Combining spectral and lateral information during the evaluation of hyperspectral data

Faidra Amargianou

Master of Science Thesis

TU Delft Delft University of Technology

Materials Science and Engineering

Combining spectral and lateral information during the evaluation of hyperspectral data

Master of Science Thesis

In fulfilment of the requirements for the award of the degree of Master of Science in Materials Science and Engineering at Delft University of Technology.

Thesis committee:

Prof. Dr. Joris Dik
Dr. Matthias Alfeld
Dr. Siddhant Kumar

to be defended publicly on Friday March 26, 2021 at 12:00.
An electronic version of this thesis is available at
http://repository.tudelft.nl/.

Faidra Amargianou

March 22, 2021

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology

Abstract

Hyperspectral techniques have found application for the investigation of objects and samples in many scientific fields. Data evaluation approaches commonly either process them as stacks of randomly ordered spectra or as flat images. This thesis aims at combining both, lateral and spectral, aspects of hyperspectral data during data evaluation. Different approaches, such as lateral distance and neighbor pixel augmentation, will be explored with modern factorization techniques, such as t-stochastic neighbor embedding and self-organizing maps, and supported by artificial neural networks. The routines developed are applied on selected X-ray fluorescence data (XRF) sets from the field of materials science and cultural heritage.

The altered augmented data set concept consists in augmenting the spectrum of a central pixel with the mean spectrum of its eight neighboring pixels. This new method is applied to the data set of a bi-modal Ti-6Al-6V-2Sn alloy. Clustering the optimized augmented data set with t-SNE reveals the existence of four components: phase with strong titanium signal, phase with strong vanadium signal, the first phase on top of the second phase and the second phase on top of the first phase. The later was not identified by clustering with t-SNE the data set augmented with all the spectra of its eight neighboring pixels.

Clustering the XRF data set of a 13th century B.C. Egyptian mural painting with Self Organising Map (SOM) indicates areas with different thicknesses of copper and iron containing pigments. Clustering with Fast interpolation-based t-SNE (FIt-SNE) also reveals information about the painting sequence and the composition of a mixture of pigments. The pigments are not separable with a simple observation of the XRF elemental distribution images. Moreover, FIt-SNE dramatically accelerates t-SNE and can provide well-separated clusters. These properties are especially useful for large data sets, containing mixture of pigments.

Finally, I introduce an optimised Artificial Neural Network (ANN) for training the photograph of a painting with the elemental distribution images, obtained by its XRF data set. The elemental distribution images are upscaled with Laplacian Pyramid Super-Resolution Network. In this way, we can identify with more accuracy the exact location of the damages and retouches on a 17th century A.C. easel painting (portrait of Hortense Manchini), since the result image is of high resolution.

Table of Contents

Acknowledgements

The present Master of Science Thesis represents my academic choice to serve the field of Materials Science in the most fulfilling and creative way possible. This would have not been possible without the contribution of certain people to my personal and professional development.

I would like to express my thanks to the members of my committee: Prof. Joris Dik, Dr. Matthias Alfeld and Dr. Siddhant Kumar, for accepting to be part of the thesis committee. The constant concern and understanding of Dr. Matthias Alfeld, to guide me though a modern scientific approach of exploring with a broad set of tools the field of Materials Science, equipped me with scientific persistence and knowledge.

I want to express my deep gratitude to my mother Stavroula, my father Panagiotis and my sister Myrto for their support and influence on my core beliefs.

To my friends living in Greece, Canada and Netherlands, whose presence in my life proves that friendship does not recognize distance, nationality and gender, thank you.

My words come short to express the appreciation for Tudor, my flatmate in Delft.

Delft University of Technology opened up new perspectives to knowledge. I will always admire Netherlands for its culture, that I experienced through my visits to its museums. It is a life-changing experience.

Delft, University of Technology
March 22, 2021

Sincerely,
Faidra Amargianou

"I sometimes found myself writing letters to that girl containing nothing but lines of crosses, one after another. One of those letters later fell into the hands of some people who had an interest in penetrating the mystery, but they could never find fault with such a devout cipher."

— *Hortense Manchini, Memoirs*

Chapter 1

Introduction

1-1 The importance of machine learning in cultural heritage data processing

Machine and deep learning are powerful tools that influence dramatically the advances of many sectors, from medicine to cultural heritage. Its value would not be equally appreciated without the scientific examination of paintings, archaeological treasures and other objects of historical value. ML algorithms coupled with experimental data unravel the richness of cultural heritage all around the world. The elemental mapping of a painting reveal its mysteries, from its origin and authenticity to its process of creation. Image processing has versatile functions, since it can classify the paintings or it can constitute a guide for restoration.

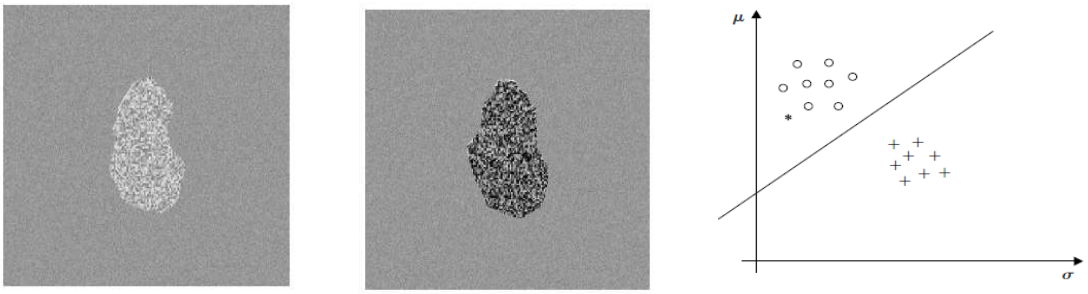
Especially during the cultural heritage campaigns, more and more data of a certain category of artistic creations and other cultural heritage objects are extracted. How can all these data transformed into information? ML algorithms have long proven to provide the ability of processing of data from the cultural heritage field. On one hand, the methods of ML applied to cultural heritage data follow the general ML advances and hence the reproducibility of these algorithms are of crucial importance [1].

On the other hand, in a smaller degree the special demands and evolution of the cultural heritage scientific research define the kind of ML algorithm compatible for analysing the relevant data. Unsupervised machine learning algorithms, such as Principal Component Analysis (PCA), are very efficient for the investigation of spectral data. However, supervised learning with methods such as artificial neural networks can be the key for the processing of big data. The main problem of supervised ML in the field of interest is the lack of suitable learning data for algorithm's training. In the future, the increasing number of data from the cultural heritage field will create groups of similar data set, suitable for training. Therefore, supervised ML seem as a promising answer to the demands of cultural heritage data interpretation and presentation.

1-2 The fundamentals of machine learning

Machine learning is a field of study, where the computers are not explicitly programmed but they learn a task from experience. The algorithms, provided by machine learning, are employed to solve a broad set of problems in pattern recognition. Image classification is an area in which machine learning is of importance.

Initially, an example of supervised machine learning task is presented, a simple case of image classification. Fig. 1-1 (a)-(b) presents two images, each having a distinct area. Fig. 1-1 (a) presents an area, which belongs to class A and Fig. 1-1 (b) presents an area, which belongs to class B. We also have available a set of images, that we know belong to class A or class B. Initially, we need to identify the measurable quantities, that distinct the two regions. Fig. 1-1 (c) assumes that these measurable quantities are the mean value and the standard deviation. Therefore, after estimating the mean values and standard deviations from the available data set with images, we plot the mean value of the intensity in each region of interest versus the corresponding standard deviation around this mean. Each point corresponds to a different image from the available database. The straight line is considered an option for separating the two classes. The measurements used for the classification, the mean value and the standard deviation in this case,are known as *features*.



(a) Class A **(b)** Class B **(c)** Plot of features

Figure 1-1: (a) Image region corresponding to class A, **(b)** Image region corresponding to class B and **(c)** Plot of the mean value versus the standard deviation for a number of different images originating from class A (circle) and class B (plus). A straight line separates the two classes [2].

Generally, l features $x_i = 1, 2, ..., l$ form the *feature vector*

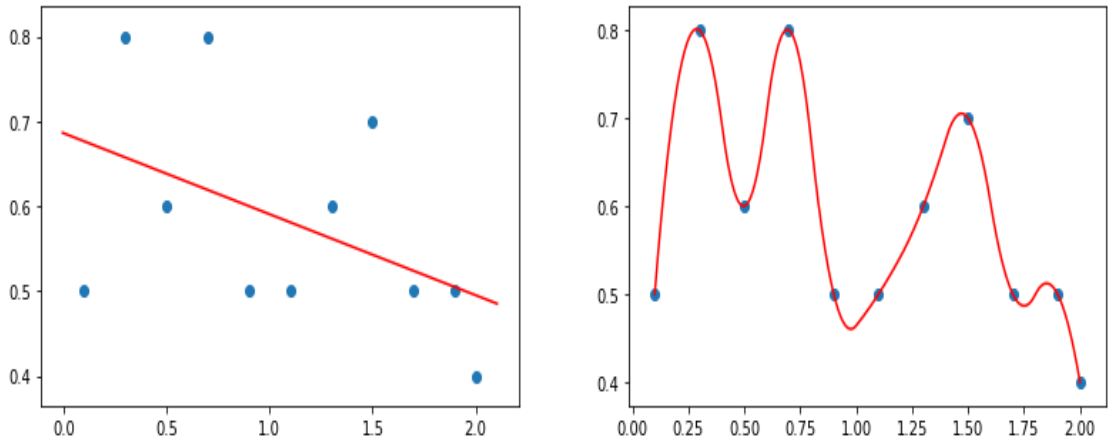
$$\mathbf{x} = [x_1, x_2, ..., x_l].$$

Each of the feature vectors identifies uniquely a single pattern object. The straight line in Fig. 1-1(c) is known as the *decision line*, and it constitutes the *classifier*, whose role is to divide the feature space into regions that correspond to either class A or class B. If a feature vector $\mathbf{x}$, corresponding to an unknown pattern, falls in the class A region, it is classified as class A, otherwise as class B. If the decision is not correct, then *missclassification* has occurred. The patterns (feature vectors), whose true class is known and which are used for the design of the classifier, are known as *training feature vectors*, constituting the *labeled data*. The classifier, performing slightly better than the random guessing, is called *weak learner* and the one, achieving high efficiency, is called *strong learner*.

The issues, required to be handled in a classification task, are the below:

- The *feature generation stage* concerns the selection of features, according to the given classification task.
- The *feature selection stage* concerns the number l of features.
- The *classifier design stage* includes the type of required classifier and the design of classifier, based on an *optimality criterion*. The linear classifier does not always have the best performance and the type of the non-linearity of the classifier should be identified.
- Finally, the *system evaluation stage* undertakes the task of the performance assesement of the classifier, through estimation of the classification error rate.

In the previous example of image classification, we assumed that a set of training data were available, and the classifier was designed by exploiting this a priori known information. This is known as *supervised learning*.



(a) Underfitting **(b)** Overfitting

Figure 1-2: A simplified example for the bias and variance error in prediction models, leading to overfitting and underfitting.

As any prediction model, supervised machine learning algorithms have prediction errors (bias and variance). *Bias* is the difference between the average prediction of our model and the actual prediction. Examples of high bias ML algorithms are Linear Regression, Linear Discriminant Analysis and Logistic Regression, since they cannot capture the characteristics of all the training data, but capture a simplified relation between the features. Low bias algorithms are Decision Trees, k-Nearest Neighbors and Support Vector Machines. *Variance* is the variability of model prediction for a given data point or a value. Even if high variance algorithms perform very well on training data set, this may lead to bad predictions as far as test data are concerned. Non linear algorithms and decision trees have high variance.

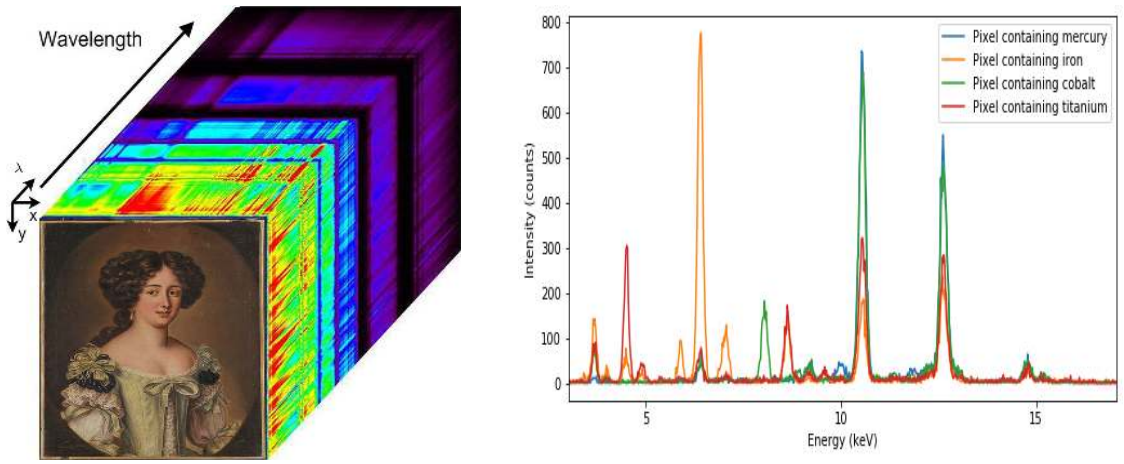
In Fig. 1-2, a simplified example of high bias and variance is presented. High bias and low variance algorithms can lead to *underfitting*. This problem arises with limited training data

and the use of linear algorithms to describe non linear data. On the other hand, low bias and high variance can lead to *overfitting*. In this case, we may capture the noisy data set together with the underlying pattern in the data. Therefore, a trade-off between bias and variance is necessary.

In high dimensional data, the training data are usually linearly separable. Then, we can select two parallel hyperplanes that separate the two classes of data, so that the distance between them is as large as possible. The region bounded by these two hyperplanes is called the *margin*, and the maximum-margin hyperplane is the hyperplane that lies halfway between them. This hyperplane defines the *maximum margin classifier* or else *hard classifier*. In overlapping data, some data are allowed inside the margin, so that the best fit is achieved. In this case, the classifiers are called *soft margin* classifiers.

However, training data are not always available, and there is another type of machine learning tasks for which training data, of known class labels, are not available. In these cases, we are given a set of feature vectors $\mathbf{x}$ and the goal is to unravel the underlying similarities and *cluster* (group) similar vectors together. This is part of the unsupervised learning and is called *clustering*.

In this introduction, we will focus on presenting the kind of data, usually found on the investigation of paintings, the hyperspectral data, as well as defining clustering and matrix factorisation. Fig. 1-3 (a) shows an hyperspectral cube, where each pixel of the image has a spectrum, characteristic of the elemental character of the material in this pixel. Hyperspectral imaging is defined as simultaneous acquisition of spatially coregistered images, in many spectrally contiguous bands. Therefore, it measures continuous spectral bands, as opposed to multispectral imaging which measures spaced spectral bands. Fig. 1-3 (b) presents the full spectrum of four pixels. The first contains high concentration of mercury, the second iron, the third cobalt and the last one titanium.



(a) Hyperspectral cube **(b)** Spectra of selected pixels

Figure 1-3: (a) A painting of unknown artist mapped with X-ray fluorescence (XRF). Each of the 389×290 pixels with 2048 channels has a characteristic spectrum. **(b)** A plot of the XRF spectra of four pixels that contain high concentration of an important element (mercury, iron, cobalt, titanium).

In terms of matrices, the data matrix D, containing n columns (pixels) and d rows (spectral channels) can be decomposed into the matrices W and H. The basis vectors W (loads, endmembers, archetypes) contain the spectral information in a matrix of k bases by d channels, where k<d. The coefficients in H (scores) are represented by a k bases by n pixels, containing the lateral information that corresponds to the k bases. The selection of the remaining data is related with the goal of the data processing, but generally its minimization is recommended. According to the Factor Analysis, the residues are represented by a matrix d by n:

$$R^{dxn} = D^{dxn} - W^{dxk} H^{kxn} \tag{1-1}$$

The complexity of the data set is reduced ($k < d$), in order to increase the computational performance and mainly to evaluate the data through a representation, understandable by humans. According to the way of determining the matrices W and H, different approaches for matrix factorization can be applied, with the most popular one in the cultural heritage field being the Principal Component Analysis (PCA). PCA, as well as other dimensionality reduction techniques, are also used for clustering.

A clustering algorithm can be employed to reveal the groups in which feature vectors are clustered in the l-dimensional feature space or the reduced space. For instance, the groups are formed based on the chemical character of the pigments or the compositional similarities in a painting.

Two issues of unsupervised machine learning are the choice of an appropriate way to measure 'similarity' between feature vectors and of the proper algorithm or algorithmic combination based on the similarity measure.

Semi-supervised learning can be used either for classification or clustering tasks. In the case of classification tasks, set of patterns of unknown class origin (unlabeled data), in addition to the training patterns (labeled data),whose true class is known. If we have limited labeled data at hand, it is important to recover additional information from the unlabeled samples, related to the general structure of the data. In the case of clustering tasks, labeled data are used as constraints. The clustering task is constrained to assign certain points in the same cluster or to exclude certain points of being assigned in the same cluster.

1-3 Aim and structure of the thesis

This thesis aims to develop and optimize methods for combining spectral and spatial information during the investigation of hyperspectral data from the field of material science and cultural heritage. The current studies on spectroscopic images are based only on the exploration of the spectral contribution of the data set and ignore the spatial information. With the goal of exploring more challenging and complex XRF data sets, the spectral information is combined with the spatial information, which is presented either in the form of photograph or in the form of the neighboring pixels of the XRF data sets.

Chapter 1 presents the importance of machine learning during the evaluation of data from the field of cultural heritage and introduces some basic information about machine learning methods. Chapter 2 is a literature review of conventional methods, applied in this thesis and of spectral/spatial data fusion approaches, found in the current bibliography. Conventional

methods for hyperspectral data processing with supervised machine learning, such as typical artificial neural networks and convolutional neural networks, are presented. The dimensionality reduction and clustering methods, performed in this thesis, are briefly explained together with the relevant bibliography. Chapter 3 presents the four examined data sets and describes the developed methods for combining spectral and lateral information for the examination of the titanium alloy's and paintings XRF data sets. Chapter 4 contains the results of applying the advanced methods to the above data sets. Some of these results are compared with the results obtained by classical methods or by simulated data sets. The presentation of the results is followed by their interpretation and discussion. Finally, in Chapter 5 the conclusions and suggestions for future research are presented.

Chapter 2

Literature review

2-1 Conventional methods for hyperspectral data processing with supervised machine learning

2-1-1 Artificial Neural Networks

At the core of artificial neural networks, we can find the *neurons*, or else nodes or units. Each neuron receives the inputs, multiplies each input by a parameter, called weight, sums the values of the weighted inputs and inserts the value into an activation function. As shown in Figure 2-1 [3], neural networks have an input layer with n nodes, such as an observation with n features, and an output layer with nodes representing the classes. Between the input and the output layer, we have the hidden layers that transform the features values from the input layer to something that, one processed by the output layer, it resembles the target class.

The weights of artificial neural networks are initialized to small random numbers. In forward propagation, the initial observations propagate through the hidden layers and give output value or values, which are compared with the true values of the observations through a loss function. The process of *backpropagation* refers to the propagation of an algorithm backwards to identify the degree that each parameter contributed to the error between the true and predicted values. The weights are adjusted according to an optimisation algorithm, called gradient descent. *Epoch* is the process of one forward and backward propagation. The training consists of multiple epoques. Obviously, the input observations are training data, since in order to train the algorithm to classify correctly an unlabelled observation, we need to optimise its parameters based on known observations. Therefore, we create a structure that successfully labels unknown data.

2-1-2 Convolutional Neural Networks

A *convolutional neural network* is a special type of deep learning network, designed to be used on spatial data and consists of three types of hidden layers: the convolutional layers, the pooling layers and the fully connected layers.

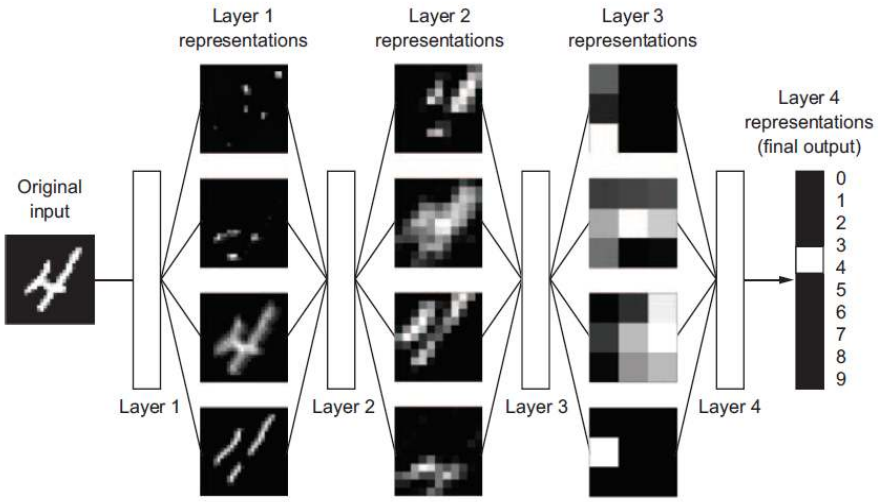


Figure 2-1: Illustration of a digit-classification model. The original input (input layer) is a number that needs to be recognised by the system. The model has three hidden layers, which are representations of the original input. The output layer consists of 10 neurons, that correspond to the numbers from 0 to 9.

Convolutional kernels are the filters, contained in the convolutional layers, that when applied to the input layer (usually image), they produce the output feature map. Filter is a matrix of integer numbers applied to a subset of the input image. Each pixel of the subset area of the input image is multiplied by the corresponding integer number of this matrix, that is the convolutional kernel. Then the values of all pixels of the subset image are summed up, giving a single value to this subarea. The filter is then moved to another subset area of the input image and repeats the same process until the whole input image is filtered. Stride is called the magnitude of the glide of the filter between the different areas of the input images. In this way, the output feature map is created.

The advantage of convolutional kernels is that they capture the spatial and temporal dependencies in an image, since they do not simply transform the image into an one column feature vector. The integer values of the matrix of the filter constitute the weights, which with proper training capture properly the spatial information. Figure 2-2 (a) presents the *valid padding*, where a kernel filters an area of the input image with the same size. The convoluted matrix will be reduced. On the contrary, *same padding* refers to the filtering of an augmented initial image with kernel such that the convoluted matrix has the same dimensions as the input image. Furthermore, the convolutional layers can be more than one. The first convolutional layer captures the low level features of the image, such as the edges and the color, while additional levels capture high level features, creating a more detailed representation of the input image. However, the increase of layers is translated to higher computational cost.

Max pooling (Figure 2-2(b)) returns the maximum value from the portion of the image covered by the Kernel, while *average pooling* returns the average equivalent. The benefits of pooling is the reduction of the spatial size of the convolved features and subsequently the decrease of the computational power required for data processing. Even if the dimensionality is reduced, the effective training of the model is maintained. The extracted dominant features are rotational and positional invariant. This effect is important because when we manipulate the data

structure to include spatial information, it is not straightforward to ensure that the features are rotational and positional invariant.

Finally, the *fully connected layer* (Figure 2-2(c)) is the same as in the artificial neural networks. The neurons are fully connected to all the the activations from the previous layers. Specifically, the representation of the input image after padding and pooling is flattened into a column vector. The flattened output is fed to a feed-forward neural network and backpropagation applied to every iteration of training.

The most recent and critical applications of convolutional neural networks (convets), related to the focus of this literature review, will be presented. Specifically, the presented papers are about convets for the processing of spectral data and convets for the examination of paintings.

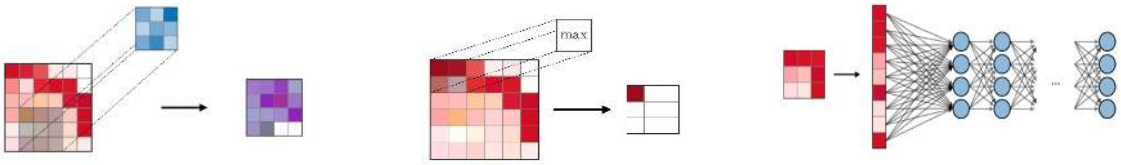
The recent advances in hyperspectral image proccessing with convets focus on hyperspectral image classification. The usual case is the use of spectral images and not raw spectral data. This approach has the disadvantage of loss of spectral information, since different spectra may produce similar images of different objects and create a false similarity of spectral features between these different objects. The weaknesses of traditional methods, such as random forests and SVMs, compared to convets are the difficulty in spatial-spectral feature extraction for big data and the superficial feature extraction of hyperspectral images.

Han *et al.* [4] presents the most advanced effort so far for joint spatial-spectral hyperspectral image classification, based on convets. The main challenge, associated with convets for hyperspectral image classification, is the absence of a commonly-used methodology for correlation between the spectral and spatial information, as well as the lack of the training samples. Both these obstacle are being addressed in the paper. As can be seen from the Figure 2-3, a two-stream convolutional network is designed to learn the spatial-spectral features of the different-scale spatial pixel blocks around the center pixel. By selecting the blocks with sizes of 3x3 and 5x5 around the center pixel containing all the hyperspectral bands, more spatial information is included. The problem of limited data is addressed by a spatial enhancement method. This mode of data augmentation entails the insertion of spatially rotated and row-column transformed hyperspectral data into a two-stream convolutional network.

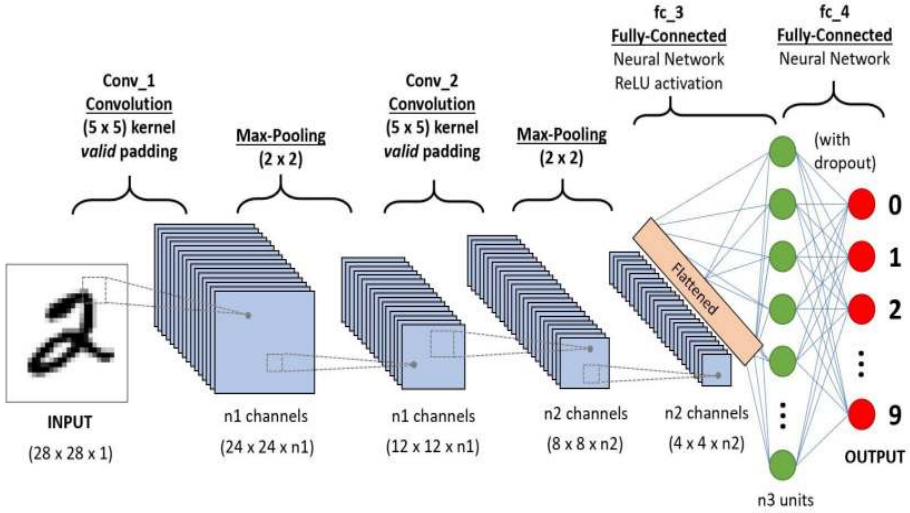
Deep neural network technology has been applied to datasets from electron energy loss spectroscopy (EELS), scanning probe microscopy [5] and Fourier-transform infrared spectroscopy (FTIR) [6].

Chatzidakis *et al.* [7] use electron energy-loss raw spectra of Mn^{2+}, Mn^{3+} and Mn^{4+} to test and train a convolution neural network to recognise the existence of manganese and their different oxides in samples. The extension of this study can be beneficial for the examination of pigments in paintings. The accurate identification of different oxides can discriminate between retouches and original paint. The advantage of CNNs is that they are automated feature extractors and compared to texture descriptors for feature extraction entail less human bias. The novelty of this paper lies on the achievement of translation invariance. This is an important feature, since a practical classifier would be one trained from one data set acquired from one spectrometer but potentially tested on another instrument with different calibration.

Oviedo *et al.* [8] propose a fast and interpretable classification of small X-ray diffraction datasets using data augmentation and deep neural networks. The paper presents a methodology of data augmentation of raw XRD data in the frame of convets.



(a) Padding. **(b)** Pooling. **(c)** Fully connected.



(d) Demonstration of a convolutional neural network for the classification of digits.

Figure 2-2: An illustration of the three main stages of a convolutional neural network. **(a)** During valid padding or no padding, a kernel 3x3 is applied to a 3x3 subarea of a 6x6 input image, generating a 4x4 output feature map. **(b)** During max pooling, the 6x6 matrix is reduced to a 3x3 output feature map by considering the maximum value of each of the 6 subareas with 4 pixels. **(c)** After valid padding and pooling, we have a layer whose neurons are fully connected with all the activations from the previous layers. The fully connected layer functions as the layers of an artificial neural network. **(d)** Demonstration of a convolutional neural network for the classification of digits with two convolutional and two pooling layers. In this example, an activation function ReLU (rectified linear unit) is applied to the values of the previous layers to provide non negative input data for the fully connected layer, by zeroing the negative values [4].

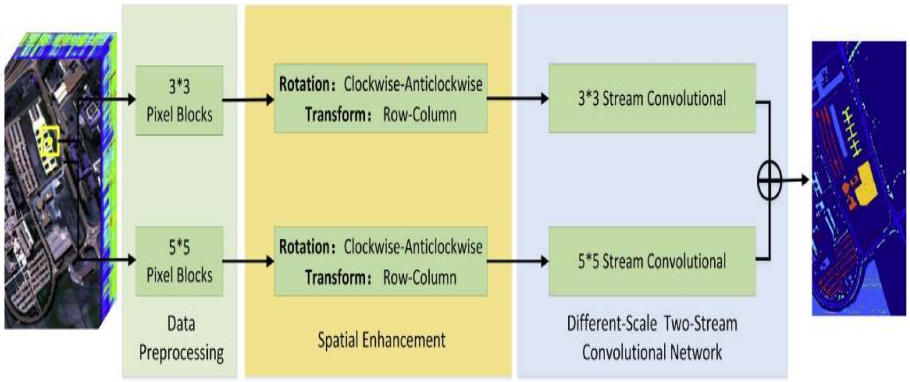


Figure 2-3: The overall structure of the method for hyperspectral image classification with convolutional neural networks. Spatial information is included by selecting blocks with sizes of 3x3 and 5x5 around the center pixel containing all the hyperspectral bands. Data augmentation is achieved with insertion of spatially rotated and row-column transformed hyperspectral data [4].

Finally, the paper of Zenf *et al.* [9] is a recent example of a convolutional neural network for pixel-wise reconstruction of paintings. Specifically, the article investigates the reconstruction of degraded drawings of Van Gogh. Reproductions are used as reference information for the original appearance of drawings. The advantage of convets over traditional methods, such as k-nearest neighbours, is their better performance in learning colour information and the use of only some of the pixels in training dataset.

2-2 Conventional methods for hyperspectral data processing with unsupervised machine learning

After reporting the literature review for the supervised classification, now the unsupervised case is examined.

Dimensionality reduction deals with the inability of processing effectively data, due to the big size on the feature space. The dimensionality reduction algorithms project high-dimensional data to a low-dimensional space, retaining as much of the salient information as possible, while removing redundant information. As a result, more patterns are identified effectively because noise has been reduced. Clustering compares how similar the data for one observation is to data for other observations and groups.

In this chapter, the most important dimensionality reduction techniques, Principal Component Analysis (PCA), Indepedent Component Analysis (ICA), Non-negative matrix factorisation (NMF) and Simplex Volume Maximisation (SiVM), are presented. Additionally, the most popular and promising clustering algorithms, Hierarchical Clustering Analysis (HCA), hard clustering, t-distributed stochastic neighbor embedding (t-SNE) and self organising maps (SOM), are introduced with the relevant literature review.

For a better explanation of clustering and dimensionality reduction techniques, we use the XRF or RIS experimental data of two paintings. The main data used for the examination of the painting of Figure 2-4 (left image) [10] were extracted by its XRF signal. This elemental analysis determines the age of the painting between late 17th to early 18th century. Knowledge

Figure 2-4: Both images are photographs of paintings. The left one is a 17th to early 18th century AC easel painting of unknown artist. The right one is a 13th century BC mural in the tomb of Nakhtamun, showing the deceased (left) negotiating with three gods for his passage into the netherworld [10].

of the historical context allows the identification of the painting as a portrait of Hortense Mancini, Duchesse de Mazarin. The absence of underdrawing fits the claim that the examined painting is probably a copy of another portrait. However, a more detailed experimental and data analysis work can reveal the particularities of the pigments and painting methods of the artist.

The second painting, Figure 2-4 (right image), is a mural in a Ramesside tomb (13th century BC) of the Theban Necropolis in Egypt. Both RIS and XRF data were acquired and t-SNE offered an effective way of presenting interprentable data.

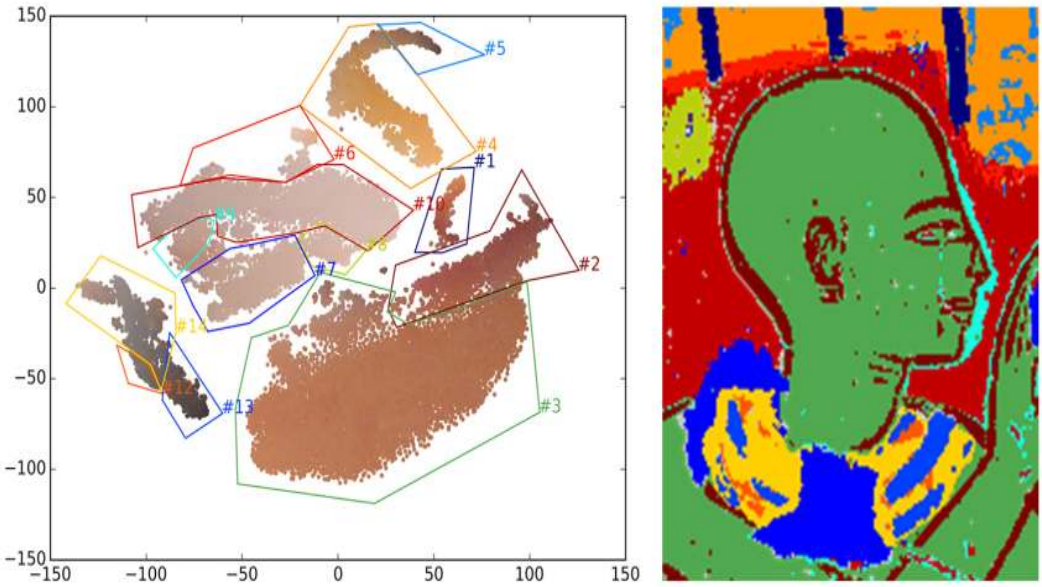


Figure 2-5: Scatter plot of t-SNE representation with manually selected clusters to the left and the corresponding cluster image to the right [10].

2-2-1 Clustering algorithms

t-SNE

t-distributed stochastic neighbor embedding (t-SNE) is a technique that visualises high dimensional data by creating a single map of datapoints, that reveal the structure at multiple scales. Van der Maarten and Hindon show that the visual performance of t-SNE is higher than many other visualisation techniques [11]. t-SNE, as a dimensionality reduction technique, uses repulsive and attractive forces between datapoints to preserve the original high dimensional clusters.

Visualising datapoints with t-SNE is based on three important concepts: the *perplexity parameter*, the *matrix of similarity scores* and the *student's t-distribution function*.

Initially, the similarity between all datapoints of the same cluster is calculated by projecting the datapoints into the low dimensional space through the t-distribution. Since the t-distribution will be wider for different clusters, the similarities should be properly scaled. The density or scarcity of the datapoints in an area is portrayed by the perplexity parameter, which reflects the effective number of neighbours around a datapoint. The use of the t-distribution deals with the issue of outliers or remote clusters. The benefit of t-SNE over SNE is that at the same time considers the effect of outliers and the effect of distances between datapoints in the low dimendional space. Dissimilar points with small pairwise distance strongly repel each other, but not infinitely.

Based on this article [11], many efforts to optimise, mainly the computational complexity of $O(n^2)$ with n the number of datapoints, have been undertaken. Van der Maarten accelerated t-SNE by using the variant of the Barnes-Hut algorithm to approximate some features of t-SNE (BH t-SNE), requiring $Nlog(N)$ computations [12]. Altered formulation of t-SNE, such

as Hierarchical Stochastic Neighbor Embedding (HSNE) for visual analysis of mass cytometry [13], are implemented to satisfy the demands of different fields. Linderman et al. [14] present the fastest implementation of t-SNE, called Fast interpolation-based t-SNE (FIt-SNE), with computation time reduced by at least ten times compared to BH t-SNE.

The benefits of data visualisation with t-SNE have been exploited by Alfeld *et al.* [10], applied to RIS and XRF fused data. Specifically, the XRF image with the most dominant elemental peaks was used as a reference image for the data fusion. A feature detection algorithm, the scale-invariant feature transform (SIFT), was used for the fusion of the reference image and the RIS image and a subsequent fusion of data. t-SNE, applied to fused data, provides a data representation, which contains the visualisation of information from different techniques without the need of image alignment.

Figure 2-5 [10] is a t-SNE representation with manually selected clusters of fused data. This representation highlights Egyptian painting techniques. Three stages can be identified: (a) Fe- containing pigments are used for the skin and the yellow background is created, then (b) white paint, probably, calcite is applied to the background, covering part of the dark (cluster 9) and of the yellow paint (cluster 6) and (c) Fe- based lines (dark red) are painted.

Another technique that presents promising characteristics for the analysis of fused data is the Self Organising Maps (SOMs), discussed later in this Chapter. In a future work, the performance of SOMs and t-SNE to fused data can be analysed and compared.

Self organising maps

The self-organising maps are competitive-cooperative learning algorithms. The main concept of SOMs is the below:

- SOMs consist of nodes, each of which is represented by the vector of weights and its position in the map space.
- Initially, the values of the vectors are assigned randomly.
- The neuron map is trained through an iterative process from the input data. Specifically, at each step, a single spectrum is taken from the input data and the distance between the input datapoint and all the neurons is calculated.
- The overall Best Matching Unit (BMU)- the node with the smallest distance from all calculated ones- is the winning neuron and adjusts its weight according to the input data. At the same time the matrix is adjusted by altering accordingly the weights of the neighbours of the BMU.
- This whole process is repeated for as many times as the representative subset of spectra. Representations as many as the spectra are created and therefore all spectra are mapped to the trained SOM.

These depictions are related with elemental representations since in certain wavelengths, specific elemental peaks dominate. The corresponding representation reveals the place (nodes) of a selected element. Additionally, a cumulative trained SOM indicates the nodes that describe the majority of spectra.

The SOMs were introduced by Kohonen in the 1980s [15]. The analysis of LIBS data with SOMs is common in geochemistry. In 2015, Pagnotta et al. [16] compare two brass alloys with SOM representations based on LIBS data.

SOMs are applied to paintings and archaeological objects some decades later than their invention [17, 18]. Based on the quantity of silver attained with XRF from an Etruscan Gold Coin Collection, the coins were classified in groups with the self-organsising mapping method [17]. Pagnotta et al. chooses the SOM elemental representations of thousands of LIBS data, due to the rapid aquisition of the elemental results [18]. The same year chemical mapping of various geological samples is also successful with SOM [19]. Archaeology and geochemistry research seem to provide each other with information for LIBS and SOM algorithms.

Recently, increasing interest was shown for the application of SOMs in paintings and manuscripts. SOM can reduce the dimensionality of the XRF data set, by creating clusters that share the similar spectra. XRF spectra of a peruvian watercolour painting from the Getty Research Institute collection are investigated with SOM clustering ([20]). Observing the mean spectrum per cluster gives information about the materials of each cluster. Although X-ray images are two-dimensional representations of three-dimensional objects, this information can indicate layers and substrates. These indications are later confirmed with Raman spectroscopy in selected areas of the painting.

2-2-2 Dimensionality reduction algorithms

Principal Component Analysis

Principal Component Analysis (PCA) is the most popular and frequenly applied chemometric method, especially to the experimental data of paintings. In PCA, the scores are orthonormal and the loads orthogonal.

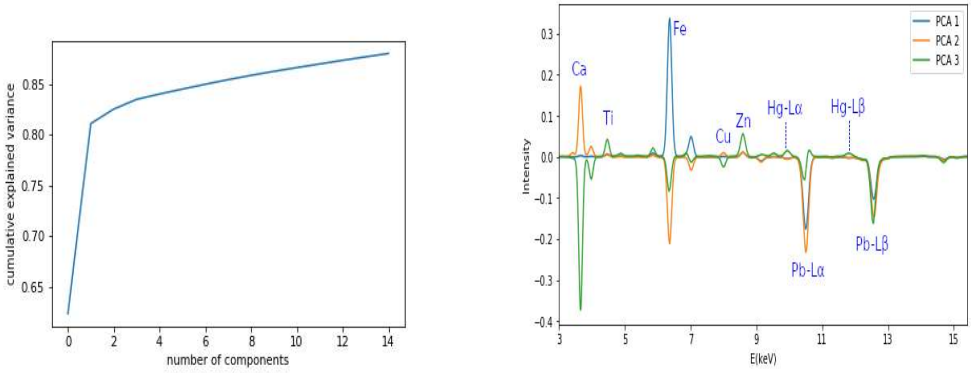
In Figure 2-6 (d), we can see the five representation given by the five first principal components covering 0.958 of the cumulative explained variance. This means that most features can be represented by the first four PCAs. However, high variance does not always coincide with high degree of information, since PCA can disregard in its first components the information of an area corresponding to only few pixels.

According to Figure 2-6 (e), the first principal component (PCA 1, Figure 2-6 (a)) has the rising edge of iron and lead is in the negative part of the representation. PCA 2 (Figure 2-6 (b)) has calcium in the positive part and lead with iron peaks in the negative one. PCA 3 (Figure 2-6 (c)) has among others vermillion, titanium and zinc in positive part and calcium in the negative part. This is in line with the Figure 2-6 (c), where we can observe dark blue spots on the forehead and certain areas of the dress.

There is a free software, PyMCA, that provides the elemental maps extracted from the experimental data, as presented in Figure 2-7.

The elements identified are sulphur, chlorine, arsenic, potassium, calcium, titanium, chromium, manganese, iron, cobalt, copper, zinc, arsenic, rhodium, tin, mercury and lead. The most amazing optical illusion in the painting is the color of the dress. It appears light green, even if there is not a single green pigment identified. It is actually a shadow of yellow, created by lead tin yellow type I (Pb_2SnO_4) on the top of black carbon ($CaPO_4$). The elements can

(a) PCA 1 **(b)** PCA2 **(c)** PCA3



(d) Quality of description **(e)** Loads

Figure 2-6: Representation of the XRF data by PCA. (a),(b),(c) The bases are shown as images and (d) the cumulative explained variances as a function of the first fifteen PCAs. (e) Plot of the corresponding loads.

be associated with certain pigments, based on the current bibliography (Table 2-1 and Table 2-2). The use of lead tin yellow for the dress and black carbon, containg calcium for the shadows of the dress, is well captured by PCA 3, as mentioned before. For the coloring of the cheeks and the lips, vermilion red (HgS) is chosen, an information also contained in PCA 3. The existence of potassium in the hair may suggest the presence of prussian blue, even if this is not highly probable.

Cerulean blue ($CoO \cdot nSnO_2$), titanium white (TiO_2) and zinc white (ZnO), shown in Figure 2-7, are colors of the 19th century, used as restoration pigments. There is also a possibility that gypsum ($CaSO_42H_2O$) is used locally for restoration purposes, due to the abundance of sulfur in the white upper lace of the dress near the chest. Sulfur is barely visible in the lips, due to the presence of vermilion (HgS).

Conclusively, XRF and microscopic analysis detects cerulean blue, titanium white, chromium pigment and zinc white, which are pigments of the 19th century. However, all of them are used for retouches. Gypsum, earth pigments, lead tin yellow, lead white, vermilion, azurite, calcite(chalk white) and bone black are pigments connected with the history of painting and constitute the main palette of the painting. Therefore, the examined painting is a portrait of Hortense Mancini, dated between late 17th to early 18th century. XRF and other analytical techniques allowed the identification of pigments- already used before the 18th century- as the initial (principal) pigments.

Independent Component Analysis

PCA creates features mutually uncorrelated to avoid information redundancies. For certain application, Independent Component Analysis (ICA) is superior. ICA assumes a linear combination of the base vectors and demands statistical independence and not simply uncorrelatedness of output data. With this demand more information can be exploited.

This independence may have some similarity with the physical function of neural outputs. The higher performance is only relevant for non-Gaussian random variables. Otherwise, there is no unique solution, as imposed by PCA through the orthogonality. Even if it responds to this physical reality, the base vectors as linear combination do not respond to the reality of the hyperspectral data in the VIS and NIR region.

In the case of the representations of Figure 2-8 (a), (b) and (c), new information about the painting is not revealed with ICA. Hence, we are presenting the loads for the 13th century BC mural.

From Figure 2-8 (g), we can make the below observations. ICA 1 (Figure 2-8 (d)) includes predominantly the information of iron in hematite (Fe_2O_3) and is very similar with the image of the elemental distribution of iron. ICA 2 (Figure 2-8 (e)) has copper, contained in $CaCuSi_4Si_4O_{10}$, in the positive part and calcium green ($CaCuSi_4Si_4O_{10} + Si4O_2$), also in the positive part, but with very low intensity. This difference of peak height create a contrast in the depiction of copper and green calcite. Finally, ICA 3 (Figure 2-8 (f)) contains the three dominant peaks of the XRF signal for the mural, calcium, iron and copper. This concentration of dominant peaks in one component does not help the better understanding of data.

Pigments	Chemical compound
White pigments	
Antimony white	Sb_2O_3
Chalk	$CaCO_3$
Gypsum	$CaSO_4 \cdot 2H_2O$
Lead white	$2PbCO_3Pb(OH)_2$
Lithopone	$ZnO + BaSO_4$
Permanent white	$BaSO_4$
Titanium white	TiO_2
Zinc white	ZnO
Yellow pigments	
Auri pigmentum	As_2S_2
Cadmium yellow	CdS
Chrome yellow	$2PbSO_4 \cdot PbCrO_4$
Cobalt yellow	$K_3[Co(NO_2)_6] \cdot 1.5H_2O$
Lead-tin yellow type I	Pb_2Sn0_4
Lead-tin yellow type II	$PbSn_{1-x}Si_x0_4$
Massicot	PbO
Naples yellow	$Pb(SbO_3)_2/Pb_3(SbO_4)_2$
Strontium yellow	$SrCrO_4$
Titanium yellow	$NiOSb_2O_3 \cdot 20TiO_2$
Yellow ochre	$Fe_2O_3 \cdot nH_2O(20-70\%)$
Zinc yellow	$K_2O \cdot 4ZnO \cdot 4CrO_3 \cdot 3H_2O$
Red pigments	
Cadmium red	$CdS + CdSe$
Cadmium vermillion	$CdS + HgS$
Chrome red	$PbO \cdot PbCrO_4$
Molybdate red	$7PbCrO_4 \cdot 2PbSO_4 \cdot PbMoO_4$
Realgar	As_2S3
Red lead	Pb_3O_4
Red ochre	Fe_2O_3
Vermillion	HgS

Table 2-1: Chemical compounds (second column) with the associated name of white, yellow and red pigments (first column) [21].

Pigments	Chemical compound
Green pigments	
Malachite green	$Cu_2CO_3(OH)_2$
Basic copper	$Cu_x(SO_4)_y(OH)_2$
Chromium oxide	Cr_2O_3
Chrysocolla	$CuSiO_3 \cdot nH_2O$
Cobalt green	$CoO \cdot 5ZnO$
Emerald green	$Cu(CH_3COO)_2 \cdot 3Cu(AsO_2)_2$
Gusgnent green	$Cr_2O_3 \cdot H_2O + H_3BO_3$
Verdigris	$Cu(CH_3COO)_2 \cdot Cu(OH)_2$
Blue pigments	
Azurite	$2CuCO_3 \cdot Cu(OH)_2$
Cerulean blue	$CoO \cdot nSnO_2$
Cobalt blue	$CoO \cdot Al_2O_3$
Cobalt violet	$Co_3(PO_4)_2$
Light cobalt violet	$Co_3(AsO_4)_2$
Egyptian blue	$CaO \cdot CuO \cdot 4SiO_2$
Manganese blue	$BaSO_4 \cdot Ba_3(MnO_4)_2$
Prussian blue	$Fe_4(Fe(CN)_6)_3$
Smalt	$Co - glass(K_2O + SiO_2 + CoO)$
Ultramarine (Lapis Lazuli)	$Na_{6-10}Al_6Si_6O_{24}S_{2-4}$
Blue verditer	$Cu(CO_3)_2(OH)_2$
Black pigment	
Antimony black	Sb_2O_3
Black iron oxide	$FeO \cdot Fe_2O_3$
Carbon or Charcoal black	$C(95\%)$
Cobalt black	CoO
Ivory black	$C + Ca(PO_4)_2$
Manganese oxide	$MnO + Mn_2O_3$

Table 2-2: Chemical compounds (second column) with the associated name of green, blue and black pigments (first column) [21].

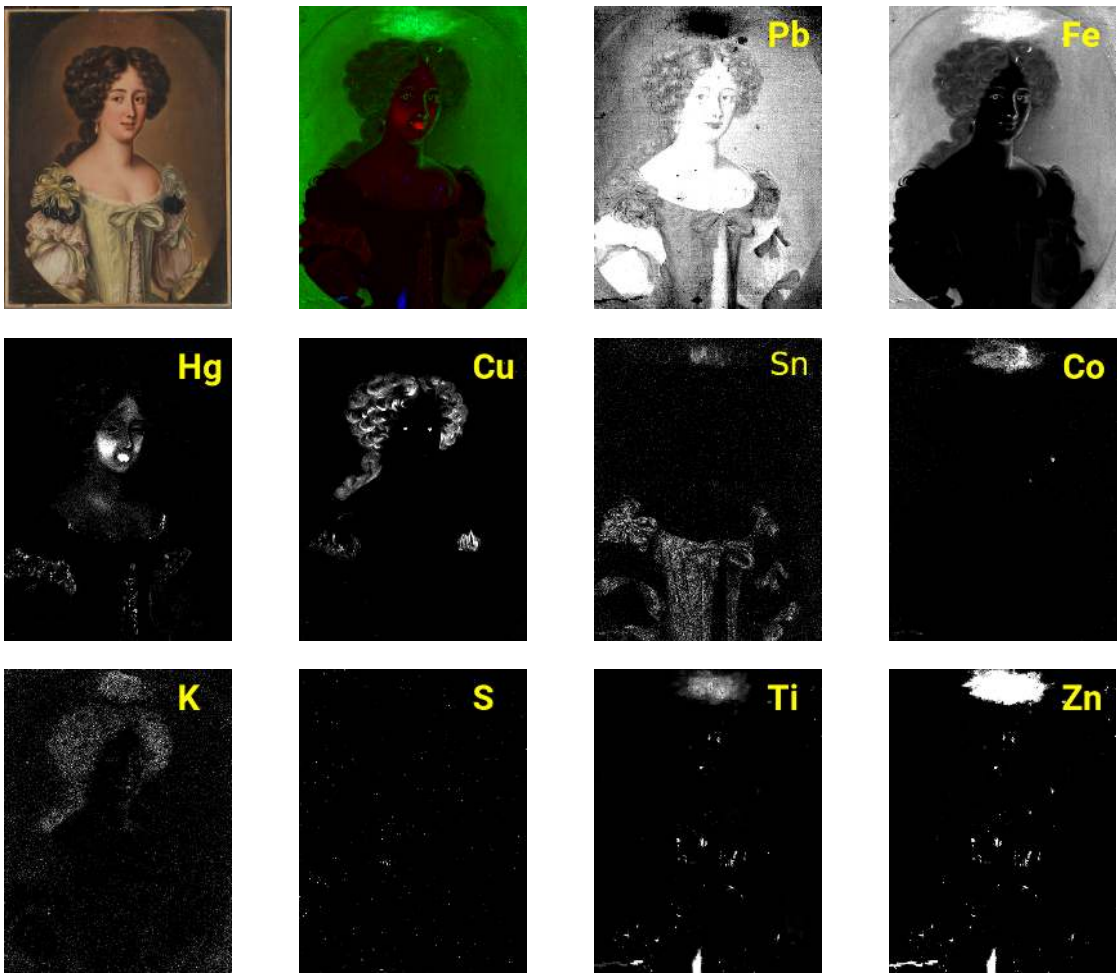
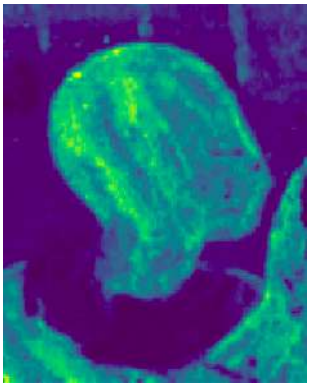


Figure 2-7: Elemental maps for the portrait of Hortense Manchini (unknown artist). Next to the photography of the painting, there is an elemental map of vermillion (red), iron (green) and titanium (blue). The rest of the images are the elemental distributions of lead, iron, mercury, copper, tin, cobalt, potassium, sulfur, titanium and zinc, respectively.

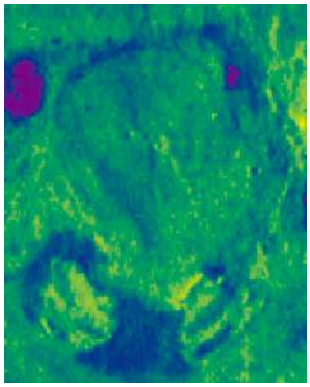
(a) ICA 1

(b) ICA 2

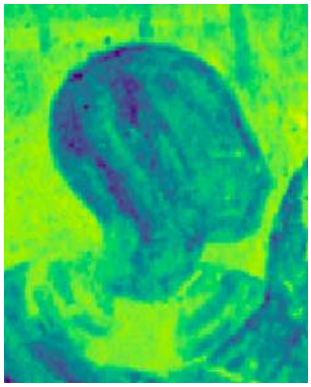
(c) ICA 3



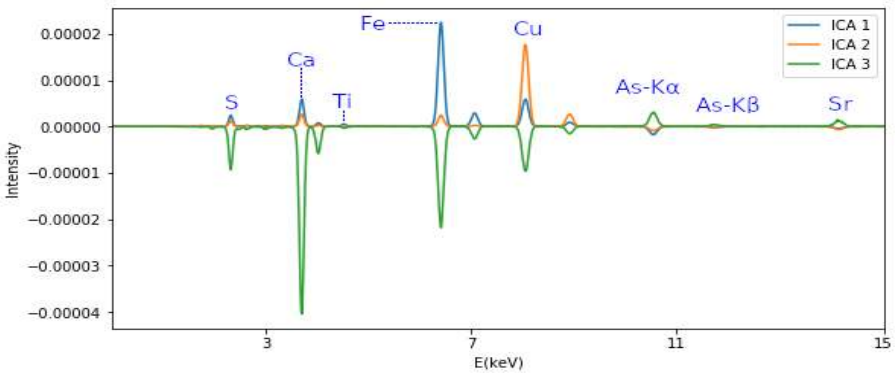
(d) ICA 1



(e) ICA 2



(f) ICA 3



(g) Loads

Figure 2-8: Representation of XRF data sets by Independent Component Analysis (ICA). (a), (b) and (c) Bases are shown as images for the portrait of Hortense Mancini and (d), (e) and (f) for the mural. (g) Plot of the corresponding loads for the mural.

The popularity of ICA has considerably increased in the cultural heritage field since the publication of the article concerning the digital image analysis to uncover the underwritten text in the Archimedes palimpsest by Salerno et al. [22]. However, a comparative article between blind Source separation algorithms reveals that the orthogonalization algorithm uncovers the hidden text in a painting [23]. The research around hidden details in ancient manuscripts and paintings is still in its original phase. The review of Tonazzini et al. [24] proposes, besides PCA and ICA, Fisher linear discriminant analysis (LDA), self-organizing maps (SOMs) and spectral angle mapping (SAM) for analysis of data from various experimental techniques.

Non-negative Matrix factorisation

Non-negative Matrix factorisation (NMF) guarantees the nonnegativity of the elements of the resulting matrix factors, more close to the physical reality of the reflectance spectra [25]. The computational cost increases significantly and the number of bases is not straightforward. As was with ICA, NMF has the advantage of the blind source separation [26] and as a result the K-M method can also be applied here. As far as face recognition is concerned, NMF has superior performance compared to PCA and ICA.

In the case of the representations of Figure 2-9 (a), (b) and (c), new information about the painting is not revealed with ICA. Hence, we are presenting the loads for the 13th century mural.

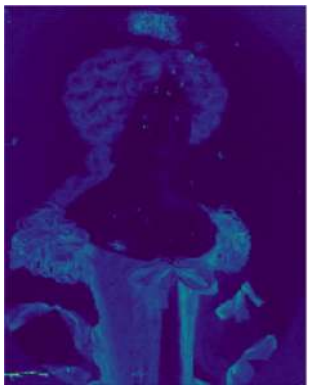
In this case, NMF analysis is superior compared to PCA and NMF, mainly due to two factors. The first one is the extraction of spectral information due to the preservation of the non-negative character of the data. The second one is that in this case, we get an efficient separation of pigments.

Qualitatively, we can say that the three NMFs (Figure 2-9 (h)) achieve a separation of the peaks describes in ICA 3 (Figure 2-9 (g)) . NMF1 has the calcium peak, NMF 2 the iron peak and NMF 3 the copper peak. NMF 4 has the arsenic peak.

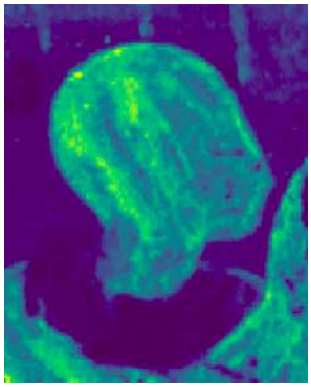
NMF is a fast and robust unsupervised technique, which can give insights in semi-known systems by supporting the interpretation of XRF imaging data. Alfeld et al. [27] used NMF to understand more in depth the transformation kinetics during continuous heating of a double phased titanium alloy. Finally, the manuscript, known as the Salterium, was illuminated by the data from the XRF and Raman spectroscopy. NMF highlighted hidden texts [28]. NMF analysis of Real-time MA-XRF imaging spectroscopy of the Virgin with the Child by Antonello de Saliba directed its future restoration [29].

(a) NMF 1

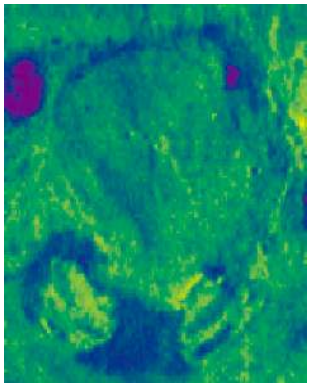
(b) NMF 2



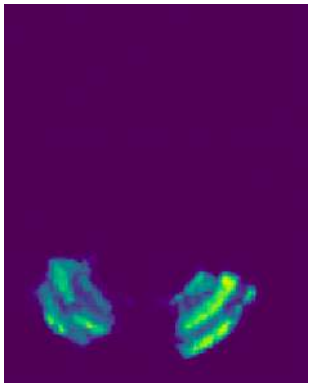
(c) NMF 3



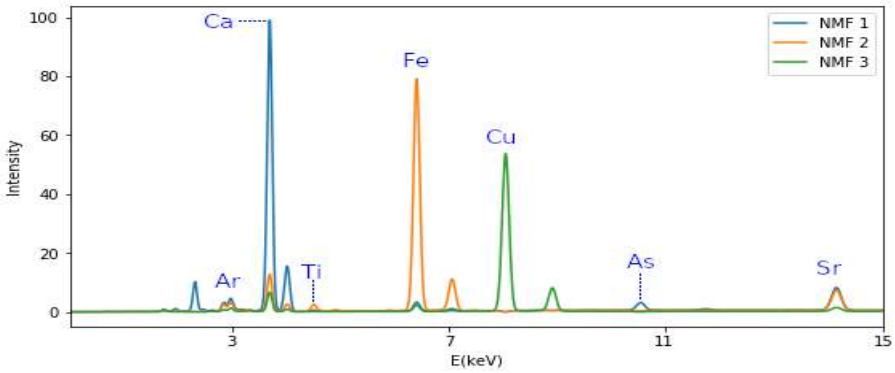
(d) NMF 1



(e) NMF 2



(f) NMF 3



(g) Loads

Figure 2-9: Representation of the XRF data by Negative Matrix Factorisation (NMF). (a), (b) and (c) Bases are shown as images for the portrait of Hortense Mancini and (d), (e) and (f) for the mural. (g) Plot of the corresponding loads for the mural.

2-3 Data fusion

Multi-modal imaging and data fusion treats the case of more than one data set. The terms *multi-modal imaging* and *data fusion* are used under different contexts.

Multi-modal imaging may refer to the parallel acquisition of different spectroscopic techniques. This technique offers fast acquisition of data. However, it presents one major drawback. The use of one instrument for many experimental tasks at the same time yield inferior results. Instruments that specialize in a specific spectroscopic task have higher performance.

Another more common practice is the acquisition of data separately evaluated but jointly interpreted. It does not refer to the combination of the raw data but the combination of the results in order to derive the final interpretation and conclusion. The fusion of algorithmic methods for the analysis of data tries to exploit with one algorithm the benefits of the standard algorithmic techniques.

The data fusion or else the pixel by pixel fusion refers to the separate acquisition but combined evaluation of fused data. When the datasets are from a different object-area by the same technique, the data fusion is called *horizontal data fusion*. When the datasets are from the same area with different techniques the data fusion is called *vertical data fusion*. The vertical data fusion is a field that is constantly developing and is highly promising. By fusion of the raw data, the acquisition of results is faster and the synergy effects of both spectroscopic techniques are exploited. However, there are two main disadvantages. The first one is the difficulty in aligning with a single pixel precision and derive an accurate representation of data. The other issue is the determination of weights on the spectral features of interest. The SIFT descriptors can occasionally present a reliable solution, but cannot be part of a common method in data fusion.

Another issue for discussion is the benefits of the fusion of the image information (intensity of pixels) and spectral information. Obviously, the feature vectors contain the spectral information, but the spatial information at least directly is not included in the feature vectors. Data augmentation can also refer to different types of data treatment. One possible data augmentation is the one where each feature vector is expanded in dimensions to include the spatial information. The former can take the form of pixel intensity or of neighbouring pixel's spectrum.

Combination of spectral and spatial data It was mentioned that t-SNE and SOM can capture some spatial information. In the previous Section, I presented a CNN for joint spatial-spectral hyperspectral image classification.

Instead of using another technique for capturing more information, neighbouring pixel data augmentation is a simple way of fusing spectral and spatial information [30] (*Method I*). As shown in Fig 2-10 [30], in neighbouring pixel augmentation, the spectral data of the neighbouring pixels are added to the spectral data of the central pixel. The benefits of this technique are the decrease of the signal to noise ratio and the revelation of more fine structures of the examined object. Therefore, instead of varying the ML techniques, we can alter the data structure and achieve better results.

Finally, there is another promising way to create augmented data sets, which include both spatial and spectral information. One way is to include the spatial information by exploiting

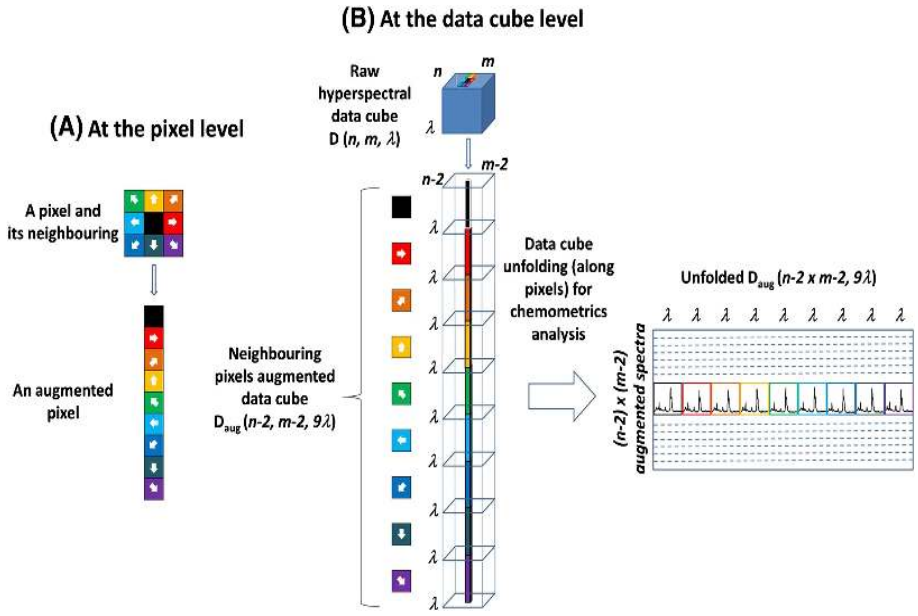


Figure 2-10: The spectrum of a pixel is usually represented by one cube. In the pixel data augmentation, each pixel is corresponded with nine spectra, since the spectra of the eight nearest neighbours of each pixel are considered.

the Wavelet transform algorithm (SWT). This algorithm uses low and high pass filters. For each elemental distribution image, four images of wavelet details are obtained: approximation, horizontal, vertical and diagonal coefficients. Spatial and spectral information is combined by creating an augmented data set with the original elemental distribution images followed by the images with the wavelet details [31].

Chapter 3

Material and methods

3-1 Data sets description

Four data sets are presented: a test data set, a X-ray fluorescence data set of a titanium alloy and two X-ray fluorescence data sets of oil paintings.

3-1-1 Test data set

The nine dimensional test data set with 60 × 60 pixels includes three main elements: a small rectangle, a big rectangle and a background (Figure 3-1). For instance, the first image consists of two areas with no information and a background with data from the normal distribution. This oversimplified data set serves the initial evaluation of the developed methods. The methods with interesting results for this data set are applied to real data sets, such as the XRF data set of the titanium alloy.

3-1-2 XRF data set of a titanium alloy

The Ti-6Al-6V-2Sn alloy with thickness 10mm and area 4 mm × 4 mm is heated with 5 °C min^{-1} and is quenched with helium at 530 °C. The concentration of minor components in the bulk material was given by the manufacturer as Al 6 wt%, V 6 wt%, Sn 2 wt%, Cu 0.75 wt% and Fe 0.35-1 wt%. The sample was investigated by micro-XRF at the microprobe end-station of beamline P06 at Deutsches Elektronen-Synchrotron (DESY) in Hamburg, Germany. The primary beam was 11.5 keV and the Vortex EM detector had active area of $50mm^2$. Areas of $10\mu m \times 10\mu m$ were scanned with $0.5\mu m$ step size and a dwell time of 0.2 s per pixel. Details on XRF imaging for this sample can be found in previous study ([10]) of the titanium alloy's XRF data set with Non-Negative Matrix Factorisation and Monte Carlo simulation.

In Figure 3-2 a), elemental distribution image of the sample for iron (Fe-K) for an area of 201 × 205 pixels is shown and the region of interest (36 × 40 pixels) is outlined. The sample presents a matrix of β phase, rich in vanadium, and α phase, rich in titanium. The two phases

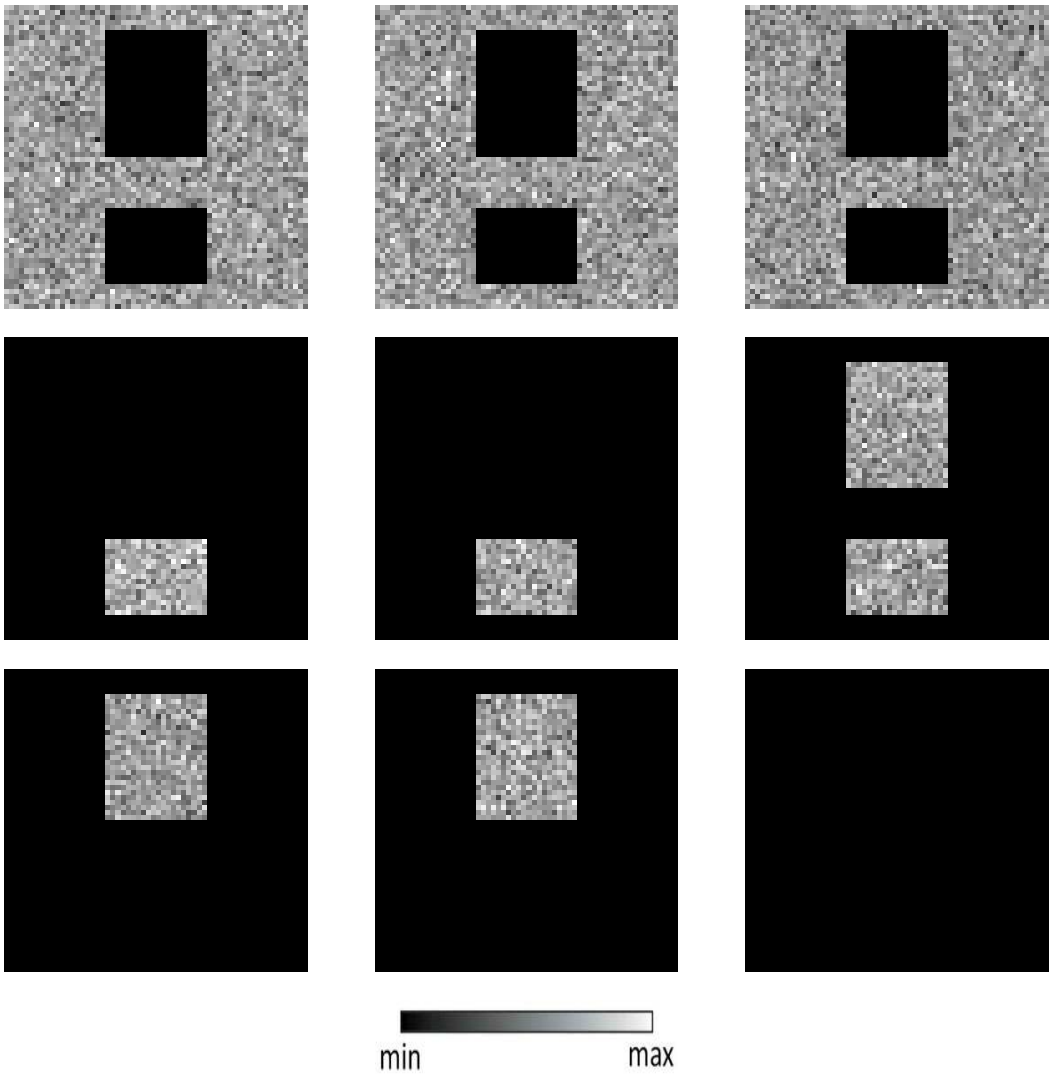


Figure 3-1: The test data set consists of images with 60 $\times$ 60 pixels. The main elements are one small square, one bigger square and a standard normal distribution.

are obvious in the Figure 3-2 b). The areas of high intensity Fe-K (light areas) are mainly attributed to β phase and the dark areas to α phase. The noisy appearance of the Ti-K is caused by variations of the intensity of the primary beam (Figure 3-2 c)). Theoretically, iron, vanadium (Figure 3-2 d)), nickel and copper (Figure 3-2 e)) are stabilisers of β phase, while titanium and tin are stabilisers of α phase. After normalizing each distribution image to the square root of its mean value (dividing each feature vector by Frobenius norm), the dimensions of the elemental distribution images are reduced to typical coefficient images obtained by NMF. Figure 3-2 g) and h) can be attributed to β and α phase, respectively. The interpretation of the other component (Figure 3-2 f)) is not straightforward.

The region of interest for the titanium alloy is simulated with Monte Carlo. The simulated data set, analyzed with NMF (Figure 3-3), confirms the nature of the first typical coefficient component (Figure 3-2 f)) as the presence of α above β phase. In summary, Base 0 represents the superposition of α on the β phase, Base 1 the α phase and Base 2 the β phase. The differences between real and artificial data set is the abundance of iron in the α phase and Base 0 also contributes strongly to the area, where β phase is found. Even if there are quantitative differences, the simulation confirms the presence of α phase on top of β phase.

3-1-3 XRF data set of an easel painting

After creating an oversimplified test data set and selecting a simple region of interest from a titanium alloy's data set, the conventional techniques and new methods are applied to XRF data sets from the field of cultural heritage. The XRF data set of the oil painting, the portrait of Hortense Mancini, consists of 389 pixels in height and 290 pixels in width. After fitting the data, 26 elemental images are created and the most important between them are shown in Figure 3-4. As mentioned in Chapter 2, the elemental distribution images do not reveal any hidden drawing or any hidden artistic composition. The restoration pigments are easily identified and contain cobalt, zinc, titanium and manganese to restore the damage all over the painting. The interpretation of the elemental image is straightforward. Hence, additional information can be extracted if the information of the high resolution image and the elemental distribution images is combined.

3-1-4 XRF data set of a mural panting

The last data set is an XRF data set of a mural's segment in a Ramesside tomb (13th century B.C.) of the Theban Necropolis in Egypt. The photograph of the mural painting shows the deceased on his way to the afterlife facing three gods [10]. The XRF data set is obtained by a XRF instrument with a Pd anode. It operated at 30 kV and 50 μA and a silicon drift detector with an active area of 25 mm^2 and a nominal thickness of 500 μm. The pixel size was 0.5 $\times$ 1.0 mm (h $\times$ v), a dwell time of 200 ms/pixel in 106 minutes and a size of the primary beam of about 1.1 mm. The details of the instrument [32] and the acquisition mode [10] are described in detail in the referenced articles.

The methods of this thesis project are applied only to a certain segment of the mural painting. The purpose of using the advanced methods for segments of this painting is not the overall examination of the painting. The main aim is to reveal hidden information or optimize the quality of information, obtained by segment of the painting, due to the application of

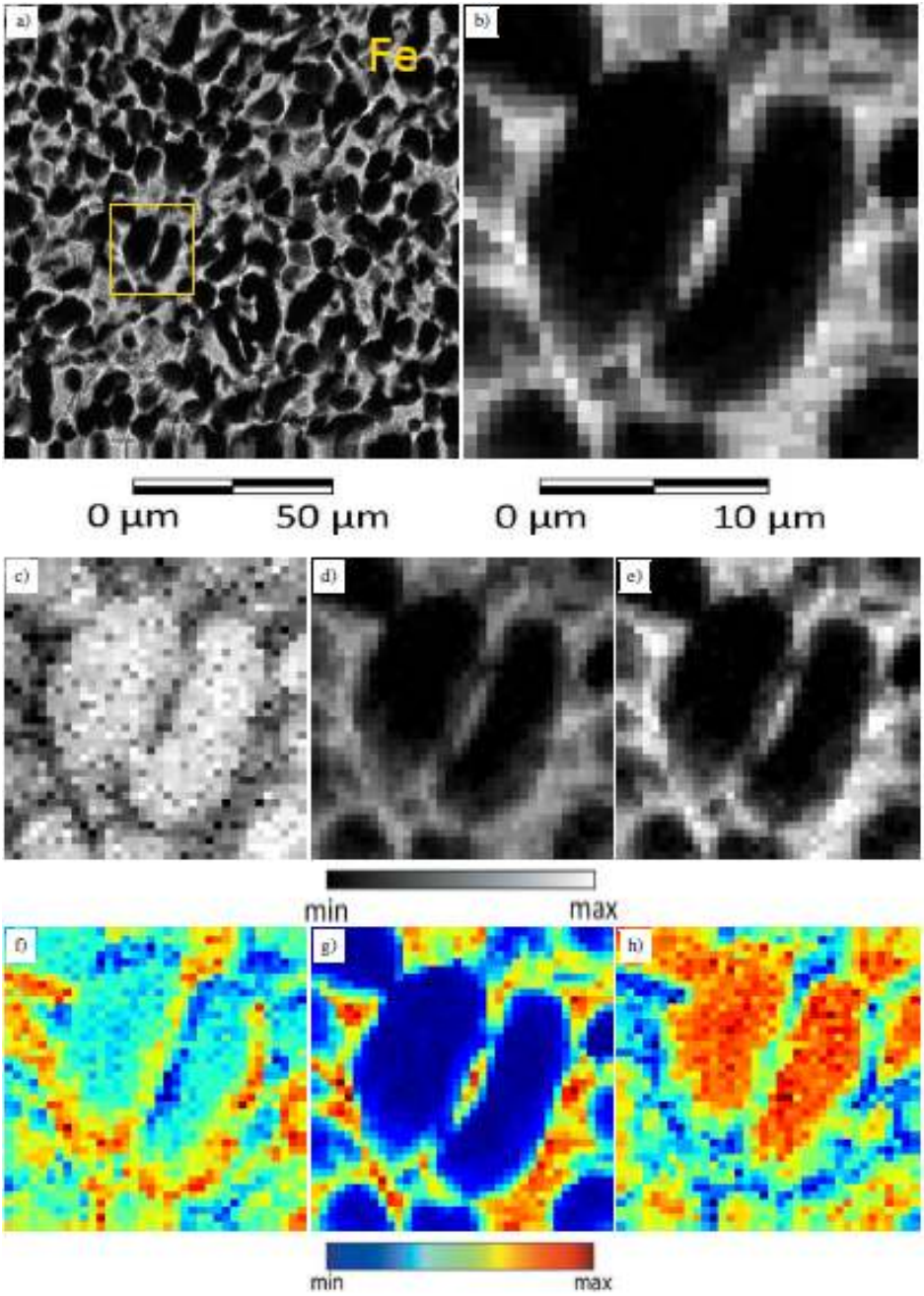


Figure 3-2: Elemental distribution image of 10 mm $\times$ 10 mm with a step size of 0.5 mm and a dwell time of 0.2 s of a titanium alloy sample (a) for the iron element. (b), (c), (d) and (e) Elemental distribution images of two primary grains for iron (Fe-K), titanium (Ti-K), vanadium (V-Kα) and copper (Cu-Kα). (f), (g) and (h) Typical coefficient images obtained by NMF from the data set of this sample.

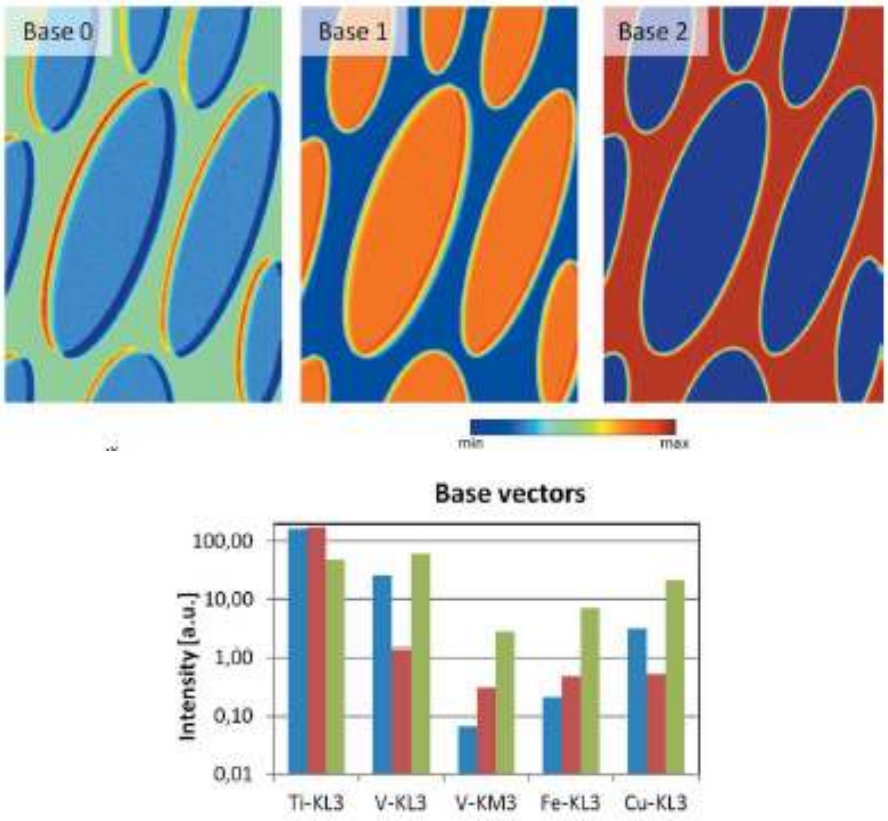


Figure 3-3: Top: Typical coefficient images from the artificial data set, based on the results of the Monte Carlo simulation. Bottom: Intensities of the spectra from the artificial data set [10].

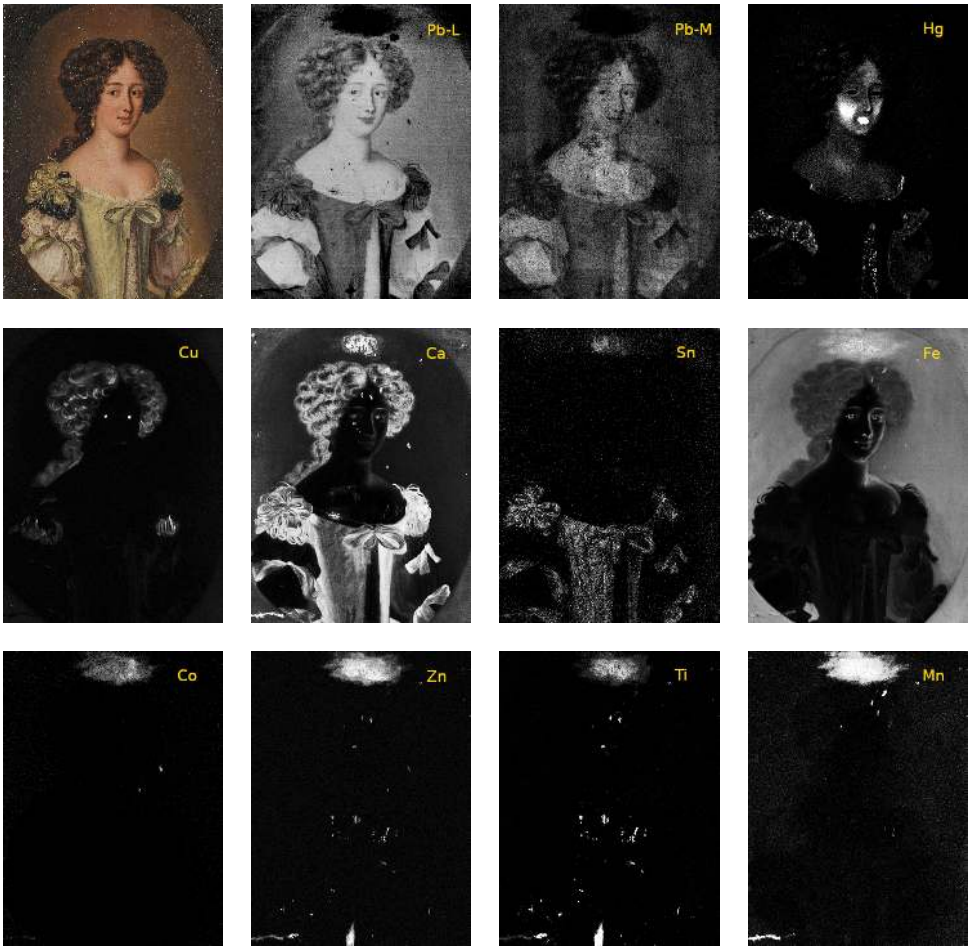


Figure 3-4: The first image is a photograph (h3880 $\times$ w3740 pixels and 60 $\times$ 73.2 cm) of the oil painting, followed by elemental distribution images (h389 $\times$ w290 pixels) for authentic pigments, including lead, mercury, copper, calcium, tin and iron. The last row images are elemental distribution images for restoration pigments, including cobalt, zinc, titanium and manganese.

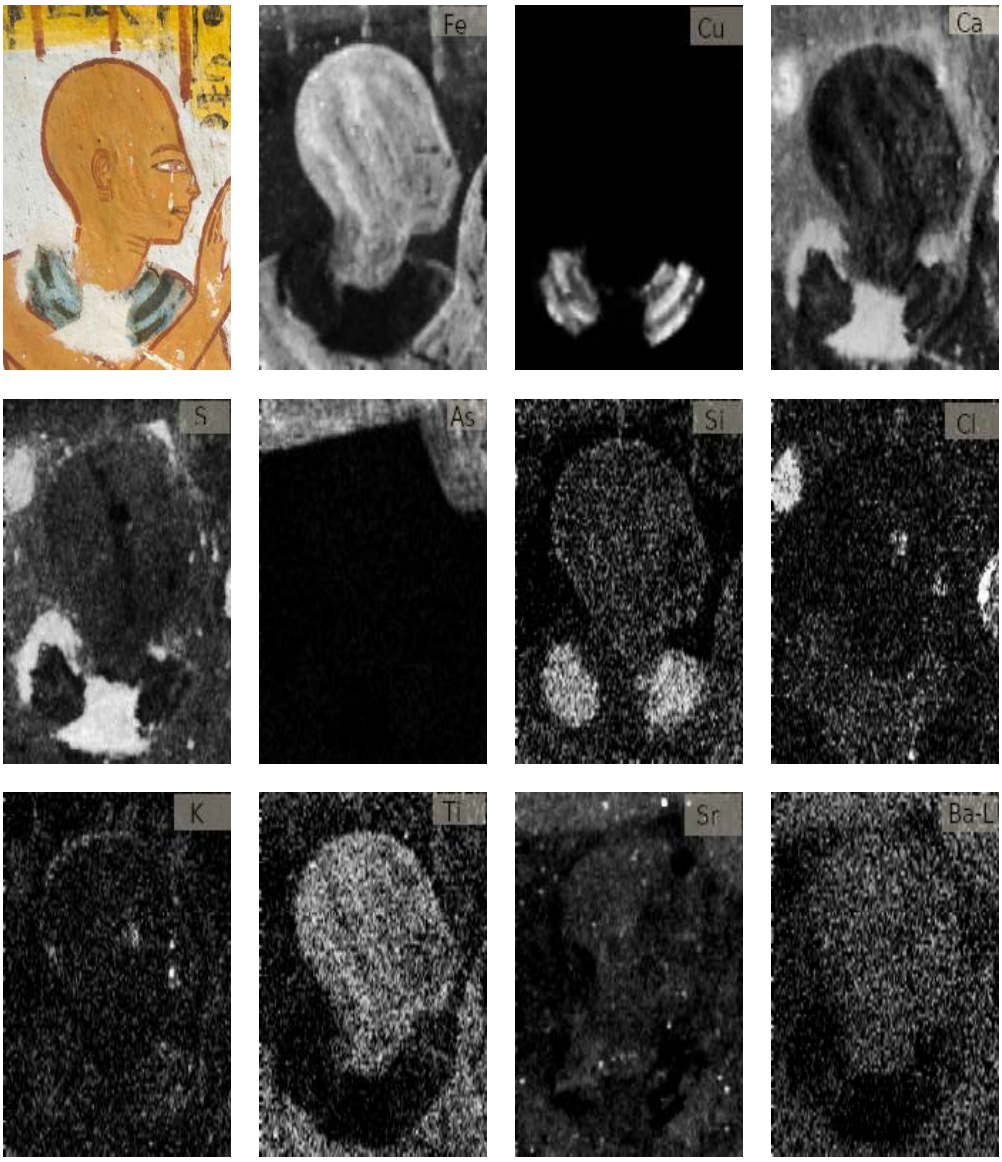


Figure 3-5: Photograph of the investigated segment of the 13th century B.C. Egyptian mural painting together with the elemental distribution images for authentic and restoration pigments.

advanced methods. The selection of this XRF data set is suitable for enhancing the amount of the information, obtained by SOM and t-SNE clustering. The painting contains iron- and copper- based pigments, not easily separable simply by the elemental distribution images.

Figure 3-5 shows the photograph of 13^{th} century B.C. mural's segment and its XRF distribution images. The site of data acquisition is the Theban Necropolis in Egypt from the tomb of Nakhtamun. The region of interest is investigated by XRF and RIS elsewhere [10]. In this thesis, three main issues of the mural painting are addressed:

- different pigments, which are not separable with a simple observation of elemental distribution images with XRF. This is explained by the fact that only some elements are detectable by the XRF instrument.
- the degree of damage in several parts of the investigated small area and restoration areas.
- painting techniques, based on the substrates and mixture of pigments, indicated by XRF.

X-ray attenuation is related to the atomic number and the physical thickness of the investigated area. The challenge is a three-dimensional interpretation of a two-dimensional representation, provided by XRF data. The goal is the investigation of aforementioned issues, by means of SOM and t-SNE- based clustering of only XRF data set. This approach is useful for a preliminary study of the mural painting and it can direct the future study, since it identifies specific regions of interest (ROIs) and it facilitates the further selection of experimental destructive or not techniques.

The information obtained for the authentic pigments by a simple observation of elemental distribution images (Figure 3-5) are the below:

- iron-based pigments are used for the skin.
- copper-based pigments are used for the collar.
- arsenic, contained in yellow orpiment (As_2S_3) or realgar (AsS_4), is used for the yellow background of the hieroglyphic text.
- calcium, contained in many white paints such as calcite ($CaCO_3$), is used for the white background.

The information obtained for the authentic pigments by a simple observation of elemental distribution images (Figure 3-5) is the use of calcium combined with sulfur, probably gypsum, for repairs on the white background. The elemental distribution image for titanium shows that this element is not contained in white paints in this mural painting. Titanium seems to be present in iron- containing areas.

The relevant bibliography, the XRF elememental distribution images and a previous study of the investigated area with RIS data set [10] suggest the below:

- a mixture of goethite ($\alpha - FeO(OH)$) and hematite (Fe_2O_3) is a common egyptian paint. Goethite is a yellow ochre. Once heated, a red pigment is produced. Natrojarosite ($KFe_3(OH)_6(SO_4)_2$) can also be added to produce different hues, but the not significant contribution of potassium suggests the use of Fe-containing pigments other than natrojarosite [33].
- a mixture of goethite and ilmenite ($FeTi0_3$) is also present in egyptian paintings. This can explain the presence of titanium in iron-containing areas.
- hematite is commonly present in Egyptian red inks [33].
- copper-based pigments can be found on different thicknesses and concentration in the area of the collar. A thin layer of Egyptian blue ($CaCuSi_4O_{10}$) is suggested below thicker bands of Egyptian blue. The RIS data also suggest the existence of bands of Egyptian green ($CaCuSi_4O_{10} + SiO_2$), altering with bands of Egyptian blue in the original composition. Despite the apparent thicker Egyptian blue bands, only some traces of Egyptian green. is proposed by the study of RIS data set [10].
- darkening of the Egyptian blue can be explained by three phenomena. A rough surface can entrap dirt. As expected, a finer layer or pigment tends to have less surface dirt. Other causes are the formation of black copper oxide (tenorite, CuO) and gypsum crusts, discolored by dirt. The degradation of Egyptian blue can lead to green areas [34].

The investigation of the small area of the mural painting aims to reveal additional information only with clustering of the XRF data set. The results are discussed and compared with the relevant bibliography and the results from the RIS data set.

3-2 Data augmentation methods

The data augmentation methods are mainly inspired by the article of neighbouring pixel data augmentation, described in Chapter 2 (*Method I*). In this section, three additional data augmentation methods are presented. Figure 3-6 is an illustration of the *Method II*. Each pixel is described by one feature vector: the spectrum of this pixel, followed by the mean spectrum of its eight neighbouring pixels. The augmented data set is normalized by Frobenius norm (the square root of the sum of the absolute squares of the matrix elements).

This method aims to combine two benefits: include the spatial information by augmenting the spectrum of data pixel with the spectrum of the neighbourhood and decrease the noise of the augmented data set by including the average of the spectra of the neighbouring pixels. The Appendix A-1-1 presents the code.

Method III includes the following steps:

- The neighbouring pixels of each pixel are indexed.
- The euclidean spectral distances between the central pixels are estimated. The euclidean spectral distances between the neighbouring pixels of one central pixel with the neighbouring pixels of all the other central pixels are also estimated.

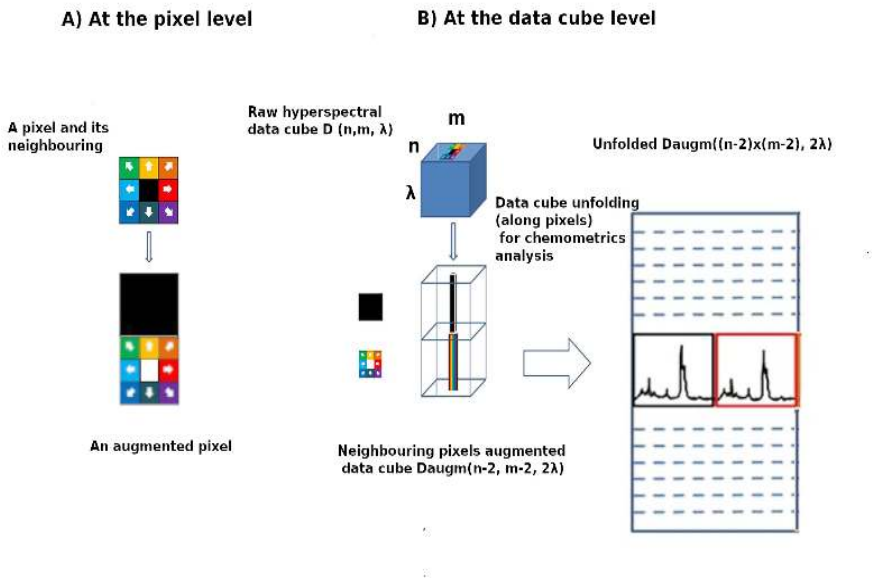


Figure 3-6: The spectrum of a pixel is represented by one cube. In the pixel data augmentation, each pixel is corresponded with two spectra, since the spectrum of the central pixel and the mean spectrum of the eight nearest neighbours of each pixel are considered.

- The sum of the eight minimum euclidean spectral distances between the neighbourhoods is calculated.
- Finally, the data points of this distance matrix are clustered with t-SNE with precomputed metric.

The Appendix A-1-2 includes the code.

3-3 Stationary wavelet transform algorithm

Wavelet transform algorithm, as part of a spectral and spatial data fusion approach, has been performed to an artificial data set, as mention in Chapter 2. In this thesis project, a two-dimensional stationary wavelet transform algorithm is applied to the XRF data of the titanium alloy. The steps of this method (labeled as *Method IV*) followed are the below:

- The original data cube is analysed with the simple Haar wavelet transform algorithm. Each elemental distribution image produces four images of wavelet coefficients for the first level decomposition.
- Spectra and wavelet coefficients have different properties. The wavelet coefficients are auto-scaled.
- Both spectral and coefficient images are normalised by Frobenius norm.

3-4 Optimization of t-SNE

In Section 2-2-1, Fast interpolation-based t-SNE (FIt-SNE) [14], an implementation of t-SNE with computation time reduced by at least ten times compared to BH t-SNE, is presented. As a result, FIt-SNE allows to process entire data sets of paintings without downsampling.

Another modification to t-SNE is *late exaggeration* [35]. The goal is to separate the clusters in t-SNE and allow their easier identification without user's intervention or a priori knowledge. This modification can be useful for the data sets of paintings. They often contain mixture of pigments and t-SNE with *late exaggeration* can perform a better separation of clusters. Conventional t-SNE uses the *early exaggeration* to deal with the problem of convergence rate getting slower for bigger data sets. The attraction term between similar data points during t-SNE is multiplied with a constant $a > 1$ during the first 250 iterations. This leads to tighter clusters, moving faster and avoiding getting trapped in local minima. Later exaggeration refers to enforcing exaggeration during the last hundreds iterations. As a result, the clusters contract and can be interpreted more easily.

3-5 Self Organising maps with different neighborhoods

There are at least two ways of clustering the data set with SOM. For simple data sets, such as the XRF data set of the titanium alloy, clustering with the use of a U-map can yield useful results. As mentioned in 2-2-1, the neuron map, describing the input data set, consists of neurons, each of which is described by the vector of weights. Each node of U-matrix is the average distance between neuron's vector and that of its four or eight neurons. In this thesis, it seems that the selection of proper neighborhood plays a crucial role in the identification of clusters in the data set of the titanium alloy.

The complex XRF data sets, such as the ones corresponding to segments of the oil painting, are clustered in a simpler way. The number of neurons, specified by the user, equals the number of clusters, since the pixels with the same neuron are clustered together. The number of clusters is determined by the user and as a result the selection is biased.

The aim of applying SOM clustering to a complex XRF data set is the identification of clusters that contain the same combination of elements but different relative strength. This might lead to separation of original from restoration pigments, which include the same elements.

3-6 Artificial neural network for paintings

A simple artificial neural network (code found on Appendix A-2) is used to combine the information contained in the photograph of the painting and the equivalent XRF data. The photographs of the paintings are of higher resolution than the equivalent XRF data sets. Instead of reducing the resolution of the photograph, we can upscale the elemental distribution images. This is achieved by upscaling each image with super resolution based on deep learning methods. The performed convolutional neural network for high-quality reconstruction of an image is the Laplacian Pyramid Super-Resolution Network (LapSRN) ([36]). This model

is slower than Efficient Sub-Pixel Convolutional Neural Network (ESPCN) and Fast Super-Resolution Convolutional Neural Network (FSRCNN), and the accuracy is worse than the one of Enhanced Deep Residual Networks (EDSR). However, LapSRN can perform 8x super-resolution, which is preferable because the resolution of the photograph is much higher than the resolution of the XRF data sets.

Chapter 4

Results and Discussion

4-1 Results

4-1-1 Methods with augmented data set

Method I and II

Test data set The data pixel augmentation methods are applied to the test data set. The test data consists of 60×60 pixels with a $n = 9$ dimensional feature vector for each pixel, as presented in Chapter 3. Figure 4-1 (a) presents the result for the test data set, clustered with t-SNE. As expected, three clusters are formed, belonging to the upper rectangle, the lower rectangle and the background. Figure 4-1 (b) presents the result for the test data set augmented with Method II, clustered with t-SNE. The augmented test data consists of 60×60 pixels with a $(1 + 1) \times n = 18$ dimensional feature vector for each pixel. Each augmented pixel is described by the feature vector of the central pixel and the mean value of the feature vectors of its 8 neighboring pixels. The clustered augmented data set presents seven clusters, describing the upper and lower rectangles, the background and the borders of the rectangles. The clusters outline the rectangles isotropically. Figure 4-1 (c) presents the result for the test data set augmented with Method I, clustered with t-SNE. The augmented test data consists of 60×60 pixels with a $(1 + 8) \times n = 18$ dimensional feature vector for each pixel. Each augmented pixel is described by the feature vector of the central pixel and the feature vectors of its 8 neighboring pixels. The clustered augmented data set presents seven clusters, describing the upper and lower rectangles, the background and the borders of the rectangles. The clusters outline separately all faces of the rectangle.

Data set of the titanium alloy Figure 4-2 (first row) presents the result for the data set with 36×40 pixels and 6 useful elemental signals of the titanium alloy, clustered with t-SNE. The three identified components are:

- α phase (blue color) with high intensity of the $Ti - K$ signal.

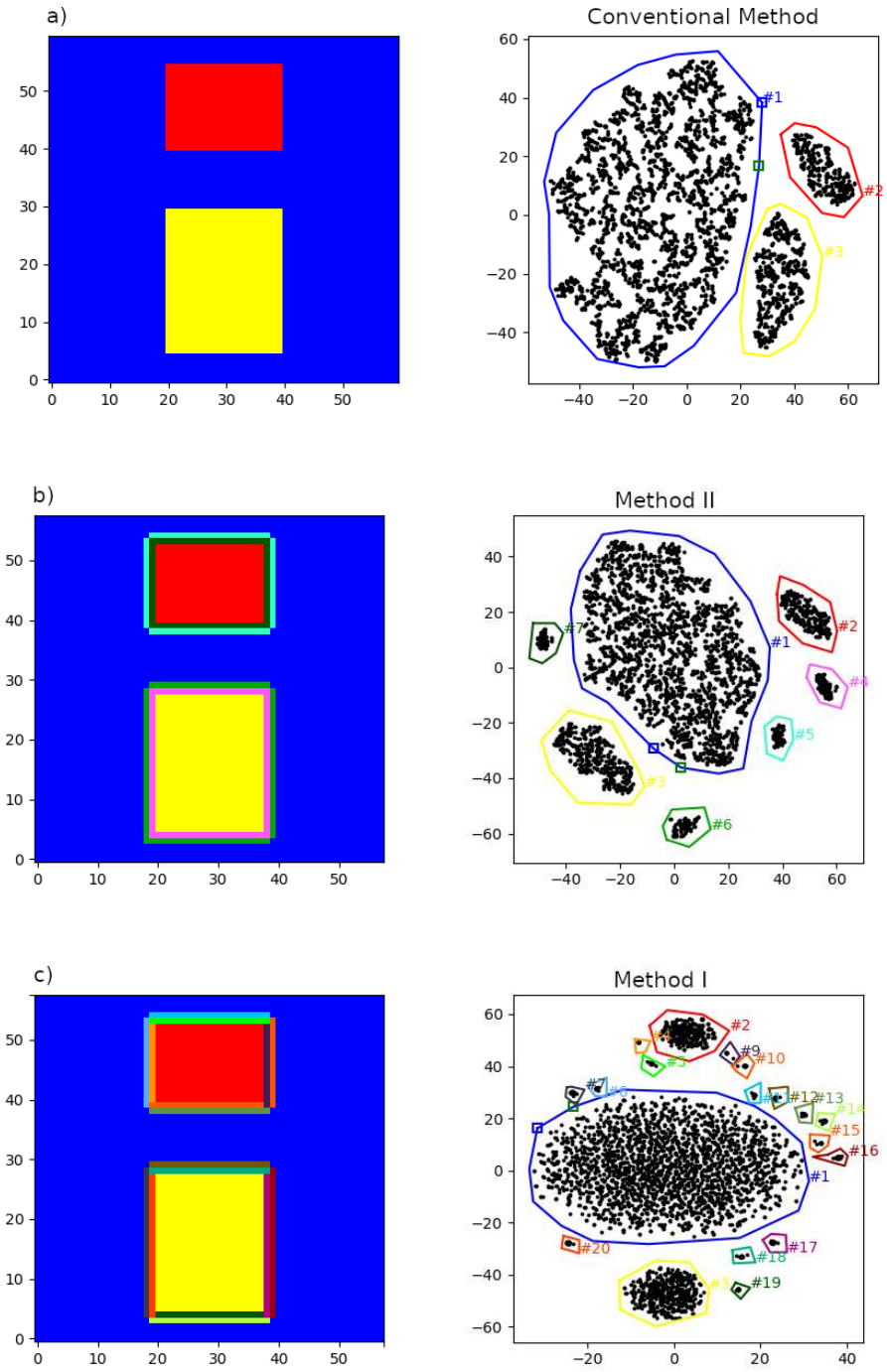


Figure 4-1: Clustered images of t-SNE representation for the a) test data set, b) test data set augmented with Method II and c) test data set augmented with Method I. The second column shows the corresponding scatter plots of the data points. Pixels and data points, belonging to the same cluster, have the same color.

- β phase (red color) with high intensity of the $V-K$, $Fe-K$, $Ni-K$ and $Cu-K_\alpha$ signals.
- α/β phase (yellow color) with intensity of signals between the one of phase α and β.

Figure 4-2 (second row) presents the result for data set augmented with the Method II of the titanium alloy, clustered with t-SNE. The data set augmented with this approach concept consists in augmenting the spectrum of a central pixel with the mean spectrum of its eight neighboring pixels. It consists of (36-2) $\times$ (40-2) pixels and $(1+1)\times 6 = 12$ elemental peaks. After normalizing each image before and after augmentation by the square root of its mean value, this method of neighboring pixel data augmentation reveal the below information:

- the superposition of α on β phase (yellow area) with intensity of signals between the one of phase α and β. The signal intensity of $V-K_\alpha$ is stronger for α/β than β/α component.
- superposition of β on α phase (orange area) with intensity of signals between the one of phase α and β. The signal intensity of $V-K_\beta$ is stronger for β/α than α/β component.

Figure 4-2 (third row) presents the result for the augmented data set with Method I of the titanium alloy, clustered with t-SNE. This augmented data set concept consists in augmenting the spectrum of a central pixel with the spectra of its eight neighboring pixels.The data set augmented with Method I consists of (36-2) $\times$ (40-2) pixels and $(1+8)\times 6 = 54$ elemental peaks. Besides performing the same normalization process, the data points of the data set augmented with Method I do not form obvious clusters. Therefore, the clusters, as formed by Method II, are projected on the data points of the data pixel augmentation method with nine spectra per pixel. It does not seem likely to separate α/β phase from β/α phase with this method.

The identification of 4 components is also possible in different areas (Figure 4-3 (a)) of the same titanium alloy sample (see Appendix A-3). This is possible with a priori knowledge of the NMF results. To achieve better separation of clusters, the clustering of aforementioned XRF data set is performed with FIt-SNE.

In Figure 4-3 (a), NMF 1 and NMF 2 represent primary α and β phase, respectively. The overlap of two phases is represented by NMF 3. All methods are capable of separating the α phase (blue color area) from the β phase (red color area) and the overlapping phases (rest of the areas) (Figure 4-3 (b), (c) and (d)). However, they are not efficient in separating α/β component from β/α component, as expected with Method II (Figure 4-3 (c)).

Method III: Interaction between neighborhoods

Another neighboring pixel data augmentation method is the one, considering the similarities between the neighborhoods of the central pixels (Method III). This method seems as efficient as Method II for the data set of the titanium alloy. This method is also able to separate all four components (Figure 4-4). The identification of four clusters is not straightforward and the a priori knowledge of the existence of four components, based on the Monte Carlo simulation, played an important role in selecting the clusters.

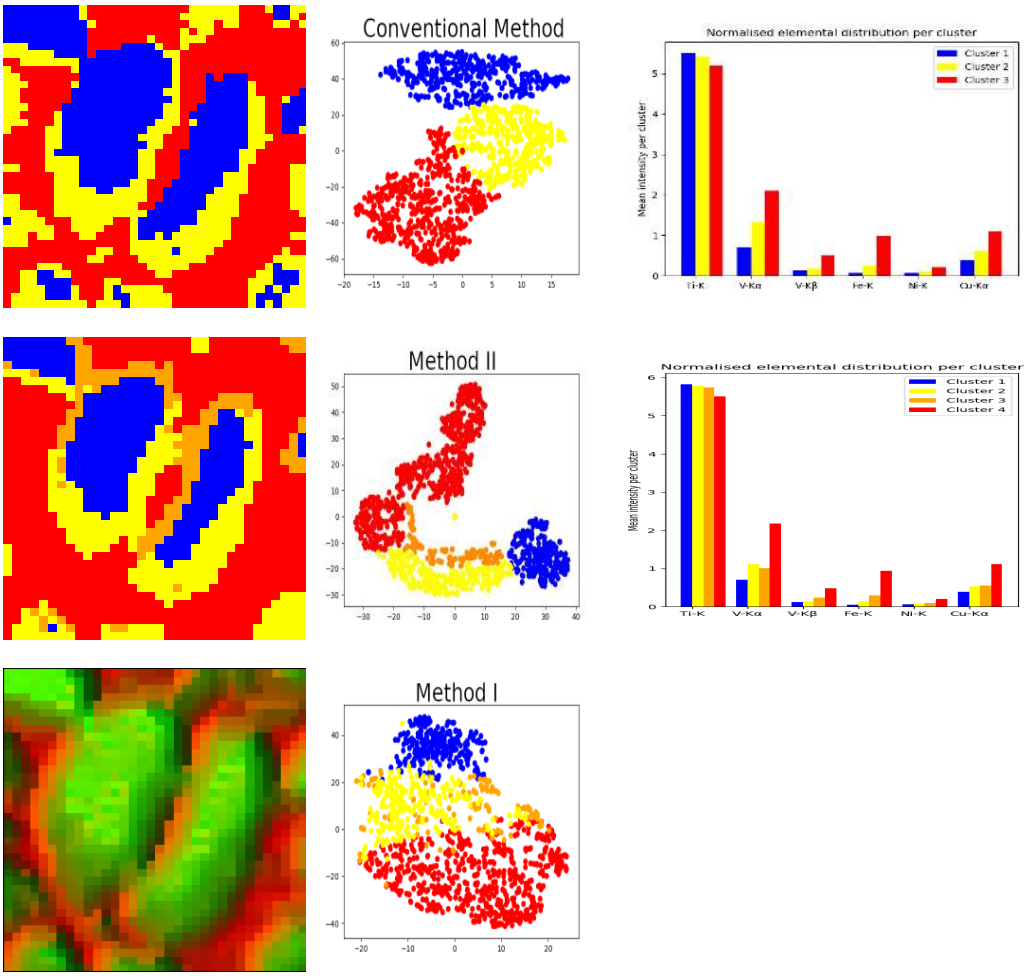


Figure 4-2: The first column shows clustered images of t-SNE representation for the data set of the titanium alloy (first row), data set augmented with Method II (second row) and data set augmented with Method I (third row). The second column shows the corresponding scatter plots of the data points and the third column the mean fitted spectra per cluster. Pixels and data points, belonging to the same cluster, have the same color.

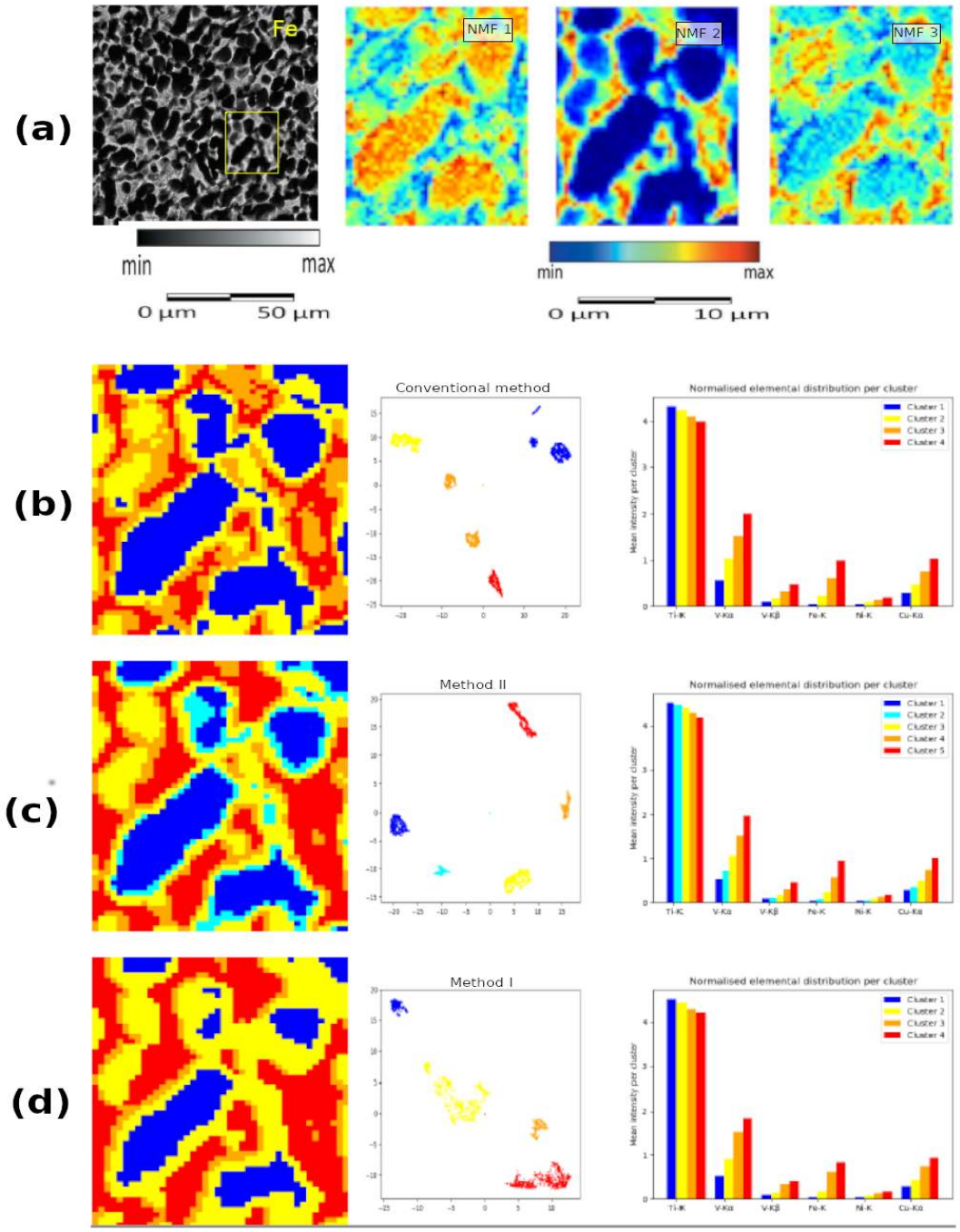


Figure 4-3: (a) Elemental distribution image for the $Fe - K$ line, whose yellow outline includes the region of interest, along with three typical coefficient images obtained by NMF from the above data set. Clustered images, scatter plots and fitted spectra of FIt-SNE representation for (b) the data set, (c) data set augmented with Method II and (d) data set augmented with Method I for part of the titanium alloy.

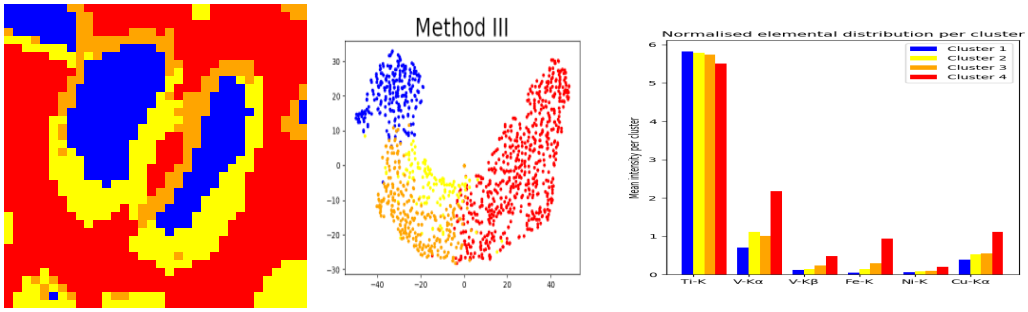


Figure 4-4: Clustered image of t-SNE represention for the titanium data set with Method III, scatter plot and mean fitted spectra per cluster.

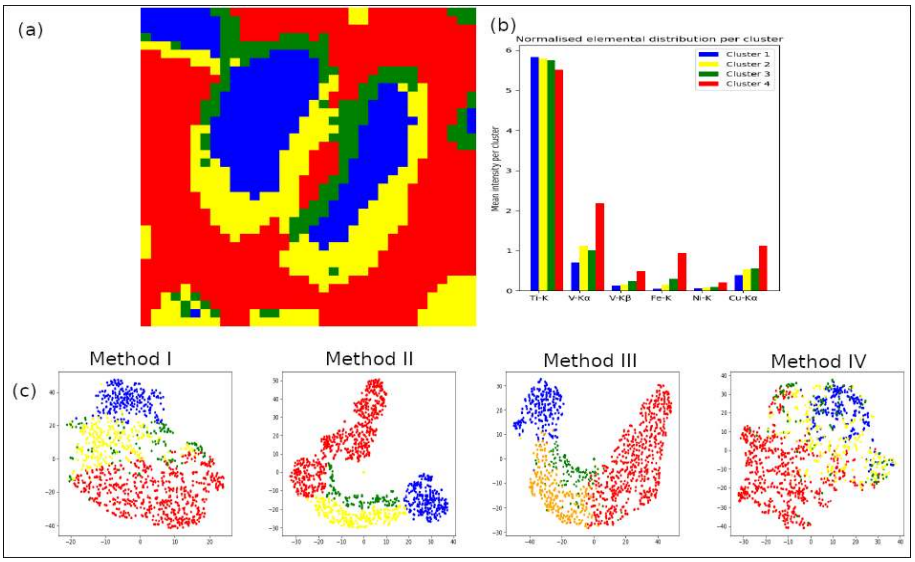


Figure 4-5: (a) Clustered image of t-SNE represention for the titanium XRF data set augmented with Method II, (b) mean fitted spectra per cluster and (c) the scattered plots for all presented methods. The data points of the scattered plots are colored based on the clustering of data set with Method II.

Method IV: Wavelet transform algorithm

Wavelet transform algorithm does not offer any additional information for the data set acquired from a sample of a bi-modal Ti-6Al-6V-2Sn alloy. This method seems to have similar performance as Method I, for the data set of the titanium alloy. It offers an approximation and enhances horizontal, vertical and diagonal details. When the number of elemental images is limited, simpler methods and noise reduction methods present higher performance. Figure 4-5 includes the results of clustering with t-SNE the data sets with all the pixel augmentation methods.

4-1-2 Self Organising Maps

Data set of titanium alloy

SOM clustering with U-map has the benefit of unbiased identification of the number of clusters and unbiased selection of pixels belonging to a certain cluster. On the other hand, U-map is not a sophisticated way of creating a clustered image. SOM clustering identifies three components with unbiased way compared to t-SNE clustering, since the user selects the neurons and not the clusters. According to clustering with U-map, at least three components are identified (Figure 4-6 first row). Qualitatively, the yellow area of the U-map corresponds to the α phase, the green area to the β phase and the red area to the α/β. However, the neurons, functioning as separators, form the clusters, but they do not offer any significant information in the clustered image. After this first observation, a more detailed examination is required.

Figure 4-6 (second row) presents an image of the titanium alloy's data set, automatically clustered with SOM of three neurons. This neural network identifies successfully the existence of at least three components. Figure 4-6 (third row) is the optimal visualisation with SOM of the titanium alloy's data set, clustered with SOM of four neurons. The interpretation of this last representation is the following:

- α phase (blue color) with high intensity of the $Ti - K$ signal.
- β phase (red color) with high intensity of the $V - K$, $Fe - K$, $Ni - K$ and $Cu - K_\alpha$ signals.
- overlapping phases (yellow color) with intensity of signals between the one of phase α and β.
- SOM clustering offers an automatic clustering, but it is unable to separate between α/β and β/α phase.

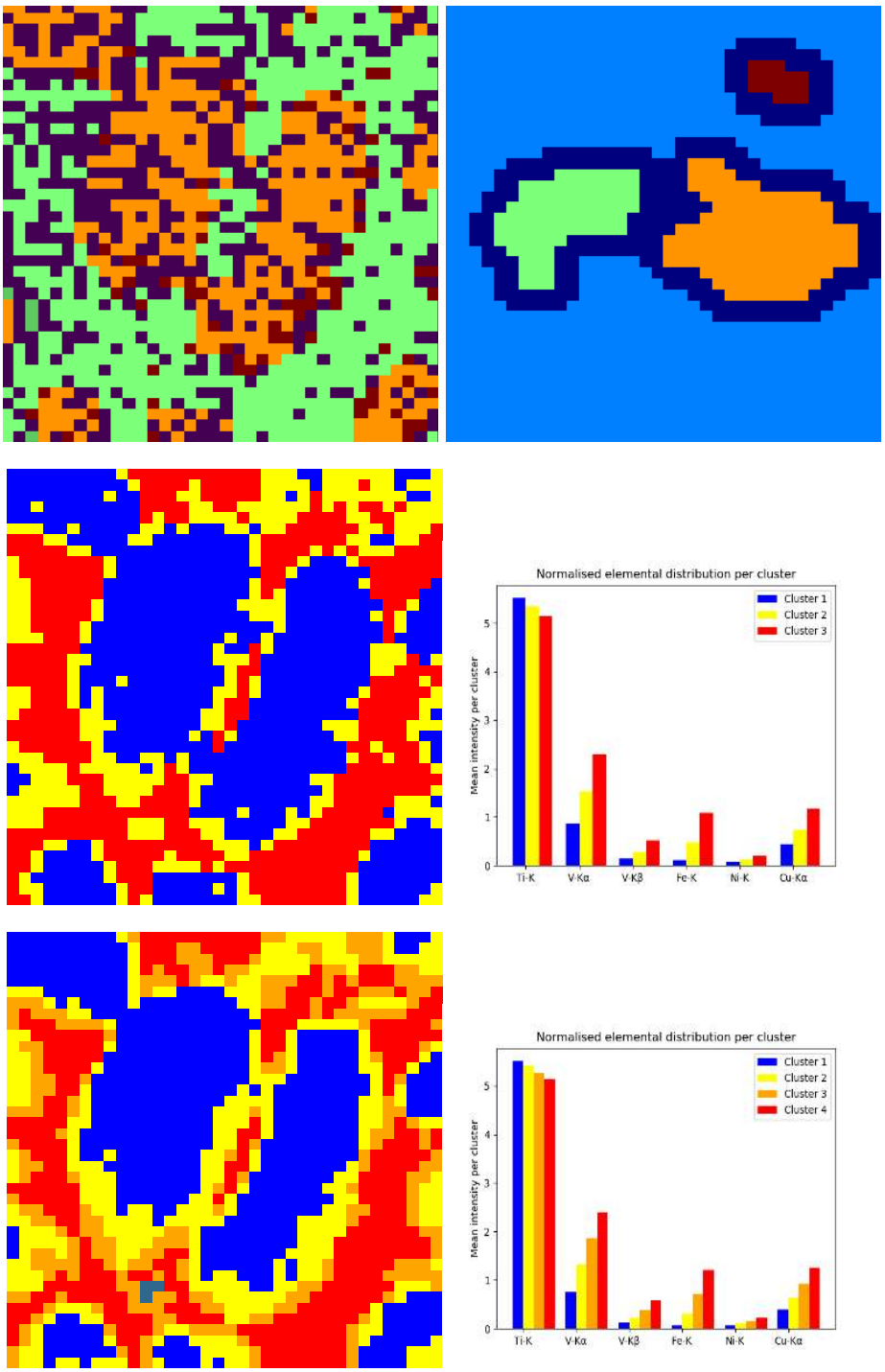


Figure 4-6: SOM with 40 $\times$ 40 neurons and clustering with U-map identifies three components with unbiased way (first row). Clustered image with SOM of 3 (second row) and 4 neurons (third row) along with their mean fitted spectra per cluster.

Data set of mural painting

Figures 4-7 and 4-8 present the results of SOM clustering for different parameters. Figure 4-7 (Top) presents the results of SOM clustering for 6 neurons and contains the below information:

- Cluster 1 mainly represents the white pigments of the mural painting.
- Cluster 2 and 3 mainly represent the copper-based pigments. The green-colored clustered area presents higher intensity lines, compared to the blue-colored one for all elements, except for sulfur.
- Cluster 4 with a strong arsenic line describes part of the yellow background, probably painted with orpiment.
- Cluster 5 and 6 represent the Fe-based pigments of the mural painting. The dark brown area has higher iron concentration than the beige area. A comparison with the photograph indicates that the lower iron concentration areas (Cluster 5) coincide with some apparent damaged areas. This suggests that these clusters do not reveal different Fe-based pigments, but difference in thickness.

The neccessity of more focused investigation of Fe-based pigments for the skin tone led to SOM clustering of 6 neurons for the segment of the face (Figure 4-7 Bottom). The later figure contains the below information:

- Cluster 1 mainly represents the white background of the mural painting.
- Cluster 2 presents strong arsenic and calcium lines. These signals can be explained by a mixture of calcite and orpiment.
- Cluster 3 presents strong calcium and iron lines. These signals can be explained by a mixture of Fe-containing pigments and calcite.
- Cluster 4 (beige), 5 (light brown) and 6 (dark brown) represent the main Fe-based pigments of the mural painting. They contain increasing concentration in iron and titanium and decreasing in calcium, contained in calcite. This suggests that these clusters do not reveal difference in thickness of Fe-containing pigments due to damages.

The comparison of spectra of Cluster 3 with Cluster 4 and 5 suggests that Cluster 3 functions as a dark red outline of the face.

SOM clustering with 6 neurons does not capture some essential information, as revealed by the elemental distribution images, such as the restoration pigments. Therefore, the number of neurons is increased with the aim of finding additional information about the mural painting. Figure 4-8 presents the results of SOM clustering for 12 neurons and contains the below information:

- Cluster 1 (dark red) presents strong calcium and iron lines. These signals can be explained by a mixture of Fe-containing pigments and calcite. The damaged areas are used as a guideline to align the clustered image with the photograph. The dark red outline seems to extend outside the outline of the face in the photograph. This suggests the existence of calcite (white background) on top of a Fe-containing pigments.

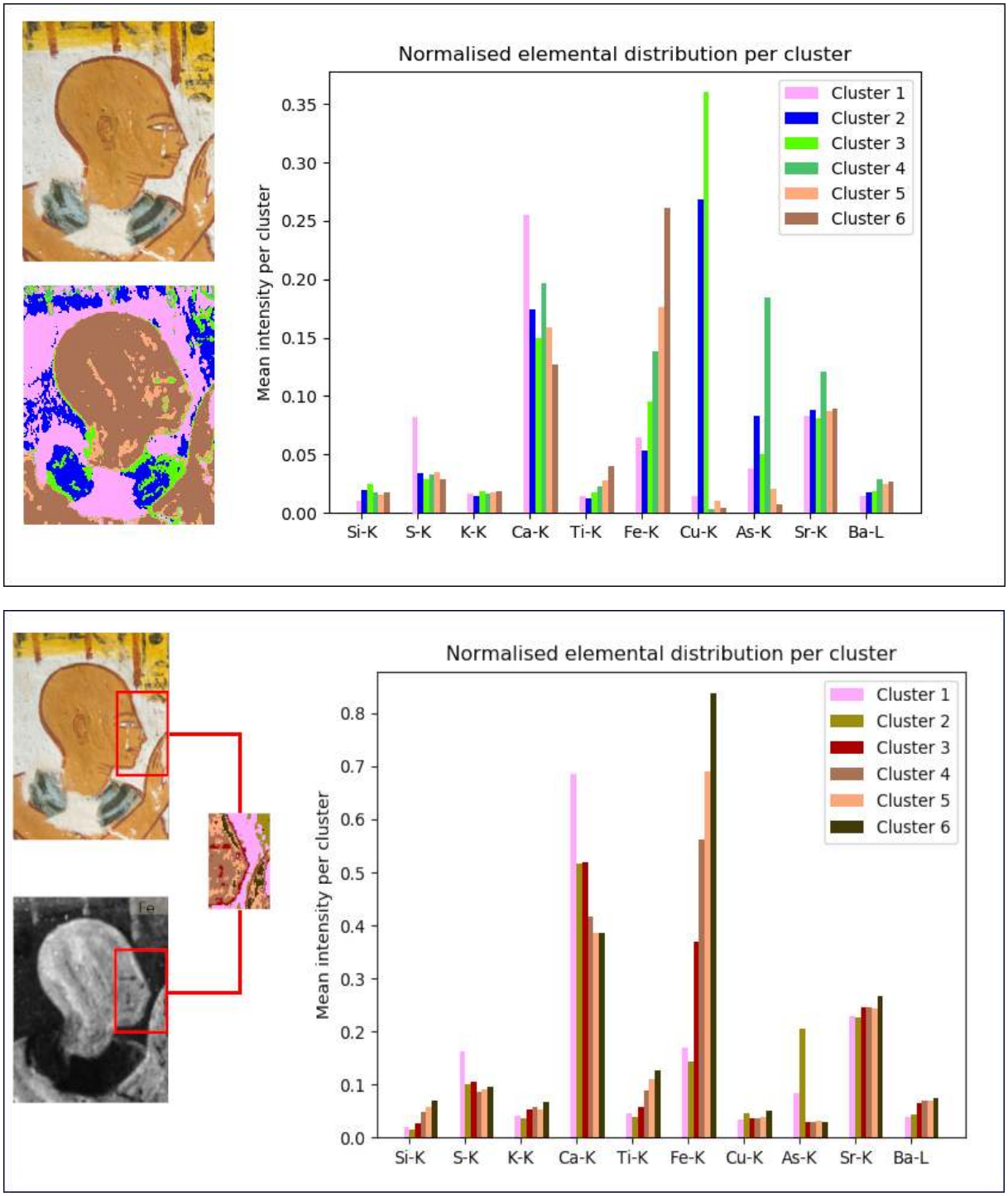


Figure 4-7: Clustered image with SOM of 6 neurons and its mean fitted spectra per cluster for the mural painting (Top) and a segment of the face (Bottom).

- Cluster 2 mainly describes copper-based pigments.
- Cluster 4 presents strong calcium signal, indicating the use of calcite for the white background. Cluster 6 even higher calcium signal. This difference cannot be attribute to different thickness of the calcium, since Cluster 6 (cyan area) is located on small areas. Therefore, the cyan area is probably painted with gypsum, used as a restoration paint.
- Cluster 6 and 7 represent the Fe-based pigments of the mural painting. The dark brown area has higher iron concentration than the beige area. A comparison with the photograph indicates that the lower iron concentration areas (Cluster 6) coincide with some apparent damaged areas. This suggests that these clusters do not reveal different Fe-based pigments, but difference in thickness.
- Cluster 3 and the rest of clusters describe the yellow background of the hieroglyphic text. This area mainly consists of orpiment, but also of Fe-based pigments. The interpretation of these clusters is not straightforward.

Comparison of SOM with FIt-SNE The results of SOM clustering for the mural painting are compared with those of t-SNE and FIt-SNE for the same area. The goal is double: confirm or contradict the preliminary study of mural painting with SOM clustering and present the benefits of clustering XRF data sets of paintings with FIt-SNE. The methodology is the selection of apparent clusters without a priori knowledge of elemental distribution images and RIS data set. Figure 4-9 (Top) presents the results for clustering with t-SNE the XRF data set for mural painting.

- Cluster 1 and 2 describes copper-based pigments. The two blue colored bands in the collar have stronger copper signal than the green colored area, but weaker sulfur contribution.
- Cluster 3 describes the Fe-containing areas, without capturing the difference in thickness of these pigments.
- Cluster 4 (dark red) presents strong calcium and iron lines. These signals can be explained by a mixture of Fe-containing pigments and calcite. In the yellow area of the hieroglyphic text, there is also one at least vertical lines described by this cluster.
- Cluster 5 describes the gypsum, used for repairing the white background.
- The area of Cluster 6 is separates from Cluster 7. The main difference in the mean spectra per cluster is the increased sulfur in Cluster 6. The paint in Cluster 6 functions as restoration pigment, whereas the paint in Cluster 6 is an original pigment.
- Cluster 8 features a strong arsenic signal. It describes the yellow background, mainly painted with orpiment.

The computational cost is decreased dramatically with FIt-SNE clustering. A better separation of clusters by modifying exaggeration offers a better visualisation and reveals additional

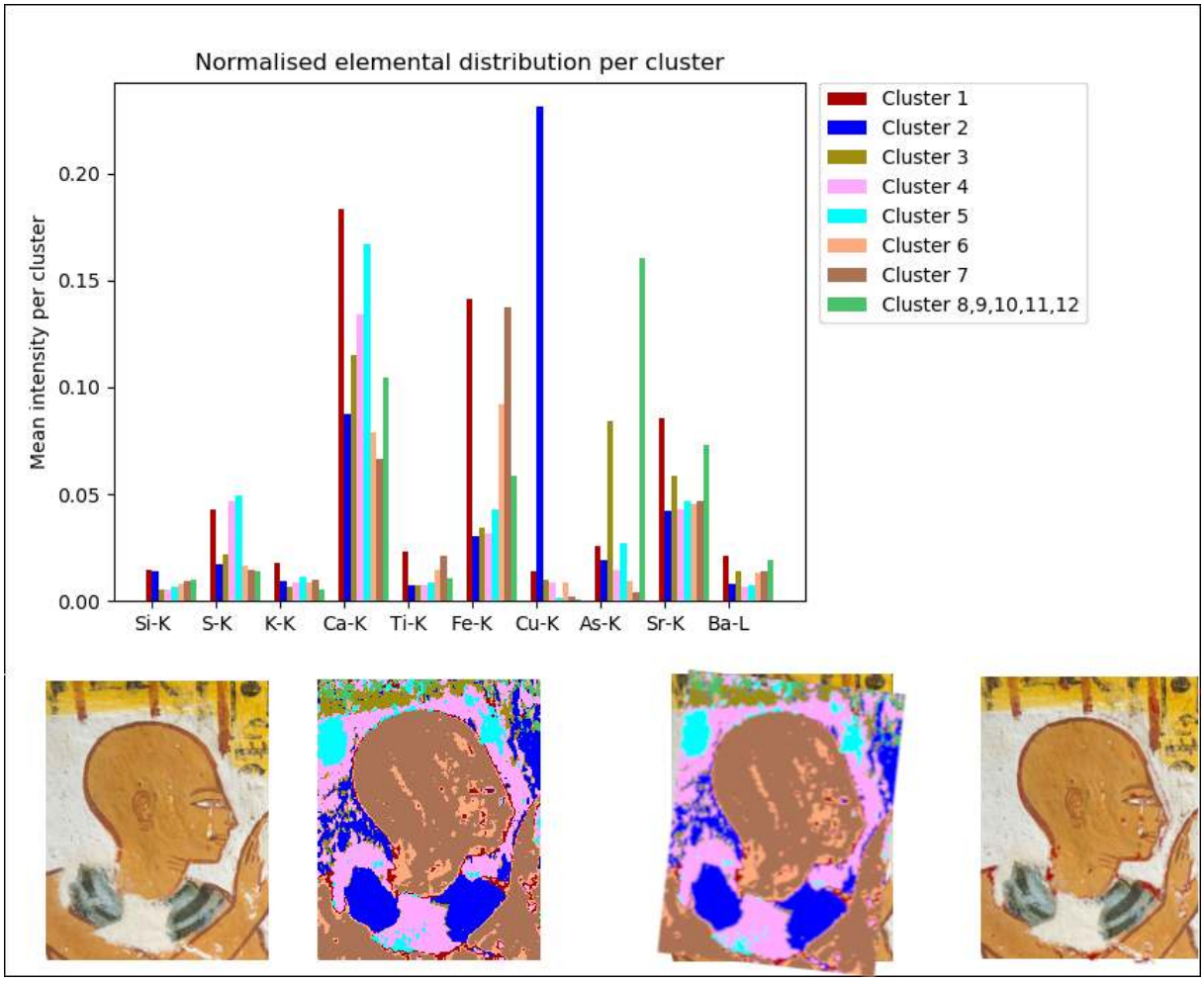


Figure 4-8: Clustered image with SOM of 12 neurons and its mean fitted spectra per cluster for the mural painting. The clustered image is aligned with the relevant photograph, based on the damaged areas.

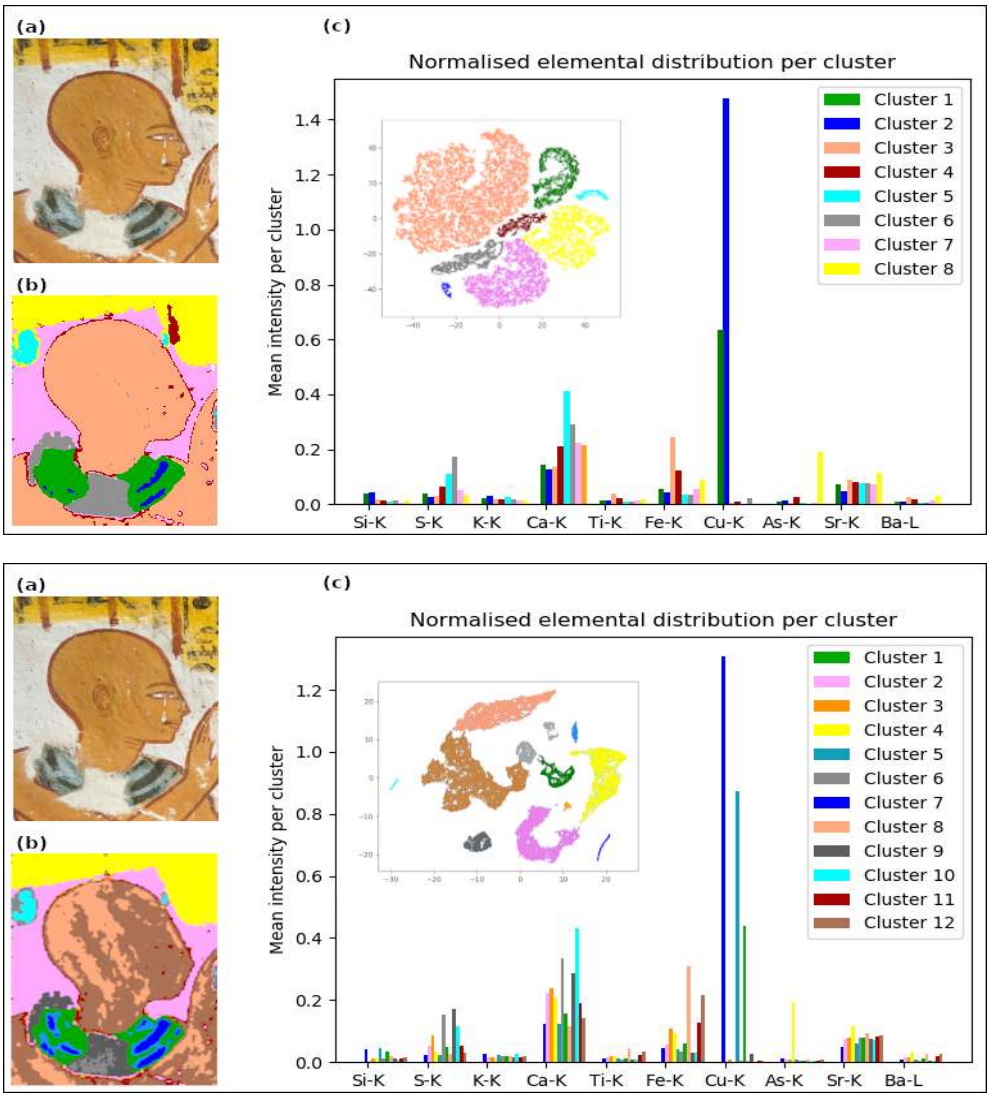


Figure 4-9: Clustered image with t-SNE (Top) and FIt-SNE (Bottom), scatter plot and mean fitted spectra per cluster for the mural painting.

information about the mural painting. Specifically, the difference in thickness of the iron-based pigments is captured in Cluster 8 and 12 of Figure 4-9 (Bottom (b) and (c)) with Cluster 8 assumed to be the higher thickness area. Moreover, the bands of copper-based pigments are more discernible and agree with the photograph Figure 4-9 (Bottom (a)). Once again more sulfur is observed in the green colored areas than the blue ones. Finally, the restoration pigments are visualized with more detail.

4-1-3 Artificial Neural Network for the easel painting

A simple artificial neural network can offer a visualization, concentrating most of the valuable information of the elemental distribution images in only one image. The spatial information of the photograph is combined with the spectral information, obtained by XRF technique. This is achieved by training a random number of photograph's pixels with a random number of pixels from the elemental distribution images. The result image is the prediction of this training for all the pixels of the photograph. However, the elemental distribution images are of lower resolution than the photograph. This inconsistency in resolution can be overcome by reducing the resolution of the photograph, as shown in Figure 4-10 (middle image of first row). The elements included in the data set, used as training set, are calcium, iron, copper, cobalt, tin and lead.

Another way to achieve higher resolution and capture more details in the result image is to upscale the elemental distribution images. The convolutional neural network for high-quality reconstruction of an image (LapSRN) has been evaluated as accurate for various photographs. However, it has not been applied for upscaling elemental distribution images. Nevertheless, the result image of ANN with upscaled elemental distribution images (Figure 4-10 last image of first row) has undeniably higher resolution than the result image of a conventional ANN. The bottom images of Figure 4-10 compare the two methods by presenting at pixel level two enlarged areas of interest. In this case, higher resolution increases the amount of information, obtained by the result image, without revealing any additional hidden information.

Figure 4-10: Photograph with 3704×4536 (w $\times$ h) of the oil painting, probably portrait of Hortense Manchini, followed by the result image with 290×389 (w $\times$ h) from the artificial neural network with training answer the photograph of the painting and training set the spectral data. The third image of the first row is the result image 2320×3109 (w $\times$ h) from the same artificial neural network, but with upscaled spectral data. The second row compares at pixel level the result image of the simple ANN and the one of the ANN with upscaled spectral data.

4-2 Discussion

Proper normalization of XRF data set augmented with Method II for the titanium alloy together with the relevant NMF results for this area facilitate the selection of four clusters:

- a cluster with pixels, whose mean spectrum has strong $Ti-K$ signal (α phase).
- a cluster with pixels, whose mean spectrum has strong $V-K$, $Fe-K$, $Ni-K$ and $Cu-K_\alpha$ signals (β phase).
- a cluster with pixels, whose mean spectrum has strong $V-K_\alpha$ signal (α/β component).
- a cluster with pixels, whose mean spectrum has strong $V-K_\beta$ signal (β/α component).

The results are aligned with the Monte Carlo simulation. The divergences of the results of the simulated data from those of the augmented data set is attributed to the overestimation of iron in α phase and other negligible quantitative differences. The reason that only one method of data pixel augmentation identifies all four components can be explained by three factors. Firstly, the superposition of β on α phase is less pronounced than of α on β phase. The $V-K_\alpha$ line (4.952 keV) is below the Ti-K edge (4.966 keV) and the $V-K_\beta$ line (5.427 keV) is above it. As a result, the $V-K_\beta$ line is much stronger attenuated. Secondly, the noisy appearance of the Ti-K and the overall noise of the limited in number elemental distribution images can be improved with mean filtering or the mean spectrum of the neighboring pixels. Thirdly, vanadium is the only element, which differentiates the superposition of β on α phase from the superposition of α on β phase. All the useful signals are related more intensely with β phase, except for the Ti-K signal. The lines of tin, which is a dominant element of the α phase, were weakly excited and strongly absorbed by the medium between sample and detector.

Separating the clusters with FIt-SNE without the a priori knowledge of NMF results reveals the presence of three bases: α, β and overlapping phases. It can't separate sufficiently the α/β from the β/α component.

SOM (40×40 neurons) clustering with U-map identifies without the selection of the user three components, but the clustered image offers only limited qualitative results. SOM clustering (3×3 neurons and 4×4 neurons), whose number of neurons is the number of clusters, offers a better visualisation and again reveal three components. These results are aligned with those of t-SNE for the same area.

Clustering the XRF data set of a 13th century B.C. Egyptian mural painting with Self Organising Map (SOM) confirms and adds to the results, obtained by RIS data. Damaged areas are identified with SOM and used properly to align the clustered with SOM image with the photograph of the mural. The outline, a mixture of calcite and iron containing pigments, of the face (Cluster 1 of Figure 4-8) extends beyond the outline of the face, as specified by the photograph. This result is aligned with a previous study [10] with RIS data set, where the below sequence is suggested. Initially, Fe-based pigments are used for the skin among others. After that, white background, probably calcite, is created. Finally, the dark red outline is painted. As a result part of the Fe- based pigments for the skin are overpainted with white paint. This leads to an inconsistency between the outline of the face as described by SOM clustering and the outline as observed in the photogrpah.

The modification of t-SNE with later exaggeration offers a more detailed clustered image. The damaged areas are visualized with more detail, the different thicknesses between pigments is captured and the Egyptian blue and probably Egyptian green areas are identified.

Finally, ANN for the investigation of complicated painting presents two main benefits: (a) a single result image, fusing the information of photograph and XRF data, and (b) a visualization, revealing the inconsistency between certain outlines of the photograph and the elemental distribution images. Therefore, ANN offer fast initial results, revealing issues for further exploration. Before performing the ANN, the photograph of the painting and the elemental distribution images with different resolution are processed accordingly. The idea of upscaling the elemental distribution images serves the purpose of exploiting the higher resolution of the photograph, when fusing spatial and spectral information with ANN. However, the convolutional neural network (CNN), used for upscaling the images, does not consider the special properties of XRF images. More advanced results require either the development of a CNN for upscaling images, which considers the XRF properties or the use of XRF detector with higher resolution.

Chapter 5

Conclusion and Future direction

5-1 Conclusion

In this study, an altered neighboring pixel data augmentation concept, SOM clustering and optimized artificial neural network framework are introduced as ways to combine spectral and spatial information. The results of these methods are presented and discussed for X-ray fluorescence data set of a bi-modal $\alpha + \beta$ Ti-6Al-6V-2Sn alloy, of an easel painting of 17^{th} century A.C., identified as portrait of Hortense Manchini and of 13^{th} century B.C. Egyptian mural painting.

The augmented data set concept consists in augmenting the spectrum of a central pixel with the spectra of its eight neighboring pixels. This method is generalized for all pixels. The altered augmented data set concept consists in augmenting the spectrum of a central pixel with the mean spectrum of its eight neighboring pixels. This new method, performed to part of the data set of the titanium alloy, reveals the existence of four components: α phase, β phase, α on top of β phase and β on top of α phase. The real data also reveal that the $V-K_\alpha$ signal is stronger in α/β, compared to β/α phase, and the opposite holds for the $V-K_\beta$ signal. In previous studies, β on top of α phase was not separately identified. Therefore, I conclude that proper normalization of the data set and clustering with t-SNE lead to the retrieval of new information that corresponds to the less pronounced superposition of β on top of α phase.

SOM clustering is performed to a simple data set of the titanium alloy and to the mural painting. I concluded that SOM clustering with U-map is useful for identifying without the requirement of cluster selection by the user α, β and α on top of β. However, U-map is not useful for the complicated XRF data set of a painting.

A simpler method of clustering the data sets with SOM is the below. The number of clusters is determined by the number of neurons, selected by the user, since the pixels belonging to the same neuron are considered part of the same cluster. Clustering the XRF data set of a 13th century B.C. Egyptian mural painting with Self Organising Map (SOM) and Fast interpolation-based t-SNE (FIt-SNE) indicate areas with different thicknesses of copper- and

iron- containing pigments and also reveal information about the painting sequence and the composition of a mixture of pigments. Moreover, FIt-SNE dramatically accelerates t-SNE and a modification in t-SNE, called *later exaggeration*, can provide well-separated clusters. Therefore, I conclude that more tedious preliminary work and advanced clustering techniques applied to the XRF data set of this painting lead to results, aligned with the results from the RIS data set. Additionally, altering bands of copper-based pigments with varying sulfur concentration and thicknesses are visualized with greater detail. A better visualization and spectral information is also achieved for the restoration areas.

Finally, an optimized artificial neural network for training the photograph of a painting with the elemental distribution images was introduced. The elemental distribution images are upscaled with Laplacian Pyramid Super-Resolution Network, which performs 8 $\times$ super-resolution. In that way, we can exploit the high resolution of the photograph of the painting. As a result, we can identify with more accuracy the exact location of the damages and retouches on the portrait of Hortense Manchini, since the result image is of high resolution.

5-2 Future direction

A deeper interpretation of a spectroscopic imaging data set with high noise level may be assisted by data pixel augmentation methods. The exact effect of these methods on the noise level of the data sets can be examined and quantified. Additionally, the data pixel augmentation methods can be applied to data sets from the field of cultural heritage with several thousands of spectra. Therefore, the efficiency of the methods in noisy data sets with several spectra and not only a few fitted data can be investigated.

After showing that SOM clustering reveals information about the painting sequence and substrate by identifying clusters with the same elemental mixture but different signal intensities, this method can be applied to more complex data sets of paintings. Moreover, combining SOM with other ANN to achieve more detailed clustering is a promising idea. We can perform clustering of a data set with SOM as a typical layer in a neural network. Furthermore, we can add a convolutional neural network in front of the SOM layer, aiming at improving the clustering and reducing the dimensionality of the data.

Finally, the algorithms for combining spectral and spatial information with ANN during the evaluation of the hyperspectral data can be optimized. Specifically, we can create a convolutional neural network for creating super-high resolution images, modified to include the properties of hyperspectral data.

Appendix A

Developed Codes

A-1 Codes for data pixel augmentation

A-1-1 Data pixel augmentation: Method I and II

```
neigh_size=3
if np.mod(neigh_size,2) == 0:
   print('neigh_size needs to be uneven')
neigh_radius=np.int32(np.ceil((neigh_size-1)/2.))
dataw=data_unfold

    #indexes for finding neighbouring points
z=(data.shape[1])*neigh_radius+1
fc=[]
for i in range(neigh_size):
    for j in range(neigh_size):
        if not (i == neigh_radius and j == neigh_radius):
           fc.append(np.ravel_multi_index((i,0),data.shape[:2])+j)


data_neigh=np.zeros(((data.shape[0]-neigh_radius*2)*(data.shape[1]-
    neigh_radius*2),8*data.shape[2]))
data_central=np.zeros(((data.shape[0]-neigh_radius*2)*(data.shape[1]-
    neigh_radius*2),data.shape[2]))
data_augm=np.zeros(((data.shape[0]-neigh_radius*2)*(data.shape[1]-
    neigh_radius*2),8*data.shape[2]))
data_neigh_av=np.zeros(((data.shape[0]-neigh_radius*2)*(data.shape[1]-
    neigh_radius*2),data.shape[2]))
data_augm_av=np.zeros(((data.shape[0]-neigh_radius*2)*(data.shape[1]-
    neigh_radius*2),2*data.shape[2]))
#for i in range (data_augm.shape[0]):
k=0

```

```
for i in range(data_neigh.shape[0]):
   x=i+(neigh_size-1)*np.int32(i/(data.shape[1]-(neigh_size-1)))
   print("x",x)
   for l in range (8):
     for j in range (data.shape[2]):
        data_central[i,j]=dataw[z+x,j]
        data_neigh[i,j+l*data.shape[2]]=dataw[fc[l]+x,j]
        data_neigh_av[i,j]=dataw[fc[:]+x,j].mean()

#augmented data set for Method I
data_augm = np.hstack([data_central,data_neigh])
#augmented data set for Method II
data_augm_av = np.hstack([data_central,data_neigh_av])
```

A-1-2 Data pixel augmentation: Method III

```
def calc_disMA_flex_new(data,neigh_size=3):
    if np.mod(neigh_size,2) == 0:
        print('neigh_size needs to be uneven')
        return
    neigh_radius=np.int32(np.ceil((neigh_size-1)/2.))
    dataw=data.reshape(np.prod(data.shape[:2]),data.shape[2])
    dis=np.zeros((dataw.shape[0],dataw.shape[0]))
    for i in range(dataw.shape[0]):
        dis[i,:]=((dataw-dataw[i,:,None].T)**2).sum(1)**.5#Simple
            euclidean distance
    if neigh_size ==1:
        return dis, dis, dis,dis

    disa=np.zeros(((data.shape[0]-neigh_radius*2)*(data.shape[1]-
        neigh_radius*2),(data.shape[0]-neigh_radius*2)*(data.shape[1]-
        neigh_radius*2)))
    disb=disa+0.
    disc=disa+0.

    #lateral distance
    x_r,y_r=np.meshgrid(range(data.shape[1]-neigh_radius*2),range(data.
        shape[0]-neigh_radius*2))
    x_r=x_r.reshape(np.prod(x_r.shape))
    y_r=y_r.reshape(np.prod(y_r.shape))
    for x in range(disa.shape[0]):
        disc[:,x]=(x_r-x_r[x])**2+(y_r-y_r[x])**2
    disc=disc**.5


    #indexes for finding neighbouring points
    z=(data.shape[1])*neigh_radius+1
    fc=[]
    for i in range(neigh_size):
        for j in range(neigh_size):
            if not (i == neigh_radius and j == neigh_radius):
                fc.append(np.ravel_multi_index((i,0),data.shape[:2])+j)

```

```
34
35
36      st0=time()
37      st1=time()
38      for x in range(disa.shape[0]):
39          ft=time()
40          print(('{0} of {1} lines in {2:5.2f},{3:5.3f}'.format(x,disa.
                shape[0],ft-st0,ft-st1)))
41          st1=time()
42          for y in range(disa.shape[0]):
43              if x < y:
44                  xo=x+(neigh_size-1)*np.int32(x/(data.shape[1]-(neigh_size
                        -1)))#For every line there are additional steps to
                        keep it synced
45                  yo=y+(neigh_size-1)*np.int32(y/(data.shape[1]-(neigh_size
                        -1)))
46                  disa[x,y]=dis[z+xo,z+yo]
47                  fca=dis[np.meshgrid(fc+xo,fc+yo)]
48                  while len(fca) > 1:
49                      disb[x,y]+=fca.min()
50                      wo=np.s_[np.unravel_index(fca.argmin(),fca.shape)]
51                      fca=np.delete(fca,wo[0],0)
52                      fca=np.delete(fca,wo[1],1)
53                  disb[x,y]+=fca.min()
54      #These are: dis2: combined central and neighbourhood
55      #   disa: Only the center
56      #   disb: neighbour and central (1/neigh_size**2)
57      #   disc: geometric distance
58      dim=dism
59      return disa+disa.T,disb+disb.T, disc, dis, dim
```

A-2 Code for artificial neural network

Each elemental distribution X-ray image can be upscaled with LapSRN as follows:

```
1  importcv2
2  from cv2 import dnn_superres
3  sr = dnn_superres.DnnSuperResImpl_create()
4  image = cv2.imread('...low_resolution_image.png')
5  #The LapSRN model can be downloaded here (https://github.com/fannymonori/
       TF-LapSRN/tree/master/export).
6  path = ".../LapSRN_x8.pb"
7  sr.readModel(path)
8  sr.setModel("lapsrn", 8)
9  # Upscale the image
10 result = sr.upsample(image)
11 #process repeated for all elemental distribution images
```

All these X-ray images are compiled to create a set of elemental distribution images, representing the spectral information. This new data set is used as training data set and the RGB image (photograph) of the painting as the training answer.

```
from tensorflow.keras.models import Sequential
import random
rand_ind=random.choices(range(np.prod(data.shape[:2])),k=50000)

#unfolded XRF data set as the trainng set
train_set=unfold(data)[rand_ind,:]
#unfolded picture (photography) of the painting as the training answer
train_answer=unfold(pic)[rand_ind,:]

model =tf.keras.Sequential()
model.add(Dense(30,input_dim=data.shape[2],activation='sigmoid'))
model.add(Dense(train_answer.shape[1],activation='sigmoid'))
sgd = SGD(lr=0.05, decay=1e-6, momentum=0.9, nesterov=True)



model.compile(loss=loss,optimizer=sgd,metrics=['accuracy'])
model.fit(train_set,train_answer,epochs=200)

vali=unfold(data)
ppre=model.predict(vali)


res=normin(refold(ppre,pic2.shape))
reso=res+0.
plt.imshow(reso[:,:,:])
```

A-3 t-SNE representation for titanium alloy

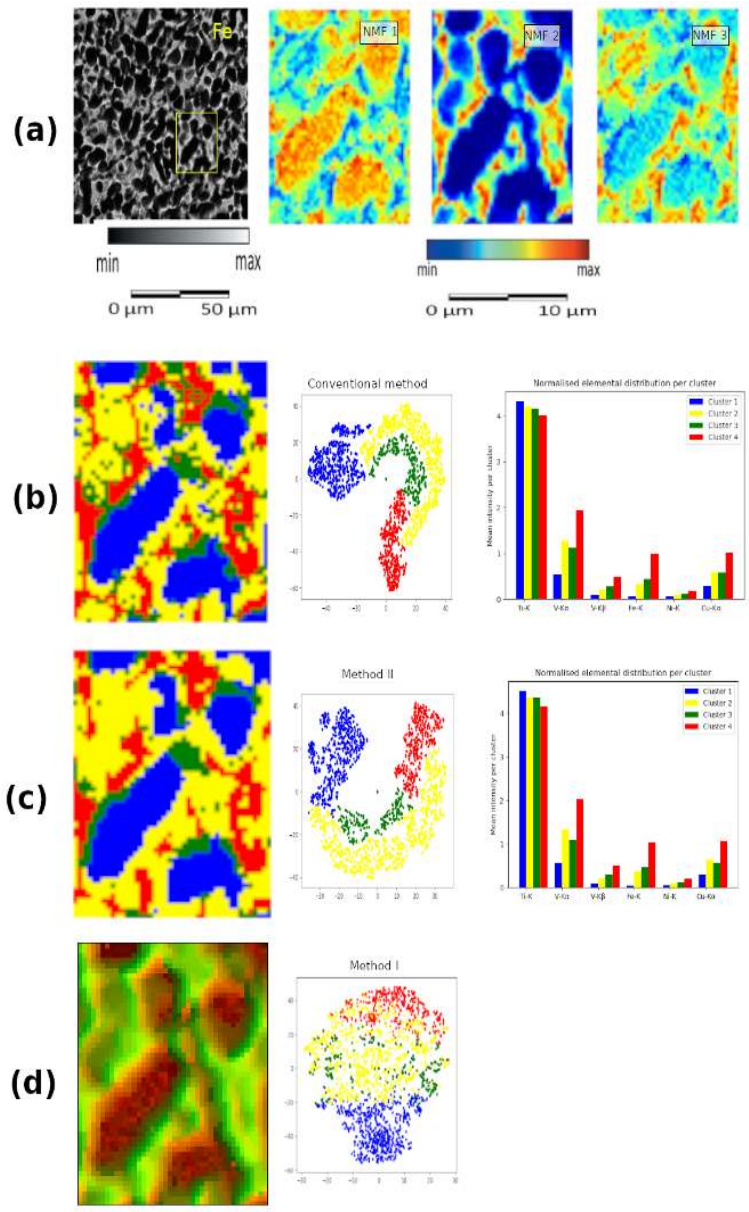


Figure A-1: (a) Elemental distribution image for the $Fe-K$ line, whose yellow outline includes the region of interest, along with three typical coefficient images obtained by NMF from the above data set. Clustered images, scatter plot and mean fitted spectra per cluster of t-SNE representation for (b) the data set, (c) the dataset augmented with Method II and (d) data set augmented with Method I for part of the titanium alloy.

Bibliography

[1] E. Raff, "A step toward quantifying independently reproducible machine learning research," 2019.

[2] S. Theodoridis and K. Koutroumbas, "Chapter 1 - introduction," in *Pattern Recognition (Fourth Edition)* (S. Theodoridis and K. Koutroumbas, eds.), pp. 1 – 12, Boston: Academic Press, fourth edition ed., 2009.

[3] F. Chollet, *Deep Learning with Python*. USA: Manning Publications Co., 1st ed., 2017.

[4] M. Han, R. Cong, X. Li, H. Fu, and J. Lei, "Joint spatial-spectral hyperspectral image classification based on convolutional neural network," *Pattern Recognition Letters*, vol. 130, pp. 38 – 45, 2020. Image/Video Understanding and Analysis (IUVA).

[5] N. Borodinov, S. Neumayer, S. V. Kalinin, O. S. Ovchinnikova, R. K. Vasudevan, and S. Jesse, "Deep neural networks for understanding noisy data applied to physical property extraction in scanning probe microscopy," *npj Computational Materials*, vol. 5, no. 1, p. 25, 2019.

[6] S. Berisha, M. Lotfollahi, J. Jahanipour, I. Gurcan, M. Walsh, R. Bhargava, H. Van Nguyen, and D. Mayerich, "Deep learning for ftir histology: leveraging spatial and spectral features with convolutional neural networks," *Analyst*, vol. 144, pp. 1642–1653, 2019.

[7] M. Chatzidakis and G. A. Botton, "Towards calibration-invariant spectroscopy using deep learning," *Scientific Reports*, vol. 9, no. 1, p. 2126, 2019.

[8] F. Oviedo, Z. Ren, S. Sun, C. Settens, Z. Liu, N. T. P. Hartono, S. Ramasamy, B. L. DeCost, S. I. P. Tian, G. Romano, A. Gilad Kusne, and T. Buonassisi, "Fast and interpretable classification of small x-ray diffraction datasets using data augmentation and deep neural networks," *npj Computational Materials*, vol. 5, no. 1, p. 60, 2019.

[9] Y. Zeng, J. C. A. van der Lubbe, and M. Loog, "Multi-scale convolutional neural network for pixel-wise reconstruction of van gogh's drawings," *Machine Vision and Applications*, vol. 30, no. 7, pp. 1229–1241, 2019.

[10] M. Alfeld, S. Pedetti, P. Martinez, and P. Walter, "Joint data treatment for vis nir reflectance imaging spectroscopy and xrf imaging acquired in the theban necropolis in egypt by data fusion and t-sne," *Comptes Rendus Physique*, vol. 19, no. 7, pp. 625 – 635, 2018. Physics and arts / Physique et arts.

[11] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 2008.

[12] L. van der Maaten, "Accelerating t-sne using tree-based algorithms," *Journal of Machine Learning Research*, vol. 15, pp. 3221–3245, 2014.

[13] V. van Unen, T. Höllt, N. Pezzotti, N. Li, M. Reinders, E. Eisemann, F. Koning, A. Vilanova, and B. Lelieveldt, "Visual analysis of mass cytometry data by hierarchical stochastic neighbour embedding reveals rare cell types," *Nature Communications*, vol. 8, pp. 1–10, 2017.

[14] G. C. Linderman, M. Rachh, J. G. Hoskins, S. Steinerberger, and Y. Kluger, "Fast interpolation-based t-sne for improved visualization of single-cell rna-seq data," *Nature Methods*, vol. 16, no. 3, pp. 243–245, 2019.

[15] T. Kohonen, "The self-organizing map," *Neurocomputing*, vol. 21, no. 1, pp. 1 – 6, 1990.

[16] S. Pagnotta, E. Grifoni, S. Legnaioli, M. Lezzerini, G. Lorenzetti, and V. Palleschi, "Comparison of brass alloys composition by laser-induced breakdown spectroscopy and self-organizing maps," *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 103-104, pp. 70 – 75, 2015.

[17] C. Arias, S. Bani, F. Catalli, G. Lorenzetti, E. Grifoni, S. Legnaioli, S. Pagnotta, and V. Palleschi, "X-ray fluorescence analysis and self-organizing maps classification of the etruscan gold coin collection at the monetiere of florence," *Applied Spectroscopy*, vol. 71, no. 5, pp. 817–822, 2017. PMID: 27154737.

[18] S. Pagnotta, M. Lezzerini, L. Ripoll-Seguer, M. Hidalgo, E. Grifoni, S. Legnaioli, G. Lorenzetti, F. Poggialini, and V. Palleschi, "Micro-laser-induced breakdown spectroscopy (micro-libs) study on ancient roman mortars," *Applied Spectroscopy*, vol. 71, no. 4, pp. 721–727, 2017. PMID: 28374609.

[19] J. Klus, P. Porizka, D. Prochazka, P. Mikysek, J. Novotny, K. Novotny, M. Slobodnik, and J. Kaiser, "Application of self-organizing maps to the study of u-zr-ti-nb distribution in sandstone-hosted uranium ores," *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 131, pp. 66 – 73, 2017.

[20] S. Kogou, L. Lee, G. Shahtahmassebi, and H. Liang, "A new approach to the interpretation of xrf spectral imaging data using neural networks," *X-Ray Spectrometry*, vol. n/a, no. n/a.

[21] R. KlockenkÃďmper, A. von Bohlen, and L. Moens, "Analysis of pigments and inks on oil paintings and historical manuscripts using total reflection x-ray fluorescence spectrometry," *X-Ray Spectrometry*, vol. 29, no. 1, pp. 119–129, 2000.

[22] E. Salerno, A. Tonazzini, and L. Bedini, "Digital image analysis to enhance underwritten text in the archimedes palimpsest," *International Journal of Document Analysis and Recognition (IJDAR)*, vol. 9, no. 2, pp. 79–87, 2007.

[23] S. Legnaioli, E. Grifoni, G. Lorenzetti, L. Marras, L. Pardini, V. Palleschi, E. Salerno, and A. Tonazzini, "Enhancement of hidden patterns in paintings using statistical analysis," *Journal of Cultural Heritage*, vol. 14, no. 3, Supplement, pp. S66 – S70, 2013. Science and Technology for the Safeguard of Cultural Heritage in the Mediterranean Basin.

[24] A. Tonazzini, E. Salerno, Z. A. Abdel-Salam, M. A. Harith, L. Marras, A. Botto, B. Campanella, S. Legnaioli, S. Pagnotta, F. Poggialini, and V. Palleschi, "Analytical and mathematical methods for revealing hidden details in ancient manuscripts and paintings: A review," *Journal of Advanced Research*, vol. 17, pp. 31 – 42, 2019. Special Issue on Celebrating JAR-1st IF.

[25] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[26] Chunghui Kuo, "Blind primary colorant spectral separation combining ica and pocs nonnegative matrix factorization," in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. iii–489, May 2004.

[27] M. Alfeld, M. Wahabzada, C. Bauckhage, K. Kersting, G. Wellenreuther, P. Barriobero-Vila, G. Requena, U. Boesenberg, and G. Falkenberg, "Non-negative matrix factorization for the near real-time interpretation of absorption effects in elemental distribution images acquired by X-ray fluorescence imaging," *Journal of Synchrotron Radiation*, vol. 23, pp. 579–589, Mar 2016.

[28] S. Mosca, T. Frizzi, M. Pontone, R. Alberti, L. Bombelli, V. Capogrosso, A. Nevin, G. Valentini, and D. Comelli, "Identification of pigments in different layers of illuminated manuscripts by x-ray fluorescence mapping and raman spectroscopy," *Microchemical Journal*, vol. 124, pp. 775 – 784, 2016.

[29] H. C. dos Santos, C. Caliri, L. Pappalardo, R. Catalano, A. Orlando, F. Rizzo, and F. P. Romano, "Real-time ma-xrf imaging spectroscopy of the virgin with the child painted by antonello de saliba in 1497," *Microchemical Journal*, vol. 140, pp. 96 – 104, 2018.

[30] F. Jamme and L. Duponchel, "Neighbouring pixel data augmentation: a simple way to fuse spectral and spatial information for hyperspectral imaging data analysis," *Journal of Chemometrics*, vol. 31, no. 5, p. e2882, 2017. e2882 CEM-16-0141.R1.

[31] A. Nardecchia, R. Vitale, and L. Duponchel, "Fusing spectral and spatial information with 2-d stationary wavelet transform (swt 2-d) for a deeper exploration of spectroscopic images," *Talanta*, vol. 224, p. 121835, 2021.

[32] L. de Viguerie, S. Rochut, M. Alfeld, P. Walter, S. Astier, V. Gontero, and F. Boulc'h, "Xrf and reflectance hyperspectral imaging on a 15th century illuminated manuscript: combining imaging and quantitative analysis to understand the artist's technique," *Heritage Science*, vol. 6, p. 11, Apr 2018.

[33] T. Christiansen, M. Cotte, W. de Nolf, E. Mouro, J. Reyes-Herrera, S. de Meyer, F. Vanmeert, N. Salvadó, V. Gonzalez, P. E. Lindelof, K. Mortensen, K. Ryholt, K. Janssens, and S. Larsen, "Insights into the composition of ancient egyptian red and black inks on papyri achieved by synchrotron-based microanalyses," *Proceedings of the National Academy of Sciences*, vol. 117, no. 45, pp. 27825–27835, 2020.

[34] V. Daniels, R. Stacey, and A. Middleton, "The blackening of paint containing egyptian blue," *Studies in Conservation*, vol. 49, no. 4, pp. 217–230, 2004.

[35] G. Linderman, M. Rachh, J. Hoskins, S. Steinerberger, and Y. Kluger, "Efficient algorithms for t-distributed stochastic neighborhood embedding," *ArXiv*, vol. abs/1712.09005, 2017.

[36] W. Lai, J. Huang, N. Ahuja, and M. Yang, "Fast and accurate image super-resolution with deep laplacian pyramid networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, pp. 2599–2613, Nov. 2019.