0.00%	0.00%	0.00%	0.85%	0.00%	0.00%	0.00%	0.00%	0.01%	0.94%	26.37%	29.96%
macass	0.00%	0.00%	0.00%	0.00%	24.54%	0.00%	0.00%	0.00%	0.00%	1.70%	0.71%	29.50%	32.88%

All magnifications

Table 61

Which run	Validation curves	Features selected
First		**5 features in total:** L44_11_22_100x glcm_ClusterProminence_d_20_250x glszm_SmallAreaHighGrayLevelEmphasis_250x L44_11_27_100x L32_8_14_100x
Second		**4 features in total:** L12_3_4_100x L8_2_5_100x glrlm_LongRunLowGrayLevelEmphasis_250x L32_8_5_100x
Third		**4 features in total:** glrlm_LongRunLowGrayLevelEmphasis_250x L16_4_5_100x first_order_90Percentile_100x L16_4_17_100x
Union features	L32_8_14_100x, L12_3_4_100x, L44_11_27_100x, first_order_90Percentile_100x, L8_2_5_100x, glcm_ClusterProminence_d_20_250x, L44_11_22_100x, L16_4_5_100x, L16_4_17_100x, glrlm_LongRunLowGrayLevelEmphasis_250x, L32_8_5_100x, glszm_SmallAreaHighGrayLevelEmphasis_250x	

Table 62

Pred. / True	anacon	becong	cadyno	cotter	fallsc	macass	nucleb	radihi	rumjun	spisye	usaesi	wismut	yeelir
anacon	49.72%	49.11%	0.00%	0.79%	0.10%	0.21%	0.00%	0.03%	0.00%	0.01%	0.00%	0.02%	0.00%
becong	2.72%	24.68%	0.00%	0.14%	0.11%	18.07%	0.00%	0.70%	53.14%	0.01%	0.00%	0.14%	0.29%
cadyno	0.10%	0.00%	0.32%	0.02%	94.81%	0.21%	0.00%	0.00%	0.00%	1.12%	0.00%	0.01%	3.40%
cotter	0.12%	0.01%	0.00%	7.54%	0.09%	0.01%	90.01%	0.07%	0.00%	1.25%	0.00%	0.88%	0.00%
fallsc	2.09%	0.01%	0.00%	4.34%	52.88%	16.88%	0.13%	0.83%	0.00%	10.51%	0.00%	3.91%	8.40%
macass	0.00%	0.00%	0.00%	5.47%	0.13%	1.39%	18.70%	13.46%	0.00%	7.75%	0.00%	52.99%	0.09%
nucleb	0.02%	0.01%	0.00%	4.35%	0.02%	0.07%	86.41%	1.93%	0.00%	1.14%	0.00%	6.06%	0.00%
radihi	0.14%	0.01%	0.00%	7.33%	0.16%	7.44%	0.37%	55.50%	1.08%	0.47%	0.00%	27.38%	0.13%
rumjun	0.01%	0.00%	0.00%	0.02%	0.00%	0.98%	0.00%	1.84%	97.02%	0.00%	0.00%	0.11%	0.02%
spisye	0.03%	0.01%	0.00%	3.15%	5.29%	12.47%	1.37%	2.88%	0.00%	43.82%	0.00%	24.04%	6.94%
usaesi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
wismut	0.11%	0.01%	0.00%	12.71%	1.34%	8.55%	2.99%	17.78%	0.02%	8.78%	0.00%	47.13%	0.59%
yeelir	0.09%	0.00%	0.00%	1.16%	33.99%	12.28%	0.16%	0.50%	0.00%	21.82%	0.00%	4.36%	25.64%

Pred. / True	anacon	becong	cadyno	cotter	fallsc	macass	nucleb	radihi	rumjun	spisye	usaesi	wismut	yeelir
anacon	86.14%	0.00%	0.00%	13.83%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%
becong	0.39%	99.09%	0.00%	0.31%	0.00%	0.00%	0.00%	0.00%	0.20%	0.00%	0.01%	0.00%	0.00%
cadyno	0.00%	0.00%	13.13%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.08%	0.00%	0.00%	86.79%
cotter	0.68%	0.00%	0.00%	99.31%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
fallsc	0.00%	0.00%	0.00%	0.00%	98.59%	0.00%	0.00%	0.00%	0.00%	1.09%	0.04%	0.28%	0.00%
macass	0.00%	0.00%	0.00%	0.00%	0.00%	99.99%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%
nucleb	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
radihi	0.00%	0.32%	0.00%	0.00%	0.00%	0.00%	0.00%	66.10%	33.58%	0.00%	0.00%	0.00%	0.00%
rumjun	0.00%	20.43%	0.00%	0.00%	0.00%	0.00%	0.00%	14.50%	65.07%	0.00%	0.00%	0.00%	0.00%
spisye	0.00%	0.00%	5.75%	0.00%	4.35%	0.00%	0.00%	0.00%	0.00%	69.83%	0.10%	0.01%	19.97%
usaesi	0.11%	0.00%	0.00%	0.00%	0.68%	0.04%	0.00%	0.00%	0.00%	0.07%	98.97%	0.13%	0.00%
wismut	0.00%	0.00%	0.00%	0.00%	0.73%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.27%	0.00%
yeelir	0.00%	0.00%	4.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.72%	0.00%	0.00%	95.27%

Pred. / True	anacon	becong	cadyno	cotter	fallsc	macass	nucleb	radihi	rumjun	spisye	usaesi	wismut	yeelir
anacon	84.39%	0.05%	0.00%	15.56%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
becong	0.05%	97.46%	0.00%	2.11%	0.00%	0.00%	0.00%	0.00%	0.38%	0.00%	0.00%	0.00%	0.00%
cadyno	0.00%	0.00%	99.90%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%
cotter	0.96%	46.06%	0.00%	52.77%	0.00%	0.00%	0.00%	0.00%	0.22%	0.00%	0.00%	0.00%	0.00%
fallsc	0.00%	0.00%	0.00%	0.00%	94.51%	0.00%	0.00%	0.00%	0.00%	3.33%	0.52%	1.62%	0.01%
macass	0.00%	0.00%	0.00%	0.00%	0.00%	99.52%	0.00%	0.00%	0.00%	0.48%	0.00%	0.00%	0.00%
nucleb	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
radihi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.34%	0.66%	0.00%	0.00%	0.00%	0.00%
rumjun	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.28%	98.71%	0.00%	0.00%	0.00%	0.00%
spisye	0.00%	0.00%	0.10%	0.00%	1.19%	0.01%	0.00%	0.00%	0.00%	76.24%	4.46%	0.01%	17.98%
usaesi	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%	0.00%	0.36%	99.61%	0.00%	0.01%
wismut	0.00%	0.00%	0.02%	0.00%	3.08%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	96.90%	0.00%
yeelir	0.00%	0.00%	13.65%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%	3.13%	0.00%	0.00%	83.20%

Table 63

Average assigned probabilities

True \ Pred.	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
Anacon	73.42%	16.39%	0.00%	10.06%	0.03%	0.07%	0.00%	0.01%	0.00%	0.00%	0.01%	0.01%	0.00%
BeCong	1.06%	73.74%	0.00%	0.85%	0.04%	6.02%	0.00%	0.23%	17.91%	0.00%	0.00%	0.05%	0.10%
CaDyno	0.03%	0.00%	37.79%	0.01%	31.60%	0.07%	0.00%	0.00%	0.00%	0.40%	0.00%	0.00%	30.09%
Cotter	0.59%	15.36%	0.00%	53.21%	0.03%	0.00%	30.00%	0.03%	0.07%	0.42%	0.00%	0.29%	0.00%
Fallsc	0.70%	0.00%	0.00%	1.45%	81.99%	5.63%	0.04%	0.28%	0.00%	4.98%	0.19%	1.94%	2.80%
Nucleb	0.00%	0.00%	0.00%	1.83%	0.04%	66.97%	6.23%	4.49%	0.00%	2.74%	0.00%	17.66%	0.03%
RadiHi	0.01%	0.00%	0.00%	1.45%	0.01%	0.02%	95.47%	0.64%	0.00%	0.38%	0.00%	2.02%	0.00%
RumJun	0.05%	0.11%	0.00%	2.44%	0.05%	2.48%	0.12%	73.64%	11.77%	0.16%	0.00%	9.13%	0.04%
SpisYe	0.00%	6.81%	0.00%	0.01%	0.00%	0.33%	0.00%	5.87%	86.93%	0.00%	0.00%	0.04%	0.01%
USAESI	0.01%	0.00%	1.95%	1.05%	3.61%	4.16%	0.46%	0.96%	0.00%	63.29%	1.52%	8.02%	14.96%
Wismut	0.04%	0.00%	0.00%	0.00%	0.23%	0.01%	0.00%	0.00%	0.00%	0.14%	99.53%	0.04%	0.00%
Yeelir	0.04%	0.00%	0.01%	4.24%	1.71%	2.85%	1.00%	5.93%	0.01%	2.93%	0.00%	81.10%	0.20%
macass	0.03%	0.00%	5.89%	0.39%	11.33%	4.10%	0.05%	0.17%	0.00%	8.56%	0.00%	1.45%	68.03%

SD of assigned probabilities

True \ Pred.	Anacon	BeCong	CaDyno	Cotter	Fallsc	Nucleb	RadiHi	RumJun	SpisYe	USAESI	Wismut	Yeelir	macass
Anacon	16.77%	23.14%	0.00%	6.59%	0.05%	0.10%	0.00%	0.02%	0.00%	0.00%	0.01%	0.01%	0.00%
BeCong	1.19%	34.70%	0.00%	0.89%	0.05%	8.52%	0.00%	0.33%	24.91%	0.01%	0.00%	0.07%	0.14%
CaDyno	0.05%	0.00%	44.23%	0.01%	44.69%	0.10%	0.00%	0.00%	0.00%	0.51%	0.00%	0.00%	40.11%
Cotter	0.35%	21.71%	0.00%	37.47%	0.04%	0.00%	42.43%	0.03%	0.10%	0.59%	0.00%	0.41%	0.00%
Fallsc	0.99%	0.01%	0.00%	2.05%	20.65%	7.96%	0.06%	0.39%	0.00%	4.02%	0.24%	1.50%	3.96%
Nucleb	0.00%	0.00%	0.00%	2.58%	0.06%	46.37%	8.82%	6.35%	0.00%	3.55%	0.00%	24.98%	0.04%
RadiHi	0.01%	0.00%	0.00%	2.05%	0.01%	0.03%	6.41%	0.91%	0.00%	0.54%	0.00%	2.86%	0.00%
RumJun	0.07%	0.15%	0.00%	3.46%	0.07%	3.51%	0.17%	18.68%	15.42%	0.22%	0.00%	12.91%	0.06%
SpisYe	0.00%	9.63%	0.00%	0.01%	0.00%	0.46%	0.00%	6.10%	15.48%	0.00%	0.00%	0.05%	0.01%
USAESI	0.01%	0.00%	2.69%	1.48%	1.75%	5.87%	0.64%	1.36%	0.00%	14.02%	2.08%	11.33%	5.73%
Wismut	0.05%	0.00%	0.00%	0.00%	0.31%	0.02%	0.00%	0.00%	0.00%	0.16%	0.43%	0.06%	0.00%
Yeelir	0.05%	0.01%	0.01%	5.99%	0.99%	4.03%	1.41%	8.38%	0.01%	4.14%	0.00%	24.04%	0.28%
macass	0.04%	0.00%	5.73%	0.54%	16.02%	5.78%	0.08%	0.24%	0.00%	9.43%	0.00%	2.05%	30.38%

Colour class 4

100x

Table 64

Which run	Validation curves	Features selected
First		5 features in total: L4_1_3 L32_8_2 L12_3_4 L8_2_2 L40_10_10
Second		5 features in total: L4_1_1 L28_7_2 L12_3_2 L16_4_6 L28_7_11
Third		5 features in total: L4_1_3 L28_7_2 L12_3_4 L44_11_44 L24_6_25
Union features	L28_7_11, L32_8_2, L24_6_25, L28_7_2, L4_1_1, L12_3_2, L40_10_10, L8_2_2, L12_3_4, L44_11_44, L16_4_6, L4_1_3	

Table 65

filenames	Pred. / True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
C100Deniso-ADU_003.tif	Deniso	99.96%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.04%
C100Ellwel-ADU_002.tif	Ellwel	0.00%	95.23%	0.00%	0.00%	0.30%	4.47%	0.00%	0.00%	0.00%
B100GabEFI-ADU_002.tif	GabEFI	0.00%	0.00%	99.71%	0.00%	0.00%	0.00%	0.00%	0.00%	0.29%
B100KMcGee-ADU_005.tif	KMcGee	0.00%	0.00%	0.00%	99.09%	0.00%	0.00%	0.00%	0.91%	0.00%
F100Millika-ADU_003.tif	Millik	0.00%	0.02%	0.00%	0.00%	98.78%	0.12%	1.08%	0.00%	0.00%
C100Ranstad-SDU_001.tif	Ransta	0.00%	0.31%	0.00%	0.00%	1.64%	96.96%	1.08%	0.00%	0.00%
C100Romani-SDU_004.tif	Romani	0.00%	1.15%	0.00%	0.00%	2.34%	74.38%	22.13%	0.00%	0.00%
C100SpaGen-ADU_002.tif	SpaGen	0.00%	0.00%	0.00%	50.71%	0.00%	0.04%	0.00%	49.24%	0.00%
A100southa-Hyd_002.tif	southa	0.00%	0.00%	10.84%	0.00%	0.00%	0.00%	0.01%	0.00%	89.14%

filenames	Pred. / True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
A100Deniso-ADU_003.tif	Deniso	92.47%	0.00%	0.00%	0.00%	0.00%	7.08%	0.45%	0.00%	0.00%
B100Ellwel-ADU_005.tif	Ellwel	0.00%	99.84%	0.00%	0.00%	0.13%	0.03%	0.00%	0.00%	0.00%
A100GabEFI-ADU_001.tif	GabEFI	0.00%	0.00%	99.98%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%
C100KMcGee-ADU_003.tif	KMcGee	0.00%	0.00%	0.00%	94.52%	0.00%	0.00%	0.00%	5.48%	0.00%
E100Millika-ADU_004.tif	Millik	0.00%	0.01%	0.00%	0.00%	97.48%	0.70%	1.81%	0.00%	0.00%
B100Ranstad-SDU_005.tif	Ransta	0.10%	0.10%	0.00%	0.00%	0.48%	98.21%	1.11%	0.00%	0.00%
B100Romani-SDU_003.tif	Romani	0.13%	0.00%	0.00%	0.00%	0.27%	0.03%	99.27%	0.00%	0.30%
B100SpaGen-ADU_003.tif	SpaGen	0.00%	0.00%	0.00%	9.45%	0.00%	0.00%	0.00%	90.55%	0.00%
B100southa-Hyd_004.tif	southa	0.04%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.96%

filenames	Pred. / True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
B100Deniso-ADU_001.tif	Deniso	99.94%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.06%
A100Ellwel-ADU_004.tif	Ellwel	0.00%	99.18%	0.00%	0.00%	0.82%	0.00%	0.00%	0.00%	0.00%
C100GabEFI-ADU_005.tif	GabEFI	0.00%	0.00%	99.93%	0.05%	0.00%	0.00%	0.00%	0.02%	0.00%
A100KMcGee-ADU_003.tif	KMcGee	0.00%	0.00%	0.03%	99.76%	0.00%	0.00%	0.00%	0.21%	0.00%
E100Millika-ADU_003.tif	Millik	0.00%	0.76%	0.00%	0.00%	99.20%	0.00%	0.04%	0.00%	0.00%
A100Ranstad-SDU_004.tif	Ransta	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%
A100Romani-SDU_003.tif	Romani	0.00%	0.00%	0.00%	0.00%	22.73%	0.00%	77.24%	0.00%	0.03%
C100SpaGen-ADU_004.tif	SpaGen	0.00%	0.00%	0.00%	3.30%	0.00%	0.00%	0.00%	96.70%	0.00%
A100southa-Hyd_004.tif	southa	0.16%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	99.84%

Table 66

Average assignes probabilities

Pred. / True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
Deniso	97.46%	0.00%	0.00%	0.00%	0.00%	2.36%	0.15%	0.00%	0.03%
Ellwel	0.00%	98.08%	0.00%	0.00%	0.42%	1.50%	0.00%	0.00%	0.00%
GabEFI	0.00%	0.00%	99.87%	0.02%	0.00%	0.00%	0.00%	0.01%	0.10%
KMcGee	0.00%	0.00%	0.01%	97.79%	0.00%	0.00%	0.00%	2.20%	0.00%
Millik	0.00%	0.26%	0.00%	0.00%	98.49%	0.27%	0.98%	0.00%	0.00%
Ransta	0.04%	0.14%	0.00%	0.00%	0.70%	98.39%	0.73%	0.00%	0.00%
Romani	0.04%	0.38%	0.00%	0.00%	8.45%	24.80%	66.21%	0.00%	0.11%
SpaGen	0.00%	0.00%	0.00%	21.15%	0.00%	0.01%	0.00%	78.83%	0.00%
southa	0.07%	0.00%	3.62%	0.00%	0.00%	0.00%	0.00%	0.00%	96.31%

SD of assigned probabilities

Pred. / True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
Deniso	3.53%	0.00%	0.00%	0.00%	0.00%	3.34%	0.21%	0.00%	0.02%
Ellwel	0.00%	2.04%	0.00%	0.00%	0.29%	2.10%	0.00%	0.00%	0.00%
GabEFI	0.00%	0.00%	0.12%	0.02%	0.00%	0.00%	0.00%	0.01%	0.13%
KMcGee	0.00%	0.00%	0.01%	2.33%	0.00%	0.00%	0.00%	2.34%	0.00%
Millik	0.00%	0.35%	0.00%	0.00%	0.73%	0.30%	0.73%	0.00%	0.00%
Ransta	0.05%	0.13%	0.00%	0.00%	0.69%	1.24%	0.52%	0.00%	0.00%
Romani	0.06%	0.54%	0.00%	0.00%	10.13%	35.06%	32.44%	0.00%	0.13%
SpaGen	0.00%	0.00%	0.00%	21.05%	0.00%	0.02%	0.00%	21.07%	0.00%
southa	0.07%	0.00%	5.11%	0.00%	0.00%	0.00%	0.01%	0.00%	5.07%

250x

Table 67

Which run	Validation curves	Features selected
First		**6 features in total:** L44_11_32 glrlm_LongRunLowGrayLevelEmphasis L4_1_2 L16_4_6 L32_8_11 glszm_ZonePercentage
Second		**5 features in total:** glrlm_LongRunLowGrayLevelEmphasis L12_3_5 L12_3_4 L4_1_3 glcm_DifferenceVariance_d_1
Third		**6 features in total:** glrlm_LongRunLowGrayLevelEmphasis L44_11_32 L12_3_5 first_order_Skewness glcm_DifferenceVariance_d_1 L4_1_3
Union features	L12_3_4, L32_8_11, glszm_ZonePercentage, first_order_Skewness, L4_1_2, glcm_DifferenceVariance_d_1, L12_3_5, glrlm_LongRunLowGrayLevelEmphasis, L16_4_6, L44_11_32, L4_1_3	

Table 68

filenames	Pred. / True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
C250Deniso-ADU_003.tif	Deniso	99.89%	0.00%	0.03%	0.07%	0.00%	0.01%	0.00%	0.00%	0.00%
C250Ellwel-ADU_002.tif	Ellwel	0.00%	91.58%	0.00%	0.00%	0.00%	0.00%	0.00%	8.42%	0.00%
B250GabEFI-ADU_002.tif	GabEFI	1.00%	0.00%	98.01%	0.99%	0.00%	0.00%	0.00%	0.00%	0.00%
B250KMcGee-ADU_005.tif	KMcGee	0.15%	0.00%	0.03%	99.79%	0.00%	0.04%	0.00%	0.00%	0.00%
F250Millika-ADU_003.tif	Millik	0.00%	0.00%	0.00%	0.00%	99.99%	0.00%	0.00%	0.00%	0.01%
C250Ranstad-SDU_001.tif	Ransta	0.12%	0.00%	0.00%	0.12%	0.00%	99.76%	0.00%	0.00%	0.00%
C250Romani-SDU_004.tif	Romani	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
C250SpaGen-ADU_002.tif	SpaGen	0.00%	11.99%	0.00%	0.00%	0.00%	0.00%	0.00%	88.01%	0.00%
A250southa-Hyd_002.tif	southa	0.00%	0.00%	0.00%	0.00%	0.05%	0.00%	0.00%	0.00%	99.95%

filenames	Pred. / True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
A250Deniso-ADU_003.tif	Deniso	99.97%	0.00%	0.00%	0.00%	0.00%	0.03%	0.00%	0.00%	0.00%
B250Ellwel-ADU_005.tif	Ellwel	0.00%	98.84%	0.00%	0.00%	0.00%	0.00%	0.00%	1.16%	0.00%
A250GabEFI-ADU_001.tif	GabEFI	0.00%	0.00%	99.85%	0.15%	0.00%	0.00%	0.00%	0.00%	0.00%
C250KMcGee-ADU_003.tif	KMcGee	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
E250Millika-ADU_004.tif	Millik	0.00%	82.35%	0.00%	0.00%	17.54%	0.06%	0.05%	0.00%	0.00%
B250Ranstad-SDU_005.tif	Ransta	0.07%	0.00%	0.00%	0.00%	0.00%	99.92%	0.00%	0.00%	0.01%
B250Romani-SDU_003.tif	Romani	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
B250SpaGen-ADU_003.tif	SpaGen	0.00%	0.07%	0.00%	0.00%	0.00%	0.00%	0.00%	99.93%	0.00%
B250southa-Hyd_004.tif	southa	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.23%	0.00%	98.77%

filenames	Pred. / True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
B250Deniso-ADU_001.tif	Deniso	99.99%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%
A250Ellwel-ADU_004.tif	Ellwel	0.00%	99.45%	0.00%	0.00%	0.00%	0.00%	0.00%	0.55%	0.00%
C250GabEFI-ADU_005.tif	GabEFI	0.00%	0.00%	99.93%	0.07%	0.00%	0.00%	0.00%	0.00%	0.00%
A250KMcGee-ADU_003.tif	KMcGee	0.00%	0.00%	0.04%	99.96%	0.00%	0.00%	0.00%	0.00%	0.00%
E250Millika-ADU_003.tif	Millik	0.06%	0.00%	0.00%	0.00%	99.88%	0.00%	0.00%	0.00%	0.06%
A250Ranstad-SDU_004.tif	Ransta	0.05%	0.00%	0.00%	0.00%	0.00%	99.95%	0.00%	0.00%	0.00%
A250Romani-SDU_003.tif	Romani	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
C250SpaGen-ADU_004.tif	SpaGen	0.00%	0.18%	0.00%	0.00%	0.00%	0.00%	0.00%	99.82%	0.00%
A250southa-Hyd_004.tif	southa	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	98.75%	0.00%	1.25%

Table 69

Average assignes probabilities

True \ Pred.	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
Deniso	99.95%	0.00%	0.01%	0.02%	0.00%	0.02%	0.00%	0.00%	0.00%
Ellwel	0.00%	96.62%	0.00%	0.00%	0.00%	0.00%	0.00%	3.38%	0.00%
GabEFI	0.33%	0.00%	99.26%	0.40%	0.00%	0.00%	0.00%	0.00%	0.00%
KMcGee	0.05%	0.00%	0.02%	99.92%	0.00%	0.01%	0.00%	0.00%	0.00%
Millik	0.02%	27.45%	0.00%	0.00%	72.47%	0.02%	0.02%	0.00%	0.02%
Ransta	0.08%	0.00%	0.00%	0.04%	0.00%	99.88%	0.00%	0.00%	0.00%
Romani	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
SpaGen	0.00%	4.08%	0.00%	0.00%	0.00%	0.00%	0.00%	95.92%	0.00%
southa	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	33.32%	0.00%	66.65%

SD of assigned probabilities

True \ Pred.	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
Deniso	0.04%	0.00%	0.01%	0.04%	0.00%	0.01%	0.00%	0.00%	0.00%
Ellwel	0.00%	3.57%	0.00%	0.00%	0.00%	0.00%	0.00%	3.57%	0.00%
GabEFI	0.47%	0.00%	0.89%	0.41%	0.00%	0.00%	0.00%	0.00%	0.00%
KMcGee	0.07%	0.00%	0.02%	0.09%	0.00%	0.02%	0.00%	0.00%	0.00%
Millik	0.03%	38.82%	0.00%	0.00%	38.84%	0.03%	0.02%	0.00%	0.02%
Ransta	0.03%	0.00%	0.00%	0.06%	0.00%	0.09%	0.00%	0.00%	0.01%
Romani	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
SpaGen	0.00%	5.59%	0.00%	0.00%	0.00%	0.00%	0.00%	5.59%	0.00%
southa	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	46.26%	0.00%	46.25%

500x

Table 70

Which run	Validation curves	Features selected
First		**7 features in total:** L4_1_3 L4_1_2 glszm_LowGrayLevelZoneEmphasis L36_9_5 L4_1_4 L28_7_6 L32_8_6
Second		**5 features in total:** L4_1_3 L4_1_2 L36_9_4 L40_10_4 glszm_LowGrayLevelZoneEmphasis
Third		**4 features in total:** L4_1_3 L4_1_2 glszm_LowGrayLevelZoneEmphasis L32_8_5
Union features	L4_1_4, glszm_LowGrayLevelZoneEmphasis, L36_9_5, L32_8_5, L36_9_4, L4_1_2, L32_8_6, L28_7_6, L40_10_4, L4_1_3	

Table 71

filenames	Pred. / True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
C500Deniso-ADU_003.tif	Deniso	77.87%	0.00%	0.00%	0.00%	0.00%	0.00%	0.38%	0.00%	21.75%
C500Ellwel-ADU_002.tif	Ellwel	0.00%	99.67%	0.22%	0.00%	0.00%	0.01%	0.00%	0.10%	0.00%
B500GabEFI-ADU_002.tif	GabEFI	0.00%	0.08%	98.54%	0.88%	0.00%	0.48%	0.00%	0.02%	0.00%
B500KMcGee-ADU_005.tif	KMcGee	0.00%	0.00%	13.80%	83.63%	0.00%	0.00%	0.00%	2.56%	0.00%
F500Millika-ADU_003.tif	Millik	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
C500Ranstad-SDU_001.tif	Ransta	0.00%	0.00%	0.00%	0.00%	0.00%	99.98%	0.00%	0.00%	0.01%
C500Romani-SDU_004.tif	Romani	50.72%	0.00%	0.00%	0.00%	0.00%	0.00%	32.79%	0.00%	16.49%
C500SpaGen-ADU_002.tif	SpaGen	0.00%	0.11%	42.42%	27.79%	0.00%	0.00%	0.00%	29.68%	0.00%
A500southa-Hyd_002.tif	southa	12.71%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	87.28%

filenames	Pred. / True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
A500Deniso-ADU_003.tif	Deniso	87.37%	0.00%	0.00%	0.00%	0.00%	0.00%	10.50%	0.00%	2.13%
B500Ellwel-ADU_005.tif	Ellwel	0.00%	99.78%	0.17%	0.00%	0.00%	0.00%	0.00%	0.05%	0.00%
A500GabEFI-ADU_001.tif	GabEFI	0.00%	0.84%	95.85%	2.02%	0.00%	1.25%	0.00%	0.04%	0.00%
C500KMcGee-ADU_003.tif	KMcGee	0.00%	0.01%	1.38%	96.67%	0.00%	0.00%	0.00%	1.93%	0.00%
E500Millika-ADU_004.tif	Millik	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
B500Ranstad-SDU_005.tif	Ransta	0.00%	0.36%	0.54%	0.00%	0.00%	99.10%	0.00%	0.00%	0.00%
B500Romani-SDU_003.tif	Romani	6.42%	0.00%	0.00%	0.00%	2.15%	0.00%	91.34%	0.00%	0.09%
B500SpaGen-ADU_003.tif	SpaGen	0.00%	1.70%	26.02%	12.31%	0.00%	0.00%	0.00%	59.96%	0.00%
B500southa-Hyd_004.tif	southa	84.36%	0.00%	0.00%	0.00%	0.00%	0.09%	6.14%	0.00%	9.41%

filenames	Pred. / True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
B500Deniso-ADU_001.tif	Deniso	32.97%	0.00%	0.00%	0.00%	0.00%	0.00%	8.34%	0.00%	58.68%
A500Ellwel-ADU_004.tif	Ellwel	0.00%	99.22%	0.34%	0.00%	0.00%	0.00%	0.00%	0.44%	0.00%
C500GabEFI-ADU_005.tif	GabEFI	0.00%	1.76%	81.28%	12.99%	0.00%	0.03%	0.00%	3.93%	0.00%
A500KMcGee-ADU_003.tif	KMcGee	0.00%	0.00%	0.19%	99.72%	0.00%	0.00%	0.00%	0.09%	0.00%
E500Millika-ADU_003.tif	Millik	0.00%	0.00%	0.00%	0.00%	99.97%	0.00%	0.03%	0.00%	0.00%
A500Ranstad-SDU_004.tif	Ransta	0.00%	0.00%	0.00%	0.00%	0.00%	99.98%	0.00%	0.00%	0.01%
A500Romani-SDU_003.tif	Romani	0.09%	0.00%	0.00%	0.00%	63.60%	0.00%	36.31%	0.00%	0.00%
C500SpaGen-ADU_004.tif	SpaGen	0.00%	0.73%	1.41%	1.34%	0.00%	0.00%	0.00%	96.53%	0.00%
A500southa-Hyd_004.tif	southa	1.80%	0.00%	0.00%	0.00%	0.00%	0.03%	0.12%	0.00%	98.05%

Table 72

Average assignes probabilities

True \ Pred.	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
Deniso	66.07%	0.00%	0.00%	0.00%	0.00%	0.00%	6.41%	0.00%	27.52%
Ellwel	0.00%	99.56%	0.24%	0.00%	0.00%	0.00%	0.00%	0.20%	0.00%
GabEFI	0.00%	0.89%	91.89%	5.30%	0.00%	0.59%	0.00%	1.33%	0.00%
KMcGee	0.00%	0.00%	5.13%	93.34%	0.00%	0.00%	0.00%	1.53%	0.00%
Millik	0.00%	0.00%	0.00%	0.00%	99.99%	0.00%	0.01%	0.00%	0.00%
Ransta	0.00%	0.12%	0.18%	0.00%	0.00%	99.69%	0.00%	0.00%	0.01%
Romani	19.08%	0.00%	0.00%	0.00%	21.92%	0.00%	53.48%	0.00%	5.53%
SpaGen	0.00%	0.85%	23.28%	13.81%	0.00%	0.00%	0.00%	62.06%	0.00%
southa	32.95%	0.00%	0.00%	0.00%	0.00%	0.04%	2.09%	0.00%	64.91%

SD of assigned probabilities

True \ Pred.	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
Deniso	23.72%	0.00%	0.00%	0.00%	0.00%	0.00%	4.35%	0.00%	23.45%
Ellwel	0.00%	0.24%	0.07%	0.00%	0.00%	0.00%	0.00%	0.17%	0.00%
GabEFI	0.00%	0.68%	7.58%	5.46%	0.00%	0.50%	0.00%	1.84%	0.00%
KMcGee	0.00%	0.00%	6.16%	6.98%	0.00%	0.00%	0.00%	1.05%	0.00%
Millik	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.01%	0.00%	0.00%
Ransta	0.00%	0.17%	0.25%	0.00%	0.00%	0.42%	0.00%	0.00%	0.01%
Romani	22.52%	0.00%	0.00%	0.00%	29.49%	0.00%	26.81%	0.00%	7.75%
SpaGen	0.00%	0.66%	16.86%	10.85%	0.00%	0.00%	0.00%	27.33%	0.00%
southa	36.62%	0.00%	0.00%	0.00%	0.00%	0.04%	2.86%	0.00%	39.49%

All magnifications

Table 73

Which run	Validation curves	Features selected
First		**8 features in total:** L4_1_3_500x L44_11_44_100x L44_11_32_250x first_order_Variance_250x L12_3_4_100x L12_3_5_100x glszm_LargeAreaHighGrayLevelEmphasis_250x glcm_ClusterTendency_d_3_100x
Second		**5 features in total:** L4_1_1_500x L44_11_32_250x L36_9_2_100x L36_9_36_100x L8_2_1_500x
Third		**4 features in total:** L12_3_4_100x glszm_GrayLevelNonUniformityNormalized_100x L40_10_40_100x L4_1_1_100x
Union features	L4_1_3_500x, L12_3_4_100x, L44_11_44_100x, L12_3_5_100x, glszm_LargeAreaHighGrayLevelEmphasis_250x, L4_1_1_500x, L44_11_32_250x, first_order_Variance_250x, L8_2_1_500x, L36_9_36_100x, L40_10_40_100x, L4_1_1_100x, glszm_GrayLevelNonUniformityNormalized_100x, glcm_ClusterTendency_d_3_100x, L36_9_2_100x	

143

Table 74

Pred. True	deniso	ellwel	gabefi	kmcgee	millik	ransta	romani	southa	spagen
deniso	**100.00%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
ellwel	0.00%	**100.00%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
gabefi	0.00%	0.00%	**100.00%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
kmcgee	0.00%	**0.16%**	0.00%	**99.83%**	0.00%	0.00%	0.00%	0.00%	0.00%
millik	0.00%	0.00%	0.00%	0.00%	**99.99%**	0.00%	**0.01%**	0.00%	0.00%
ransta	0.00%	0.00%	0.00%	0.00%	0.00%	**100.00%**	0.00%	0.00%	0.00%
romani	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	**100.00%**	0.00%	0.00%
southa	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	**100.00%**	0.00%
spagen	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	**100.00%**

Pred. True	deniso	ellwel	gabefi	kmcgee	millik	ransta	romani	southa	spagen
deniso	**45.98%**	**0.01%**	0.00%	0.00%	**53.94%**	**0.07%**	0.00%	0.00%	0.00%
ellwel	0.00%	**100.00%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
gabefi	0.00%	0.00%	**100.00%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
kmcgee	0.00%	0.00%	0.00%	**94.28%**	0.00%	0.00%	0.00%	0.00%	**5.72%**
millik	**0.10%**	**0.96%**	0.00%	0.00%	**68.16%**	**0.09%**	**30.69%**	0.00%	0.00%
ransta	0.00%	0.00%	0.00%	0.00%	0.00%	**100.00%**	0.00%	0.00%	0.00%
romani	0.00%	0.00%	0.00%	0.00%	**0.01%**	0.00%	**99.99%**	0.00%	0.00%
southa	**0.22%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	**99.78%**	0.00%
spagen	0.00%	0.00%	0.00%	**23.88%**	0.00%	0.00%	0.00%	0.00%	**76.11%**

Pred. True	deniso	ellwel	gabefi	kmcgee	millik	ransta	romani	southa	spagen
deniso	**99.59%**	0.00%	0.00%	0.00%	0.00%	**0.10%**	0.00%	**0.31%**	0.00%
ellwel	0.00%	**100.00%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
gabefi	0.00%	0.00%	**100.00%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
kmcgee	0.00%	0.00%	**0.02%**	**99.42%**	0.00%	0.00%	0.00%	0.00%	**0.57%**
millik	0.00%	**0.03%**	0.00%	0.00%	**98.31%**	**0.01%**	**1.64%**	0.00%	0.00%
ransta	**1.41%**	0.00%	0.00%	0.00%	0.00%	**98.41%**	0.00%	**0.18%**	0.00%
romani	0.00%	0.00%	0.00%	0.00%	**0.09%**	**6.29%**	**93.61%**	0.00%	0.00%
southa	**2.54%**	0.00%	0.00%	0.00%	0.00%	**2.41%**	0.00%	**95.05%**	0.00%
spagen	0.00%	0.00%	0.00%	**16.96%**	0.00%	0.00%	0.00%	0.00%	**83.04%**

Table 75

Average assignes probabilities	Pred. \ True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
	Deniso	**81.86%**	0.00%	0.00%	0.00%	**17.98%**	**0.06%**	0.00%	**0.10%**	0.00%
	Ellwel	0.00%	**100.00%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	GabEFI	0.00%	0.00%	**100.00%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	KMcGee	0.00%	**0.05%**	**0.01%**	**97.84%**	0.00%	0.00%	0.00%	0.00%	**2.09%**
	Millik	**0.03%**	**0.33%**	0.00%	0.00%	**88.82%**	**0.03%**	**10.78%**	0.00%	0.00%
	Ransta	**0.47%**	0.00%	0.00%	0.00%	0.00%	**99.47%**	0.00%	**0.06%**	0.00%
	Romani	0.00%	0.00%	0.00%	0.00%	**0.03%**	**2.10%**	**97.87%**	0.00%	0.00%
	SpaGen	**0.92%**	0.00%	0.00%	0.00%	0.00%	**0.80%**	0.00%	**98.27%**	0.00%
	southa	0.00%	0.00%	0.00%	**13.61%**	0.00%	0.00%	0.00%	0.00%	**86.39%**

SD of assigned probabilities	Pred. \ True	Deniso	Ellwel	GabEFI	KMcGee	Millik	Ransta	Romani	SpaGen	southa
	Deniso	**25.37%**	0.00%	0.00%	0.00%	**25.43%**	**0.04%**	0.00%	**0.15%**	0.00%
	Ellwel	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	GabEFI	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	KMcGee	0.00%	**0.08%**	**0.01%**	**2.52%**	0.00%	0.00%	0.00%	0.00%	**2.57%**
	Millik	**0.05%**	**0.45%**	0.00%	0.00%	**14.63%**	**0.04%**	**14.10%**	0.00%	0.00%
	Ransta	**0.66%**	0.00%	0.00%	0.00%	0.00%	**0.75%**	0.00%	**0.08%**	0.00%
	Romani	0.00%	0.00%	0.00%	0.00%	**0.04%**	**2.97%**	**3.01%**	0.00%	0.00%
	SpaGen	**1.15%**	0.00%	0.00%	0.00%	0.00%	**1.14%**	0.00%	**2.28%**	0.00%
	southa	0.00%	0.00%	0.00%	**10.03%**	0.00%	0.00%	0.00%	0.00%	**10.03%**

Colour class 5

100x

Table 76

Which run	Validation curves	Features selected
First		9 features in total: L4_1_5 L8_2_6 L16_4_15 glszm_GrayLevelVariance L4_1_4 L8_2_10 L8_2_8 glrlm_LongRunLowGrayLevelEmphasis L16_4_3
Second		4 features in total: L36_9_1 L36_9_5 L32_8_34 L12_3_11
Third		5 features in total: glrlm_LongRunLowGrayLevelEmphasis L16_4_15 L24_6_3 L16_4_10 L24_6_20
Union features	L4_1_4, L16_4_15, L24_6_20, L36_9_5, L8_2_8, L16_4_3, L36_9_1, L16_4_10, L8_2_6, L12_3_11, L32_8_34, glrlm_LongRunLowGrayLevelEmphasis, L8_2_10, glszm_GrayLevelVariance, L24_6_3, L4_1_5	

Table 77

filenames	Pred. / True	ElMesq	HDelft	StanRo	USDawn
C100ElMesq-POX_001.tif	ElMesq	**99.35%**	0.00%	0.00%	**0.65%**
B100HDelft-ADU_002.tif	HDelft	0.00%	**100.00%**	0.00%	0.00%
B100StanRo-ADU_003.tif	StanRo	0.00%	0.00%	**100.00%**	0.00%
C100USDawn-ADU_004.tif	USDawn	**9.87%**	0.00%	0.00%	**90.13%**

filenames	Pred. / True	ElMesq	HDelft	StanRo	USDawn
A100ElMesq-POX_003.tif	ElMesq	**67.99%**	0.00%	0.00%	**32.01%**
A100HDelft-ADU_002.tif	HDelft	**13.63%**	**86.37%**	0.01%	0.00%
C100StanRo-ADU_001.tif	StanRo	**0.01%**	0.00%	**99.99%**	0.00%
B100USDawn-ADU_002.tif	USDawn	**1.83%**	0.00%	**0.02%**	**98.16%**

filenames	Pred. / True	ElMesq	HDelft	StanRo	USDawn
B100ElMesq-POX_005.tif	ElMesq	**99.93%**	0.00%	0.00%	**0.07%**
B100HDelft-ADU_003.tif	HDelft	0.00%	**99.88%**	**0.12%**	0.00%
A100StanRo-ADU_004.tif	StanRo	0.00%	0.00%	**100.00%**	0.00%
C100USDawn-ADU_002.tif	USDawn	**97.15%**	0.00%	0.00%	**2.85%**

Table 78

	Pred. / True	ElMesq	HDelft	StanRo	USDawn
Average assigned probabilities	ElMesq	**89.09%**	0.00%	0.00%	**10.91%**
	HDelft	**4.54%**	**95.42%**	**0.04%**	0.00%
	StanRo	0.00%	0.00%	**100.00%**	0.00%
	USDawn	**36.28%**	0.00%	**0.01%**	**63.71%**

	Pred. / True	ElMesq	HDelft	StanRo	USDawn
SD of assigned probabilities	ElMesq	**14.92%**	0.00%	0.00%	**14.92%**
	HDelft	**6.42%**	**6.40%**	**0.05%**	0.00%
	StanRo	0.00%	0.00%	0.00%	0.00%
	USDawn	**43.16%**	0.00%	**0.01%**	**43.16%**

250x

Table 79

Which run	Validation curves	Features selected
First		4 features in total: glszm_GrayLevelVariance L20_5_1 L8_2_2 62
Second		6 features in total: glszm_ZoneEntropy L4_1_2 L8_2_2 L36_9_14 L32_8_12 L44_11_44
Third		4 features in total: glszm_GrayLevelVariance L36_9_6 L40_10_36 L44_11_1
Union features	L36_9_14, glszm_ZoneEntropy, L44_11_1, L36_9_6, L40_10_36, L20_5_1, L32_8_12, 62, L8_2_2, L4_1_2, L44_11_44, glszm_GrayLevelVariance	

Table 80

filenames	Pred. / True	ElMesq	HDelft	StanRo	USDawn
C250ElMesq-POX_001.tif	ElMesq	**48.82%**	0.14%	**11.54%**	**39.50%**
B250HDelft-ADU_002.tif	HDelft	**0.74%**	**98.47%**	0.00%	**0.79%**
B250StanRo-ADU_003.tif	StanRo	0.00%	0.04%	**95.19%**	**4.76%**
C250USDawn-ADU_004.tif	USDawn	**0.22%**	**0.03%**	**3.52%**	**96.23%**

filenames	Pred. / True	ElMesq	HDelft	StanRo	USDawn
A250ElMesq-POX_003.tif	ElMesq	**56.00%**	**3.61%**	**11.48%**	**28.90%**
A250HDelft-ADU_002.tif	HDelft	**47.42%**	**52.48%**	0.04%	0.06%
C250StanRo-ADU_001.tif	StanRo	**0.10%**	**0.26%**	**97.33%**	**2.30%**
B250USDawn-ADU_002.tif	USDawn	**0.13%**	**0.80%**	**4.59%**	**94.47%**

filenames	Pred. / True	ElMesq	HDelft	StanRo	USDawn
B250ElMesq-POX_005.tif	ElMesq	**63.76%**	**36.19%**	0.00%	**0.04%**
B250HDelft-ADU_003.tif	HDelft	**3.89%**	**63.76%**	**30.58%**	**1.76%**
A250StanRo-ADU_004.tif	StanRo	**0.07%**	**0.22%**	**80.30%**	**19.41%**
C250USDawn-ADU_002.tif	USDawn	**65.01%**	**34.03%**	**0.01%**	**0.96%**

Table 81

Average assigned probabilities	Pred. / True	ElMesq	HDelft	StanRo	USDawn
	ElMesq	**56.20%**	13.32%	7.67%	22.81%
	HDelft	17.35%	**71.57%**	10.21%	0.87%
	StanRo	0.06%	0.18%	**90.94%**	8.82%
	USDawn	21.79%	11.62%	2.71%	**63.89%**

SD of assigned probabilities	Pred. / True	ElMesq	HDelft	StanRo	USDawn
	ElMesq	**6.10%**	**16.24%**	**5.43%**	**16.67%**
	HDelft	**21.30%**	**19.57%**	**14.41%**	**0.70%**
	StanRo	**0.04%**	**0.10%**	**7.57%**	**7.55%**
	USDawn	**30.56%**	**15.85%**	**1.96%**	**44.50%**

500x

Which run	Validation curves	Features selected

First

7 features in total:

L8_2_8
L32_8_31
first_order_Minimum
first_order_Range
L24_6_24
L40_10_1
glszm_LargeAreaHighGrayLevelEmphasis

Second

6 features in total:

L4_1_1
first_order_Minimum
first_order_MeanAbsoluteDeviation
L24_6_1
first_order_Median
L16_4_1

Third

6 features in total:

first_order_Range
glszm_LowGrayLevelZoneEmphasis
glcm_DifferenceEntropy_d_10
L44_11_1
glcm_ClusterShade_d_10
L28_7_14

Union features	L44_11_1, L32_8_31, first_order_Minimum, glszm_LowGrayLevelZoneEmphasis, L4_1_1, first_order_Median, L28_7_14, L8_2_8, glcm_DifferenceEntropy_d_10, first_order_Range, L16_4_1, first_order_MeanAbsoluteDeviation, glcm_ClusterShade_d_10, glszm_LargeAreaHighGrayLevelEmphasis, L40_10_1, L24_6_24, L24_6_1

Table 83

filenames	Pred. \ True	ElMesq	HDelft	StanRo	USDawn
C500ElMesq-POX_001.tif	ElMesq	**0.04%**	**0.37%**	**91.33%**	**8.27%**
B500HDelft-ADU_002.tif	HDelft	**83.52%**	**16.47%**	0.00%	**0.01%**
B500StanRo-ADU_003.tif	StanRo	0.00%	0.09%	**93.32%**	6.59%
C500USDawn-ADU_004.tif	USDawn	**16.68%**	**21.37%**	**8.08%**	**53.88%**

filenames	Pred. \ True	ElMesq	HDelft	StanRo	USDawn
A500ElMesq-POX_003.tif	ElMesq	**3.27%**	**33.98%**	**62.18%**	**0.56%**
A500HDelft-ADU_002.tif	HDelft	**69.19%**	**25.84%**	**3.79%**	**1.18%**
C500StanRo-ADU_001.tif	StanRo	**4.89%**	**5.68%**	**79.96%**	**9.47%**
B500USDawn-ADU_002.tif	USDawn	**12.78%**	**23.94%**	**24.26%**	**39.02%**

filenames	Pred. \ True	ElMesq	HDelft	StanRo	USDawn
B500ElMesq-POX_005.tif	ElMesq	**63.46%**	**36.53%**	0.00%	0.00%
B500HDelft-ADU_003.tif	HDelft	**34.16%**	**50.78%**	**1.48%**	**13.58%**
A500StanRo-ADU_004.tif	StanRo	**0.01%**	0.00%	**98.73%**	**1.26%**
C500USDawn-ADU_002.tif	USDawn	**1.34%**	**34.51%**	**1.08%**	**63.07%**

Table 84

Average assigned probabilities	Pred. \ True	ElMesq	HDelft	StanRo	USDawn
	ElMesq	**22.26%**	**23.63%**	**51.17%**	**2.94%**
	HDelft	**62.29%**	**31.03%**	**1.76%**	**4.92%**
	StanRo	**1.63%**	**1.92%**	**90.67%**	**5.77%**
	USDawn	**10.26%**	**26.61%**	**11.14%**	**51.99%**

SD of assigned probabilities	Pred. \ True	ElMesq	HDelft	StanRo	USDawn
	ElMesq	**29.17%**	**16.48%**	**38.09%**	**3.77%**
	HDelft	**20.73%**	**14.48%**	**1.56%**	**6.14%**
	StanRo	**2.30%**	**2.66%**	**7.89%**	**3.40%**
	USDawn	**6.51%**	**5.69%**	**9.71%**	**9.91%**

All magnifications

Table 85

Which run	Validation curves	Features selected
First		3 features in total: L44_11_1_100x L44_11_3_250x L4_1_4_250x
Second		3 features in total: glcm_InverseVariance_d_3_250x L28_7_5_100x L28_7_2_100x
Third		4 features in total: L24_6_24_250x L24_6_24_100x glcm_Autocorrelation_d_10_250x L44_11_1_100x
Union features	L44_11_1_100x, glcm_Autocorrelation_d_10_250x, L24_6_24_100x, L28_7_5_100x, L44_11_3_250x, L4_1_4_250x, L28_7_2_100x, glcm_InverseVariance_d_3_250x, L24_6_24_250x	

Table 86

Pred. True	elmesq	hdelft	stanro	usdawn
elmesq	**94.41%**	**0.05%**	**1.65%**	**3.89%**
hdelft	0.00%	**99.62%**	**0.38%**	0.00%
stanro	**0.12%**	**0.83%**	**99.04%**	0.00%
usdawn	**11.91%**	**0.30%**	**0.47%**	**87.32%**

Pred. True	elmesq	hdelft	stanro	usdawn
elmesq	**18.19%**	**22.45%**	**0.01%**	**59.34%**
hdelft	**29.80%**	**69.44%**	**0.04%**	**0.73%**
stanro	0.00%	**0.01%**	**99.58%**	**0.41%**
usdawn	**0.02%**	**0.34%**	**39.10%**	**60.54%**

Pred. True	elmesq	hdelft	stanro	usdawn
elmesq	**97.54%**	0.00%	0.00%	**2.46%**
hdelft	**0.13%**	**99.87%**	0.00%	0.00%
stanro	0.00%	0.00%	**100.00%**	0.00%
usdawn	**0.30%**	0.00%	0.00%	**99.70%**

Table 87

	Pred. True	ElMesq	HDelft	StanRo	USDawn
Average assigned probabilities	ElMesq	**70.05%**	**7.50%**	**0.55%**	**21.90%**
	HDelft	**9.98%**	**89.64%**	**0.14%**	**0.24%**
	StanRo	**0.04%**	**0.28%**	**99.54%**	**0.14%**
	USDawn	**4.08%**	**0.22%**	**13.19%**	**82.52%**

	Pred. True	ElMesq	HDelft	StanRo	USDawn
SD of assigned probabilities	ElMesq	**36.69%**	**10.57%**	**0.78%**	**26.49%**
	HDelft	**14.02%**	**14.29%**	**0.17%**	**0.34%**
	StanRo	**0.06%**	**0.39%**	**0.39%**	**0.19%**
	USDawn	**5.54%**	**0.15%**	**18.32%**	**16.34%**

Colour class 6

100x

Table 88

Which run	Validation curves	Features selected
First		6 features in total: L12_3_4 L44_11_3 L4_1_6 L44_11_9 L20_5_1 L40_10_3
Second		6 features in total: L12_3_4 L12_3_5 L20_5_1 L36_9_3 first_order_Median L44_11_11
Third		6 features in total: L12_3_4 L20_5_1 L40_10_3 L8_2_6 L40_10_10 L4_1_6
Union features	first_order_Median, L44_11_11, L40_10_3, L12_3_5, L8_2_6, L12_3_4, L44_11_3, L44_11_9, L36_9_3, L4_1_6, L20_5_1, L40_10_10	

Table 89

filenames	Pred. / True	EverYe	Irigar	Rabbla	UMobil
E100EverYe-POX_003.tif	EverYe	**100.00%**	0.00%	0.00%	0.00%
C100Irigar-POX_001.tif	Irigar	0.00%	**99.09%**	0.00%	**0.91%**
C100Rabbla-POX_005.tif	Rabbla	0.00%	0.00%	**100.00%**	0.00%
A100UMobil-ADU_001.tif	UMobil	0.00%	0.00%	0.00%	**100.00%**

filenames	Pred. / True	EverYe	Irigar	Rabbla	UMobil
D100EverYe-POX_002.tif	EverYe	**100.00%**	0.00%	0.00%	0.00%
A100Irigar-POX_003.tif	Irigar	0.00%	**100.00%**	0.00%	0.00%
B100Rabbla-POX_002.tif	Rabbla	0.00%	**0.18%**	**99.82%**	0.00%
C100UMobil-ADU_002.tif	UMobil	0.00%	0.00%	0.00%	**100.00%**

filenames	Pred. / True	EverYe	Irigar	Rabbla	UMobil
D100EverYe-POX_001.tif	EverYe	**100.00%**	0.00%	0.00%	0.00%
B100Irigar-POX_005.tif	Irigar	0.00%	**100.00%**	0.00%	0.00%
A100Rabbla-POX_002.tif	Rabbla	0.00%	0.00%	**100.00%**	0.00%
B100UMobil-ADU_002.tif	UMobil	0.00%	**0.30%**	0.00%	**99.70%**

Table 90

Average assigned probabilities	Pred. / True	EverYe	Irigar	Rabbla	UMobil
	EverYe	**100.00%**	0.00%	0.00%	0.00%
	Irigar	0.00%	**99.70%**	0.00%	**0.30%**
	Rabbla	0.00%	**0.06%**	**99.94%**	0.00%
	UMobil	0.00%	**0.10%**	0.00%	**99.90%**

SD of assigned probabilities	Pred. / True	EverYe	Irigar	Rabbla	UMobil
	EverYe	0.00%	0.00%	0.00%	0.00%
	Irigar	0.00%	**0.43%**	0.00%	**0.43%**
	Rabbla	0.00%	**0.09%**	**0.09%**	0.00%
	UMobil	0.00%	**0.14%**	0.00%	**0.14%**

250x

Table 91

Which run	Validation curves	Features selected
First		3 features in total: L44_11_6 L8_2_7 L12_3_2
Second		4 features in total: L20_5_3 L8_2_2 L8_2_7 L40_10_5
Third		4 features in total: L40_10_5 L8_2_2 first_order_Range glcm_ClusterProminence_d_20
Union features	L20_5_3, L8_2_7, glcm_ClusterProminence_d_20, first_order_Range, L12_3_2, L44_11_6, L8_2_2, L40_10_5	

Table 92

filenames	Pred. True	EverYe	Irigar	Rabbla	UMobil
E250EverYe-POX_003.tif	EverYe	**99.92%**	0.00%	**0.01%**	**0.06%**
C250Irigar-POX_001.tif	Irigar	**0.01%**	**99.98%**	0.00%	**0.01%**
C250Rabbla-POX_005.tif	Rabbla	0.00%	0.00%	**98.61%**	**1.39%**
A250UMobil-ADU_001.tif	UMobil	**0.09%**	**0.09%**	**5.25%**	**94.57%**

filenames	Pred. True	EverYe	Irigar	Rabbla	UMobil
D250EverYe-POX_002.tif	EverYe	**100.00%**	0.00%	0.00%	0.00%
A250Irigar-POX_003.tif	Irigar	0.00%	**99.46%**	0.00%	**0.54%**
B250Rabbla-POX_002.tif	Rabbla	0.00%	0.00%	**98.87%**	**1.13%**
C250UMobil-ADU_002.tif	UMobil	0.00%	**0.93%**	0.00%	**99.06%**

filenames	Pred. True	EverYe	Irigar	Rabbla	UMobil
D250EverYe-POX_001.tif	EverYe	**98.75%**	**1.16%**	0.00%	**0.09%**
B250Irigar-POX_005.tif	Irigar	**0.09%**	**99.87%**	0.00%	**0.04%**
A250Rabbla-POX_002.tif	Rabbla	0.00%	0.00%	**100.00%**	0.00%
B250UMobil-ADU_002.tif	UMobil	**14.31%**	**4.19%**	0.00%	**81.50%**

Table 93

Average assigned probabilities	Pred. True	EverYe	Irigar	Rabbla	UMobil
	EverYe	**99.56%**	**0.39%**	0.00%	**0.05%**
	Irigar	**0.03%**	**99.77%**	0.00%	**0.20%**
	Rabbla	0.00%	0.00%	**99.16%**	**0.84%**
	UMobil	**4.80%**	**1.74%**	**1.75%**	**91.71%**

SD of assigned probabilities	Pred. True	EverYe	Irigar	Rabbla	UMobil
	EverYe	**0.57%**	**0.55%**	**0.01%**	**0.04%**
	Irigar	**0.04%**	**0.22%**	0.00%	**0.24%**
	Rabbla	0.00%	0.00%	**0.60%**	**0.60%**
	UMobil	**6.73%**	**1.77%**	**2.47%**	**7.45%**

500x

Table 94

Which run	Validation curves	Features selected
First		6 features in total: glrlm_LowGrayLevelRunEmphasis L4_1_1 L4_1_4 first_order_Maximum 4 26
Second		5 features in total: glrlm_LowGrayLevelRunEmphasis L4_1_1 L4_1_2 L32_8_33 glszm_GrayLevelVariance
Third		4 features in total: L4_1_2 L12_3_13 L20_5_17 L16_4_6
Union features	L20_5_17, glrlm_LowGrayLevelRunEmphasis, L4_1_4, L16_4_6, L4_1_1, L12_3_13, glszm_GrayLevelVariance, first_order_Maximum, 4, L4_1_2, 26, L32_8_33	

Table 95

filenames	Pred. / True	EverYe	Irigar	Rabbla	UMobil
E500EverYe-POX_003.tif	EverYe	**71.62%**	**1.68%**	**0.58%**	**26.11%**
C500Irigar-POX_001.tif	Irigar	**22.58%**	**77.31%**	0.00%	**0.11%**
C500Rabbla-POX_005.tif	Rabbla	0.00%	0.00%	**99.88%**	**0.12%**
A500UMobil-ADU_001.tif	UMobil	0.00%	0.00%	**68.79%**	**31.20%**

filenames	Pred. / True	EverYe	Irigar	Rabbla	UMobil
D500EverYe-POX_002.tif	EverYe	**66.41%**	0.00%	**0.01%**	33.58%
A500Irigar-POX_003.tif	Irigar	0.00%	**100.00%**	0.00%	0.00%
B500Rabbla-POX_002.tif	Rabbla	**41.40%**	0.00%	**0.09%**	58.50%
C500UMobil-ADU_002.tif	UMobil	**84.01%**	0.00%	0.00%	**15.99%**

filenames	Pred. / True	EverYe	Irigar	Rabbla	UMobil
D500EverYe-POX_001.tif	EverYe	**91.35%**	0.03%	2.28%	6.34%
B500Irigar-POX_005.tif	Irigar	0.00%	**99.96%**	0.02%	0.01%
A500Rabbla-POX_002.tif	Rabbla	**0.13%**	0.00%	**95.20%**	4.67%
B500UMobil-ADU_002.tif	UMobil	**0.88%**	**4.37%**	**20.34%**	**74.40%**

Table 96

	Pred. / True	EverYe	Irigar	Rabbla	UMobil
Average assigned probabilities	EverYe	**76.46%**	**0.57%**	**0.96%**	**22.01%**
	Irigar	**7.53%**	**92.42%**	**0.01%**	**0.04%**
	Rabbla	**13.85%**	0.00%	**65.06%**	**21.10%**
	UMobil	**28.30%**	**1.46%**	**29.71%**	**40.53%**

	Pred. / True	EverYe	Irigar	Rabbla	UMobil
SD of assigned probabilities	EverYe	**10.74%**	**0.79%**	**0.96%**	**11.49%**
	Irigar	**10.64%**	**10.69%**	**0.01%**	**0.05%**
	Rabbla	**19.49%**	0.00%	**45.98%**	**26.52%**
	UMobil	**39.40%**	**2.06%**	**28.86%**	**24.74%**

All magnifications

Table 97

Which run	Validation curves	Features selected
First		**3 features in total:** L32_8_1_100x L20_5_21_100x L16_4_8_250x
Second		**3 features in total:** L44_11_3_100x L16_4_2_250x first_order_Median_100x
Third		**3 features in total:** L4_1_2_500x L36_9_1_100x L8_2_4_250x
Union features	L16_4_2_250x, L20_5_21_100x, L4_1_2_500x, L44_11_3_100x, L32_8_1_100x, L36_9_1_100x, L8_2_4_250x, L16_4_8_250x, first_order_Median_100x	

Table 98

Pred. \ True	everye	irigar	rabbla	umobil
everye	**91.64%**	**8.36%**	0.00%	0.00%
irigar	0.00%	**99.99%**	0.00%	**0.01%**
rabbla	0.00%	0.00%	**100.00%**	0.00%
umobil	0.00%	0.00%	0.00%	**100.00%**

Pred. \ True	everye	irigar	rabbla	umobil
everye	**100.00%**	0.00%	0.00%	0.00%
irigar	**0.11%**	**99.89%**	0.00%	0.00%
rabbla	0.00%	0.00%	**100.00%**	0.00%
umobil	0.00%	**9.58%**	0.00%	**90.42%**

Pred. \ True	everye	irigar	rabbla	umobil
everye	**52.88%**	**47.11%**	0.00%	0.00%
irigar	0.00%	**100.00%**	0.00%	0.00%
rabbla	0.00%	0.00%	**99.89%**	**0.11%**
umobil	0.00%	**0.56%**	0.00%	**99.44%**

Table 99

	Pred. \ True	EverYe	Irigar	Rabbla	UMobil
Average assigned probabilities	EverYe	**81.51%**	**18.49%**	0.00%	0.00%
	Irigar	**0.04%**	**99.96%**	0.00%	0.00%
	Rabbla	0.00%	0.00%	**99.96%**	**0.04%**
	UMobil	0.00%	**3.38%**	0.00%	**96.62%**

	Pred. \ True	EverYe	Irigar	Rabbla	UMobil
SD of assigned probabilities	EverYe	**20.53%**	**20.52%**	0.00%	0.00%
	Irigar	**0.05%**	**0.05%**	0.00%	0.00%
	Rabbla	0.00%	0.00%	**0.05%**	**0.05%**
	UMobil	0.00%	**4.39%**	0.00%	**4.39%**

Spectra

Colour class 1

Which run	Validation curves	Features selected
First		2 features in total: 1319.14 1478.01
Second		2 features in total: 1319.14 1481.55
Third		2 features in total: 1319.14 1474.46
Union features	1319.14, 1474.46, 1478.01, 1481.55	

Table 101

filenames	Pred.⟍True	Aus_Oly	Can_Key	Chi_Hen	Rus_Tec	Zam_Min
QD_28-03_0168TL	Aus_Oly	**100.00%**	0.00%	0.00%	0.00%	0.00%
QD_21-05_0216BR	Can_Key	0.00%	**100.00%**	0.00%	0.00%	0.00%
QD_22-03_0224TL	Chi_Hen	0.00%	0.00%	**100.00%**	0.00%	0.00%
QD_70-01_0171TR	Rus_Tec	0.00%	0.00%	0.00%	**100.00%**	0.00%
QD_06-01_0202BL	Zam_Min	0.00%	0.00%	0.00%	0.00%	**100.00%**

filenames	Pred.⟍True	Aus_Oly	Can_Key	Chi_Hen	Rus_Tec	Zam_Min
QD_28-01_0165TL	Aus_Oly	**100.00%**	0.00%	0.00%	0.00%	0.00%
QD_21-06_0218TR	Can_Key	0.00%	**100.00%**	0.00%	0.00%	0.00%
QD_22-01_0220TL	Chi_Hen	0.00%	0.00%	**100.00%**	0.00%	0.00%
QD_70-02_0173TL	Rus_Tec	0.00%	0.00%	0.00%	**100.00%**	0.00%
QD_06-03_0207BL	Zam_Min	0.00%	0.00%	0.00%	0.00%	**100.00%**

filenames	Pred.⟍True	Aus_Oly	Can_Key	Chi_Hen	Rus_Tec	Zam_Min
QD_28-02_0167BL	Aus_Oly	**100.00%**	0.00%	0.00%	0.00%	0.00%
QD_21-04_0215TR	Can_Key	**0.02%**	**99.98%**	0.00%	0.00%	0.00%
QD_22-02_0221BL	Chi_Hen	0.00%	0.00%	**100.00%**	0.00%	0.00%
QD_70-01_0171TL	Rus_Tec	0.00%	0.00%	0.00%	**100.00%**	0.00%
QD_06-02_0203TL	Zam_Min	0.00%	0.00%	0.00%	0.00%	**100.00%**

Table 102

Average assigned probabilities	Pred.⟍True	Aus_Oly	Can_Key	Chi_Hen	Rus_Tec	Zam_Min
	Aus_Oly	**100.00%**	0.00%	0.00%	0.00%	0.00%
	Can_Key	**0.01%**	**99.99%**	0.00%	0.00%	0.00%
	Chi_Hen	0.00%	0.00%	**100.00%**	0.00%	0.00%
	Rus_Tec	0.00%	0.00%	0.00%	**100.00%**	0.00%
	Zam_Min	0.00%	0.00%	0.00%	0.00%	**100.00%**

SD of assigned probabilities	Pred.⟍True	Aus_Oly	Can_Key	Chi_Hen	Rus_Tec	Zam_Min
	Aus_Oly	0.00%	0.00%	0.00%	0.00%	0.00%
	Can_Key	**0.01%**	**0.01%**	0.00%	0.00%	0.00%
	Chi_Hen	0.00%	0.00%	0.00%	0.00%	0.00%
	Rus_Tec	0.00%	0.00%	0.00%	0.00%	0.00%
	Zam_Min	0.00%	0.00%	0.00%	0.00%	0.00%

Colour class 3 with sav.gol. filter and f.d. transf.
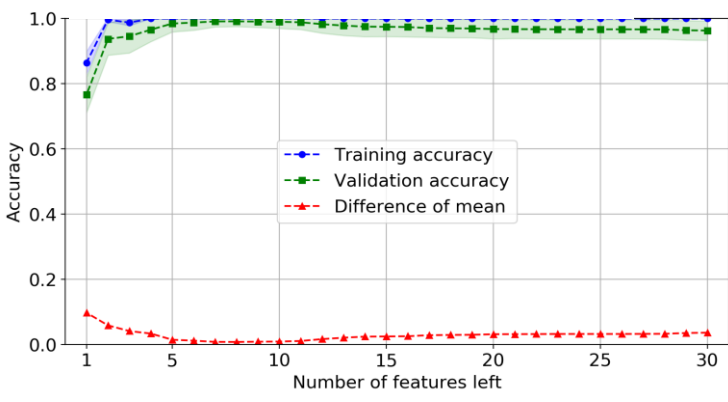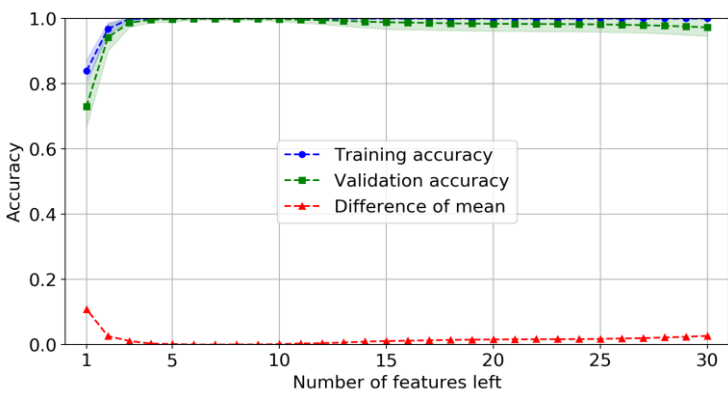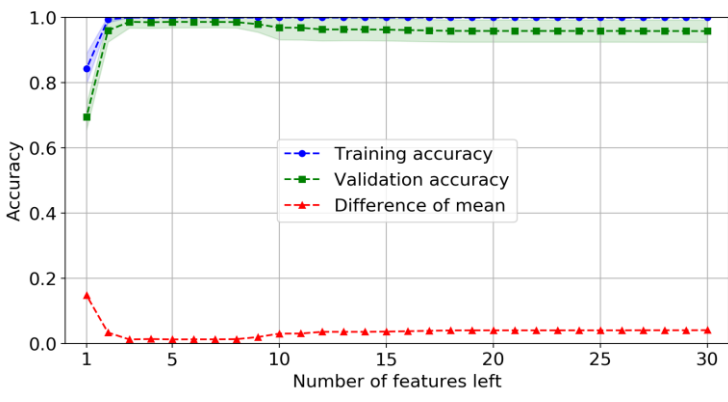
Which run	Validation curves	Features selected
First		2 features in total: 1424.92 1481.55
Second		2 features in total: 1382.55 1424.92
Third		2 features in total: 1428.46 1538.32
Union features	1382.55, 1424.92, 1428.46, 1481.55, 1538.32	

Table 104

filenames	Pred.⟍ True	Aus_Rad	Aus_Run	Bra_Nuc	Can_Dyn	Can_Far	Can_Sun	Ger_Wis	USA_Cot	USA_Fal	Yog_Rud	Yog_Spi
QD_40-02_0201BL	Aus_Rad	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_65-02_0144BL	Aus_Run	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_15-02_0066TR	Bra_Nuc	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_20-02_0094BL	Can_Dyn	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_27-03_0255BR	Can_Far	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_25-02_0242TL	Can_Sun	0.00%	0.00%	0.00%	0.00%	0.00%	99.98%	0.00%	0.00%	0.01%	0.00%	0.00%
QD_18-02_0186BR	Ger_Wis	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	80.75%	19.17%	0.03%	0.00%	0.04%
QD_07-01_0235BL	USA_Cot	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	23.54%	76.46%	0.00%	0.00%	0.00%
QD_31-1_0045BR	USA_Fal	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	80.00%	0.00%	19.98%
QD_72-01_0098BL	Yog_Rud	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
QD_29-02_0080TR	Yog_Spi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.66%	0.03%	39.94%	0.00%	59.37%

filenames	Pred.⟍ True	Aus_Rad	Aus_Run	Bra_Nuc	Can_Dyn	Can_Far	Can_Sun	Ger_Wis	USA_Cot	USA_Fal	Yog_Rud	Yog_Spi
QD_40-01_0198BR	Aus_Rad	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_65-01_0143TL	Aus_Run	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_15-03_0068BR	Bra_Nuc	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_20-01_0093BL	Can_Dyn	0.00%	0.00%	0.00%	97.78%	0.00%	2.22%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_27-02_0251BL	Can_Far	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_25-01_0240TR	Can_Sun	0.00%	0.00%	0.00%	3.99%	0.00%	96.01%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_18-03_0187TR	Ger_Wis	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	98.59%	0.00%	1.40%	0.00%	0.00%
QD_07-03_0238BL	USA_Cot	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.06%	99.94%	0.00%	0.00%	0.00%
QD_31-3_0048TR	USA_Fal	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.39%	0.00%	86.92%	0.00%	11.68%
QD_72-02_0102TR	Yog_Rud	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
QD_29-03_0081BL	Yog_Spi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	7.30%	0.00%	92.69%

filenames	Pred.⟍ True	Aus_Rad	Aus_Run	Bra_Nuc	Can_Dyn	Can_Far	Can_Sun	Ger_Wis	USA_Cot	USA_Fal	Yog_Rud	Yog_Spi
QD_40-02_0201BR	Aus_Rad	99.95%	0.05%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_65-03_0145BL	Aus_Run	0.00%	99.99%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_15-01_0065BR	Bra_Nuc	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_20-01_0093TL	Can_Dyn	0.00%	0.00%	0.00%	60.45%	0.00%	39.55%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_27-01_0247BL	Can_Far	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_25-03_0245BL	Can_Sun	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_18-01_0189BR	Ger_Wis	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	79.25%	0.00%	0.00%	0.00%	20.75%
QD_07-02_0239TL	USA_Cot	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	93.73%	6.25%	0.00%	0.00%
QD_31-2_0046TR	USA_Fal	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	17.55%	82.44%	0.00%	0.00%
QD_72-03_0105TR	Yog_Rud	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
QD_29-01_0077TR	Yog_Spi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	3.92%	0.00%	0.00%	0.00%	96.08%

Table 105

Average assigned probabilities

True \ Pred.	Aus_Rad	Aus_Run	Bra_Nuc	Can_Dyn	Can_Far	Can_Sun	Ger_Wis	USA_Cot	USA_Fal	Yog_Rud	Yog_Spi
Aus_Rad	99.98%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Aus_Run	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Bra_Nuc	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Can_Dyn	0.00%	0.00%	0.00%	86.08%	0.00%	13.92%	0.00%	0.00%	0.00%	0.00%	0.00%
Can_Far	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Can_Sun	0.00%	0.00%	0.00%	1.33%	0.00%	98.66%	0.00%	0.00%	0.00%	0.00%	0.00%
Ger_Wis	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	86.20%	6.39%	0.48%	0.00%	6.93%
USA_Cot	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	7.87%	90.04%	2.09%	0.00%	0.00%
USA_Fal	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.47%	5.85%	83.12%	0.00%	10.55%
Yog_Rud	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
Yog_Spi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.53%	0.01%	15.75%	0.00%	82.71%

SD of assigned probabilities

True \ Pred.	Aus_Rad	Aus_Run	Bra_Nuc	Can_Dyn	Can_Far	Can_Sun	Ger_Wis	USA_Cot	USA_Fal	Yog_Rud	Yog_Spi
Aus_Rad	0.02%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Aus_Run	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Bra_Nuc	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Can_Dyn	0.00%	0.00%	0.00%	15.59%	0.00%	15.59%	0.00%	0.00%	0.00%	0.00%	0.00%
Can_Far	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Can_Sun	0.00%	0.00%	0.00%	1.66%	0.00%	1.66%	0.00%	0.00%	0.00%	0.00%	0.00%
Ger_Wis	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	8.00%	3.01%	0.58%	0.00%	8.63%
USA_Cot	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	3.69%	4.08%	2.60%	0.00%	0.00%
USA_Fal	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.58%	7.30%	1.97%	0.00%	5.26%
Yog_Rud	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Yog_Spi	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.61%	0.00%	6.44%	0.00%	5.67%

Colour class 4

Table 106

Which run	Validation curves	Features selected
First		3 features in total: 1252.41 1495.73 1517.01
Second		3 features in total: 1396.67 1517.01 1527.66
Third		2 features in total: 1467.38 1492.18
Union features	1252.41, 1396.67, 1467.38, 1492.18, 1495.73, 1517.01, 1527.66	

Table 107

filenames	Pred. / True	Aus_S A	Can_Mil	Can_Nor	Ger_Hel	Rum_Rum	Spa_Gen	Swe_Ran
QD_73-02_0182BR	Aus_S A	**100.00%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
QD_64-03_0086BR	Can_Mil	0.00%	**100.00%**	0.00%	0.00%	0.00%	0.00%	0.00%
QD_41-01_0133BR	Can_Nor	0.00%	0.00%	**99.97%**	0.00%	0.00%	0.00%	**0.03%**
QD_45-03_0123TR	Ger_Hel	0.00%	0.00%	0.00%	**100.00%**	0.00%	0.00%	0.00%
QD_76-03_0197BR	Rum_Rum	0.00%	0.00%	0.00%	0.00%	**100.00%**	0.00%	0.00%
QD_03-06_0061BL	Spa_Gen	0.00%	**0.01%**	0.00%	0.00%	0.00%	**99.99%**	0.00%
QD_53-01_0108BL	Swe_Ran	0.00%	0.00%	**1.36%**	0.00%	0.00%	0.00%	**98.64%**

filenames	Pred. / True	Aus_S A	Can_Mil	Can_Nor	Ger_Hel	Rum_Rum	Spa_Gen	Swe_Ran
QD_73-03_0183TL	Aus_S A	**98.96%**	0.00%	0.00%	0.00%	**1.04%**	0.00%	0.00%
QD_64-02_0084TR	Can_Mil	0.00%	**99.88%**	0.00%	0.00%	**0.12%**	0.00%	0.00%
QD_41-02_0137BR	Can_Nor	0.00%	0.00%	**100.00%**	0.00%	0.00%	0.00%	0.00%
QD_45-02_0119BR	Ger_Hel	0.00%	0.00%	0.00%	**100.00%**	0.00%	0.00%	0.00%
QD_76-02_0194TR	Rum_Rum	**0.01%**	**0.01%**	0.00%	0.00%	**99.99%**	0.00%	0.00%
QD_03-04_0058BL	Spa_Gen	0.00%	0.00%	0.00%	0.00%	0.00%	**100.00%**	0.00%
QD_53-03_0112BR	Swe_Ran	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	**100.00%**

filenames	Pred. / True	Aus_S A	Can_Mil	Can_Nor	Ger_Hel	Rum_Rum	Spa_Gen	Swe_Ran
QD_73-01_0181TR	Aus_S A	**99.44%**	**0.56%**	0.00%	0.00%	0.00%	0.00%	0.00%
QD_64-01_0083TL	Can_Mil	**1.82%**	**97.56%**	0.00%	**0.62%**	0.00%	0.00%	0.00%
QD_41-03_0139TR	Can_Nor	0.00%	0.00%	**76.63%**	0.00%	0.00%	**0.01%**	**23.36%**
QD_45-01_0115BR	Ger_Hel	0.00%	**8.29%**	0.00%	**91.71%**	0.00%	0.00%	0.00%
QD_76-01_0192TR	Rum_Rum	0.00%	0.00%	0.00%	0.00%	**100.00%**	0.00%	0.00%
QD_03-05_0059BR	Spa_Gen	0.00%	0.00%	0.00%	0.00%	0.00%	**100.00%**	0.00%
QD_53-02_0110BR	Swe_Ran	0.00%	0.00%	**2.03%**	0.00%	0.00%	0.00%	**97.97%**

Table 108

	Pred. True	Aus_S A	Can_Mil	Can_Nor	Ger_Hel	Rum_Rum	Spa_Gen	Swe_Ran
Average assigned probabilities	Aus_S A	**99.47%**	**0.19%**	0.00%	0.00%	**0.35%**	0.00%	0.00%
	Can_Mil	**0.61%**	**99.15%**	0.00%	**0.21%**	**0.04%**	0.00%	0.00%
	Can_Nor	0.00%	0.00%	**92.20%**	0.00%	0.00%	0.00%	**7.80%**
	Ger_Hel	0.00%	**2.76%**	0.00%	**97.24%**	0.00%	0.00%	0.00%
	Rum_Rum	0.00%	0.00%	0.00%	0.00%	**99.99%**	0.00%	0.00%
	Spa_Gen	0.00%	0.00%	0.00%	0.00%	0.00%	**100.00%**	0.00%
	Swe_Ran	0.00%	0.00%	**1.13%**	0.00%	0.00%	0.00%	**98.87%**

	Pred. True	Aus_S A	Can_Mil	Can_Nor	Ger_Hel	Rum_Rum	Spa_Gen	Swe_Ran
SD of assigned probabilities	Aus_S A	**0.42%**	**0.26%**	0.00%	0.00%	**0.49%**	0.00%	0.00%
	Can_Mil	**0.86%**	**1.12%**	0.00%	**0.29%**	**0.06%**	0.00%	0.00%
	Can_Nor	0.00%	0.00%	**11.01%**	0.00%	0.00%	0.00%	**11.00%**
	Ger_Hel	0.00%	**3.91%**	0.00%	**3.91%**	0.00%	0.00%	0.00%
	Rum_Rum	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Spa_Gen	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Swe_Ran	0.00%	0.00%	**0.84%**	0.00%	0.00%	0.00%	**0.84%**

Colour class6

Table 109

Which run	Validation curves	Features selected
First		3 features in total: 1196.37 1199.87 1424.92
Second		5 features in total: 1189.37 1252.41 1414.32 1577.43 1684.44
Third		6 features in total: 1196.37 1410.79 1414.32 1446.14 1573.87 1709.49
Union features	1189.37, 1196.37, 1199.87, 1252.41, 1410.79, 1414.32, 1424.92, 1446.14, 1573.87, 1577.43, 1684.44, 1709.49	

Table 110

filenames	Pred. True	USA_Eve	USA_Iri	USA_Mob
QD_52-3_0043TL	USA_Eve	**78.55%**	4.33%	17.12%
QD_33-01_0088BR	USA_Iri	0.20%	**88.47%**	11.32%
QD_38-3_0035TR	USA_Mob	37.90%	30.54%	**31.56%**

filenames	Pred. True	USA_Eve	USA_Iri	USA_Mob
QD_52-1_0042TL	USA_Eve	**60.75%**	39.21%	0.04%
QD_33-03_0090BL	USA_Iri	0.00%	**100.00%**	0.00%
QD_38-1_0032BL	USA_Mob	9.89%	0.05%	**90.06%**

filenames	Pred. True	USA_Eve	USA_Iri	USA_Mob
QD_52-1_0042BL	USA_Eve	**99.99%**	0.00%	0.00%
QD_33-02_0089TL	USA_Iri	0.04%	**99.94%**	0.02%
QD_38-3_0035BR	USA_Mob	26.98%	0.19%	**72.83%**

Table 111

Average assigned probabilities	Pred. True	USA_Eve	USA_Iri	USA_Mob
	USA_Eve	**79.76%**	14.51%	5.72%
	USA_Iri	0.08%	**96.14%**	3.78%
	USA_Mob	24.92%	10.26%	**64.82%**

SD of assigned probabilities	Pred. True	USA_Eve	USA_Iri	USA_Mob
	USA_Eve	16.04%	17.55%	8.06%
	USA_Iri	0.09%	5.42%	5.33%
	USA_Mob	11.53%	14.34%	24.54%

III.   Codes

The codes used in this thesis can found in the authors GitHub repository
https://github.com/iLAN-Git/masterthesis (Lande, u.d.).