Machine Learning for Cultural Heritage Conservation: Decoding the Past through Analysis of Hyperspectral Data

Presentata da
LINGXI LIU

Coordinatore Dottorato
Prof. DONATELLA RESTANI

Supervisore
Prof. SILVIA MIRRI

Co-supervisore
Prof. GIOVANNI DELNEVO

ESAME FINALE ANNO 2024

This dissertation is submitted for the degree of Doctor of Philosophy in Beni Culturali e Ambientali

May 2024

The further back we go in history, the less sharp is the distinction between images and real things

— Susan Sontag

Le immagini, che costituiscono l'ultima consistenza dell'umano e il solo tramite della sua possibile salvezza, sono anche il luogo del suo incessante mancare a se stesso.

— Giorgio Agamben

Acknowledgements

Abstract

Cultural heritage conservation and restoration stand at the crossroads of art, history, science, and technological progress. Traditional conservation and restoration methods, while invaluable, are often constrained by their ability to adapt to the complexity of cultural artefacts. In the contemporary landscape, analytical instruments and computational technology are profoundly shaping the interdisciplinary field. Machine learning and data analysis, with their capacity to unravel intricate patterns and trends within extensive datasets, offer a promising avenue for enhancing conservation and restoration practices. This thesis aims to investigate the potential and limitations of machine-learning techniques in processing extensive hyperspectral data acquired from historical art objects. Through various case studies presented, we assess the effectiveness of machine learning models, from off-the-shelf techniques to self-developed algorithms, in supporting tasks ranging from material diagnostics and classification to mapping and digital restoration. Additionally, we critically evaluate the challenges and limitations associated with the implementation of machine learning, specifically constrained by CH, and explore the transferability of machine learning models to similar scenarios. Rooted in the obtained results, this thesis contributes to the ongoing dialogue on leveraging cutting-edge technologies to preserve and celebrate our diverse cultural heritage.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

In this chapter, we present the background and context of this thesis and introduce the research questions that drove the study. Then, we summarise the contents of each chapter. Finally, the academic activities carried out and the list of publications are provided.

1.1 Research Background

Cultural Heritage (CH) conservation is a multidisciplinary field that involves collaboration between historians, archaeologists, scientists, and technicians to enhance the understanding of our CH and safeguard its integrity for future generations [124]. Starting from the mid-20th century, Scientific Conservation, which is based on the principles of authenticity, objectivity, and reversibility, has become the dominant working methodology of the field [142, 96]. Material diagnostics, in this context, stand at the heart of understanding the physical composition, the current states, and the degradation mechanism of the artworks, providing fundamental information for subsequent preservation and restoration. This process is traditionally rooted in empirical and knowledge-driven methods, which rely on a meticulous examination that involves micro-sample taking from representative points and subsequent scientific experiments to decipher the composition and characteristics of the samples, hence the artworks [132, 131].

Over the last decades, the development of imaging spectroscopy has brought a new perspective on how researchers engage with CH, allowing the registration of detailed material signatures of the entire artwork simultaneously [115, 66]. Hyperspectral imaging (HSI) is one of the non-invasive analysis tools that register a spatial map with spectrum collected at each pixel position [74, 6]. It characterises detailed signatures of the material, attributed to the fine spectral resolution down to nanometer scale and wide sampling range varying from X-ray to Infrared [61]. Due to its superior capacity to document and reveal rich information from the inherently complex and delicate art materials [74], HSI has found a fruitful application to identify and characterize colourants in a variety of media, including paintings, illuminated manuscripts, murals, and prints [39, 22, 60].

However, handling the overwhelming volume of data, ranging from thousands to millions of recorded spectra, and the nuanced details they encompass presents an intricate challenge to conservators. The growing importance of imaging spectroscopies has led the way to a new era of data-driven analysis, making the conventional examination methods inadequate [133]. In this context, Machine Learning (ML) algorithms emerge as promising tools for heavy spectral data analysis, offering innovative solutions that directly correlate input data and output results, minimizing the need for extensive knowledge of chemical reactions and compounds [12]. While AI and ML have become integral to modern life in various social sectors, their integration into conservation practices has been more recent, commonly applied as a 'black-box' [72]. The limited adoption is partially attributed to the cautious approach necessitated in dealing with the delicate and irreplaceable art pieces, while also influenced by unique challenges presented in CH studies. Art materials, distinguished by their diversity in nature, heterogeneity in compositions, multi-layering in structure, and degradation influenced by preservation conditions, exhibit characteristics that vary from one piece to another. The uniqueness of problem faced by each case requires special adaptation of ML algorithms, which increases the barrier to apply state-of-the-art techniques in computer science and renders a one-size-fits-all model impossible.

Since the output of the algorithms has usually strong links to the underlying chemical nature, the interpretability of model is also greatly challenged. Furthermore, obtaining a ground truth in CH can be challenging and the publicly available labelled datasets are limited, Unlike applications such as object detection or language translation, which benefit from large standardized open datasets [27]. These distinct challenges in CH analysis make the direct application of ML tools complex, thus demanding deeper exploration and adaptation to the specific context.

Despite these challenges, a range of cutting-edge data processing techniques has already seen promising implementation in solving the existing challenges in the CH sector. Principal components analysis (PCA) and minimum noise fraction transform (MNF) are the data analysis tools most widely used by conservation scientists for their speed and simplicity in reducing the dimensionality of HSI data [36, 147, 90]. Clustering techniques and classifier have also been employed to find the trends within data, including K-means clustering [52, 36], spectral angle mapper (SAM) [55, 63], support vector machine (SVM) [80, 153], and fully constrained least square (FCLS) [78, 162]. More advanced machine learning models, from embedding techniques [156, 202] to neural networks (NNs) [123], are applied for spectral signal unmixing and paint component mapping. In this ever-evolving landscape, ML applications in CH are witnessing steady growth, contributing meaningfully to the ongoing dialogue on leveraging computational technology for the preservation and conservation of our cultural heritage.

1.2 Research Questions

Conservation of cultural heritage stands at the intersection of physical science and technological advancement. Machine learning and data analysis, with their capacity to discern patterns and trends within extensive datasets, offer a promising avenue for enhancing conservation and restoration efforts. However, cultural heritage presents essential constraints in terms of their intrinsic vulnerabilities, diversity and complexity. Challenges may range from issues related to data qual-

ity and quantity to model interpretability and transferability. This thesis delves into a nuanced exploration under the umbrella research question **RQ-1**, seeking to unravel the multifaceted ways in which machine learning can contribute to cultural heritage conservation:

How can machine learning and data analysis techniques contribute to the conservation and restoration of cultural heritage objects?

The investigation involves sub-questions that address specific aspects of this overarching theme, including conservation analysis, restoration applications, and model transferability. By delving into these sub-questions, we seek to provide a comprehensive understanding of the challenges, limitations, and applicability of these advanced computational methods to pave the way for effective implementation.

RQ-1.1 *Can machine learning techniques address complex identification and classification problems in CH?*

Understanding the materials used at the time of execution of the artwork, their distribution, and the current preservation states stand at the heart of effective conservation efforts in cultural heritage. Traditional identification methods, often relying on chemical analyses involving micro sample-taking, come with inherent limitations and become increasingly inaccessible in the field due to their destructive nature [127]. With the introduction of non-invasive and non-destructive analytical instruments, new possibilities for looking into the elemental and molecular compositions emerge [23]. However, the growing volume of big data exceeds the human capacity for the effective extraction of information. While statistical methods may be adequate for simpler problems, machine learning offers a promising avenue to address more nuanced and complex scenarios, where the labelled datasets are limited and the ground truths are intricate due to mixtures and degradations. In this way, we aim to explore the adaptability of ML in addressing these distinctive challenges.

RQ-1.2 *To what extent can machine learning models detect and mitigate the degradation of cultural heritage artefacts over time?*

Degradation is a complex process that presents significant threats to the preservation of cultural heritage objects. The assessment of chemical and structural changes is essentially challenging due to the presence of multiple degradation pathways and the co-existence of diverse degradation products, particularly in the case of large organic molecules [136]. In this sense, imaging spectroscopy and machine learning techniques emerge as promising tools, offering new possibilities to discriminate alterations and subtle changes without the need for laborious characterization processes. Moreover, while conventional restoration is time-consuming and highly dependent on the restorer's skill, digital restoration enabled by machine learning is increasingly considered a fast and subjective restoration solution. The effectiveness of machine learning algorithms in detecting, precisely mapping, and reconstructing damaged areas is to be explored, shedding light on their potential to advance restoration practices.

RQ-1.3 *How transferrable are the machine learning models developed to similar scenarios in cultural heritage conservation and what future applications can be envisioned?*

One of the most important features of cultural heritage artefacts is their uniqueness from piece to piece. Unlike standardized datasets common in many other ML applications, the large variations of CH objects, such as pigment palettes, degradation patterns, dimensions, and artistic styles, make it challenging to extend the applicability beyond every single case. Moreover, the scarcity of comprehensive databases and limited labelled datasets further restrict the models to adapt across a broader range. Despite these challenges posed by the nature of cultural heritage, the specific machine learning models remain applicable within similar scenarios that share the same problem or material composition. Thus, understanding the transferability of these models is crucial for their practical utility

and implementation. We seek to investigate the generalization ability of machine learning models, exploring their efficacy in diverse cultural heritage conservation contexts.

1.3   Outline

Given the background concepts exposed above, a detailed overview of the work presented in this thesis is provided below:

Chapter 1 The first chapter introduces the context of the thesis and presents the research questions. Lastly, it summarises the content of each chapter and the publications derived from the findings.

Chapter 2 This chapter delves into the exploration of machine learning algorithms for the task of classifying and identifying complex art materials. An innovative data processing pipeline is presented, integrating spatial and spectral clustering techniques to extract and map spectral signatures. We explore Simple Linear Iterative Clustering (SLIC) Superpixel's potential in reducing hyperspectral data while maintaining spectral richness. This facilitates the application of various machine learning algorithms, especially the soft clustering algorithms Fuzzy C-Means clustering (FCM) and Gaussian Mixture Models (GMM), to classify spectra into distinct colourant groups, identifying complex mixtures and areas of degradation. The efficacy of the proposed method is demonstrated through two highly diverse cases, one of cinematic films and one of pointillism painting. The results highlight the potential of machine learning in aiding artwork diagnostics, conservation, and restoration, with transferable models for similar scenarios. This chapter aims to answer the research questions **RQ-1.1**, **RQ-1.2**, and **RQ-1.3**.

Chapter 3 This chapter explores the application of machine learning to address degradation issues, focusing on a case study involving the digital restoration of degraded cinematic films. A novel Vector Quantization (VQ) algorithm is developed, leveraging VNIR hyperspectral data. The VQ algorithm utilizes a

what we call a multi-codebook that correlates degraded areas with corresponding non-degraded ones selected from reference frames. Comparative analysis with professional commercially available film restoration software reveals superior results in terms of color reconstruction, tested on both RGB and hyperspectral data. The restoration achieved on the single frame is further extended to a collective frames, revealing its potential applicability to a broader spectrum of images exhibiting similar deterioration patterns. The goal of the chapter is to contribute insights that further address the research questions **RQ-1.2**, and **RQ-1.3**.

Chapter 4 In this chapter, we aim to contribute to the systematic application of the state-of-the-art technique, Neural Networks, in solving cultural heritage problems. This chapter provides an exhaustive analysis of the literature related to Neural Networks (NNs) applied for hyperspectral data in the Cultural Heritage field. We outline the existing data processing workflows and propose a comprehensive comparison of the applications and limitations of the various input dataset preparation methods and NN architectures. By leveraging NN strategies in CH, this chapter aims to pave the way for a more systematic and informed application of this novel data analysis method. This chapter aims to answer the research questions **RQ-1.1**, and **RQ-1.3**.

Chapter 5 This chapter summarises the contribution of this thesis, highlights key findings for each research question, and outlines potential directions for further research.

Chapter 2

Spatial-Spectral Joined Unsupervised Segmentation for Colourant Analysis

> Can machine learning techniques address complex identification and classification problems in CH?
> To what extent can machine learning models detect and mitigate the degradation of cultural heritage artefacts over time?
> How transferrable are the machine learning models developed to similar scenarios in cultural heritage conservation and what future applications can be envisioned?
>
> — **RQ-1.1, RQ-1.2, RQ-1.3**

This chapter presents an innovative data processing pipeline that combines spatial and spectral clustering techniques for the extraction and mapping of spectral signatures in both cinematic films and pointillism painting. We explore Simple Linear Iterative Clustering (SLIC) Superpixel's potential in reducing hyperspectral data while maintaining spectral richness. This facilitates the application of various machine learning algorithms, especially the soft clustering

algorithms Fuzzy C-Means clustering (FCM) and Gaussian Mixture Models (GMM), to classify spectra into distinct colourant groups, identifying complex mixtures and areas of degradation.

2.1 Introduction

Art conservation and restoration are complex processes that demand a deep understanding of the materials and conditions of artworks [11]. Traditionally, the identification process requires careful sampling from the artwork, which is usually limited and risky, especially when dealing with precious pieces [76, 31]. In recent years, non-invasive spectral imaging tools have brought new possibilities into the analytical investigation, where the identification and mapping of diverse chemical compounds are enabled by examining characteristic spectral features [58, 20, 6]. Hyperspectral Imaging (HSI) is one of these emerging techniques in the Cultural Heritage (CH) field [152, 115]. It registers a map of the artwork with spectrum collected at each spatial location, forming a 3-dimensional datacube that contains high-resolutional information both spatially and spectrally. However, artworks usually exhibit heterogeneous compositions and complex preservation states, making the identification of material mixtures and the detection of degradation particularly difficult [215]. Furthermore, the precise mapping of such details through the extensive spectral data is a challenging task, demanding more efficient and automatic data segmentation techniques.

In the context of image segmentation, the choice between spatial and spectral features hinges on data characteristics and analytical goals. While spatial features excel in spatially contextualised data like urban planning and object detection, capturing adjacent pixel relationships, spectral features, prominent in hyperspectral data, shine in tasks requiring precise material identification. In the cultural heritage field, spectral-based data processing algorithms have thus found fruitful applications in extracting chemical information from the heavy HSI dataset. Numerous works have applied data dimensionality reduction techniques, such as Principal Component Analyses (PCA) [161], t-distributed

Stochastic Neighbor Embedding (t-SNE) [156], and Uniform Manifold Approximation and Projection (Umap) [202], to compress the high-dimensional HSI data into a lower-dimensional space, facilitating the fast visualization and subsequent classification. However, these approaches often sacrifice critical spectral details and result in condensed and overlapping data representations, complicating the accurate separation of material mixtures. Another common approach is spectrum-wise classification, where techniques like Spectral Angle Mapper (SAM) [63, 68, 14], Fully Constrained Least Square (FCLS) [78, 48, 62], or Neural Network models [123] are employed to match or unmix spectral signatures. These algorithms enable the accurate identification of pictorial materials but are often constrained by the limited availability of labelled training data and the computational cost of pixel-wise measurements. Moreover, hyperspectral data possesses rich information in both spectral and spatial scales. While the spectral aspect has seen significant development in analytical techniques, the spatial aspect is often underutilised. Notably, HSI data exhibit redundancies among spatially adjacent pixels due to similarities in material compositions at small scale. Incorporating the spatial information into classification processes offers great potential to enhance the clustering efficiency.

To address this, we developed a novel joint spatial-spectral data segmentation approach to unravel intricate details in art objects. We leverage the Simple Linear Iterative Clustering (SLIC) Superpixel algorithm [3] to spatially segment the data by grouping repetitive spectra into compact and relatively homogeneous regions, providing an optimal balance between data complexity reduction and spectral signature retention. Subsequently, a range of state-of-the-art unsupervised clustering techniques is explored to reveal deeper insights into the underlying patterns within the spectral data, especially the soft clustering algorithms such as Fuzzy C-Means clustering (FCM) [194] and Gaussian Mixture Models (GMM) [134]. The combination of spatial segmentation with clustering algorithms has already demonstrated efficacy in the computer vision field for improved segmentation accuracy with reduced computational time [110, 184]. While such methods have been mainly developed for colour images, they have

found promising applications in CH, primarily for fast object extraction and edge detection in document, book, and mural RGB images [93, 135, 209, 186, 91]. For analysing the high-dimensional HSI data, this joint clustering approach is increasingly applied in the remote sensing field to assist urban or terrestrial data classification [28, 88, 92]. Recently, it has also been extended in the domain of CH analysis, particularly in feature delineation within mural paintings in assist for the future restoration works [125, 113]. However, the full capacity of spatial-combined classification to address the challenges of complex material mixtures and degradation has yet to be fully explored.

The efficacy of our proposed approach will be demonstrated through its applications on two distinct historical artworks: one of cinematic films that exhibit non-uniform colour loss and colour balance shifts, and one 19th century Neo-Impressionist painting that presents highly diverse pigment compositions with numerous mixtures on canvas. Through this study, we introduce this novel joint spatial-spectral data segmentation technique to the cultural heritage field to explore its full potential in artwork diagnostics, conservation, and restoration.

The remainder of this chapter is organized as follows. Section 2.2 highlights the research questions to be solved in this chapter. Section 2.3 provides comprehensive information on the materials and spectral datasets utilised in this chapter. In Section 2.4, we detail the data processing methodology, starting from the developed algorithm to the visualisation strategy and comparisons with other approaches. The segmentation results and their analyses are discussed in Section 2.5. Finally, the conclusion and future perspectives are marked in Section 2.6.

2.2 Research Questions

The goal of this project is to investigate the potential of spatial clustering-based clustering algorithms and soft clustering techniques in assisting cultural heritage material classification. Through a more nuanced understanding of art materiality, by identifying distinct materials, separating their mixtures, and localising the areas of degradation, the following research questions are responded:

RQ-1.1 *Can machine learning techniques address complex identification and classification problems in CH?*

RQ-1.2 *To what extent can machine learning models detect and mitigate the degradation of cultural heritage artefacts over time?*

RQ-1.3 *How transferrable are the machine learning models developed to similar scenarios in cultural heritage conservation and what future applications can be envisioned?*

2.3 Experimental

2.3.1 Materials

Film samples used in this chapter are provided by "L'immagine Ritrovata Film Restoration Laboratory" (Cinematique of Bologna). They come from a historical film fragment manufactured by Fuji Photo Film Co., Ltd. between April and June 1986, with the authorship undetermined. In this case study, four positive film frames, each measuring 34 mm × 19 mm (± 0.5 mm) in physical dimension, are presented (Figure 2.1). The transparent support for these films is cellulose acetate, and the yellow, magenta, and cyan dyes are located in distinct layers within the emulsion, forming a trichromatic structure [208]. However, those large synthetic organic molecules are intrinsically unstable and deterioration is inevitable, making the preservation of the original colour a challenging problem. Moreover, the degradation is non-linear among the dyes, which means certain types of dyes degrade at faster rates over time, and is often spatially inhomogeneous that certain areas are more damaged than the rest. Among them, the cyan dyes are the most unstable to degrade, followed by the magenta and yellow dyes. As could be observed in Figure 2.1, the different stages of degradation are reflected in the gradual transformation of the film's colour, from its original bluish hue (as seen in Film 21) to shades of purple (notably in the left half of Film 23), followed by magenta dominance (most evident in Films 22 and 25),

and eventually a yellowish appearance (as indicated by the irregularly shaped stains observed on Films 22 and 25). The complexity of the non-linear and inhomogeneous deterioration makes the spectroscopic measurement preferential to the simple RGB values, by providing extended information on the colour loss based on high-resolution spectral signatures. In this context, digital restoration is advantageous to bypass the challenging and laborious manual restoration procedure, where the successful identification and segmentation of degradation areas are essential steps.

Fig. 2.1 The optical images of the film frames analysed.

The 19th century painting "Au temps d'harmonie (l'âge d'or n'est pas dans le passé, il est dans l'avenir)" (In the time of harmony: The golden Age is not in the past, it is in the future), oil on canvas, 300 x 400 cm by Paul Signac (ref 253 catalogue raisonné)[33] currently resides in the Town hall of Montreuil (Figure 2.2). It was executed between 1893 and 1895 and is nowadays one of the masterpieces of the neo-impressionist artistic movement. The Neo-Impressionists drew inspiration from the colour theories of Eugène Chevreul (1839) [46] and Ogden Rood (1879)[168], and as such involved juxtaposing small dots of pure

colours, sometimes complementary or of varying tones, to obtain the most luminous and vibrant paintings [182]. The so-called "pointillism" painting technique makes the spatial feature-based data reduction especially advantageous, while this large-size painting still exhibits a vast range of variations, including varying intensities rising from distinct pigment compositions and multiple mixtures on the canvas. To date, comprehensive material and colourimetric analysis of this painting has not been undertaken, despite its significance in understanding Paul Signac's creative process, his colour palette, and the practical application of colour theory on his canvas.

Fig. 2.2 The original Signac painting and the scanned part (marked in the black rectangle).

2.3.2 Data Acquisition

The spectral data of film samples were scanned in the University of Ljubljana by a custom-made hyperspectral camera, which uses the push-broom system with 2,448 across-track pixels and 2,048 wavelengths, composed of an imaging spectrograph (ImSpector V10E, Specim, Finland) and a monochromatic camera (Blackfly S, BFS-U3-51S5M-C, 5 MP, FLIR Integrated Imaging Solutions, Canada)[65]. Spectral images of the film samples were acquired in a reflectance geometry with a broadband LED light source covering the spectral range from 380 nm - 1000 nm. The spectral sampling rate is 0.3 nm in the spectral range of 322.9 - 1025.3 nm, where the considered effective range is from 380 nm to 780 nm. The effective spectral and spatial resolutions of the system were 2.9 nm and 100 μm, respectively, as evaluated by a gas discharge tube and spatial grids used for system calibration. Simple calibration and normalization of the spectra are performed at the time of registration. The initial data contained a lot of noise. To improve the spectral quality, the datacubes were binned by five spectrally and the excessively noisy wavelength range of less than 400 nm and over 750 nm was cropped. Spatially, the uninformative border and perforation areas were also removed, keeping only the image content. Each resulting datacube has a size of 1200 pixels (hor.) x 1000 pixels (ver.) x 180 channels.

Hyperspectral imaging dataset of the 19th century painting, Au temps d'harmonie, was acquired *in situ* with a VNIR (Visible near InfraRed) push-broom PFD4k camera by © Specim (Oulu, Finland) covering a spectral range from 400 to 1000 nm with 212 bands wavelength channels (spectral resolution of 3.0 nm). In order to scan this large painting, difficult to access, the camera has been equipped with a long working distance objective, i.e. OLE 140 (focal length 140 mm). The assembled system was mounted on a motorized rotation stage, fixed on a tripod, and set at a distance of 8.7 meters from the painting. The acquisition parameters were set with a scanning speed of 0.02°/s, a frame rate of 5 Hz, an exposure time of 190 ms, and a spectral binning of 4. It resulted in about a 30-minute scan, with a pixel size of 0.50 (v.) x 0.59 (h.) mm. The illumination was provided by two halogen lamps of 2 x 400 W, located 8 meters away from the painting with

a 45° incidence angle. Additionally, twenty LED lamps present within the city hall on both sides of the canvas could not be turned off and participated in the illumination of the painting under study. The data were normalized with dark and white references using the Specim plug-in in ENVI. The white reference corresponded to the white frame surrounding the painting and was saved at the same time as the acquisition. A spectralon was first compared to the frame reflectance signature to ensure the proper diffuse Lambertian properties of the white target. Four acquisitions led to the hyperspectral cube of more than three-quarters of the painting in less than 4 hours. As highlighted in Figure 2.2, only the upper part of the painting and several centimetres of the lowest part are missing due to the limited height of the tripod. The first pre-processing of data consisted of the manual stitching of the four datasets using the ENVI software, as well as a spectral crop of the dataset between 440-850 nm and a smoothing of the spectra with Savitzky-Golay filter (13 number of points, 2 polynomial degree) using the Spectronon (©Resonon) software. The final cube is composed of 6481 pixels (hor.) x 4457 pixels (ver.) x 143 channels.

2.4   Data Processing Methodology

The comprehensive data processing pipeline, illustrated in Figure 2.3, include several key steps, each detailed in the subsequent subsections. First, the acquired HSI data are spatially clustered into homogeneous and compact regions, employing the superpixel technique. Then, a suite of state-of-the-art unsupervised spectral clustering techniques, starting from the conventional hard clustering technique (Hierarchical clustering) to the more advanced soft clustering algorithms (Fuzzy C-means and Gaussian mixture models), is tested based on the centroid matrix extracted from each superpixel. This proposed approach is systematically employed and analysed across two distinct case studies. In the film case, the performance of the proposed method is compared against one grid-based data reduction approach, exemplified by the application of hierarchical clustering. In the painting case, the efficiency of the proposed technique is compared with a

widely used supervised algorithm, Spectral Angle Mapper. The outcomes of these analyses are comprehensively presented and examined through an adopted visualization strategy.
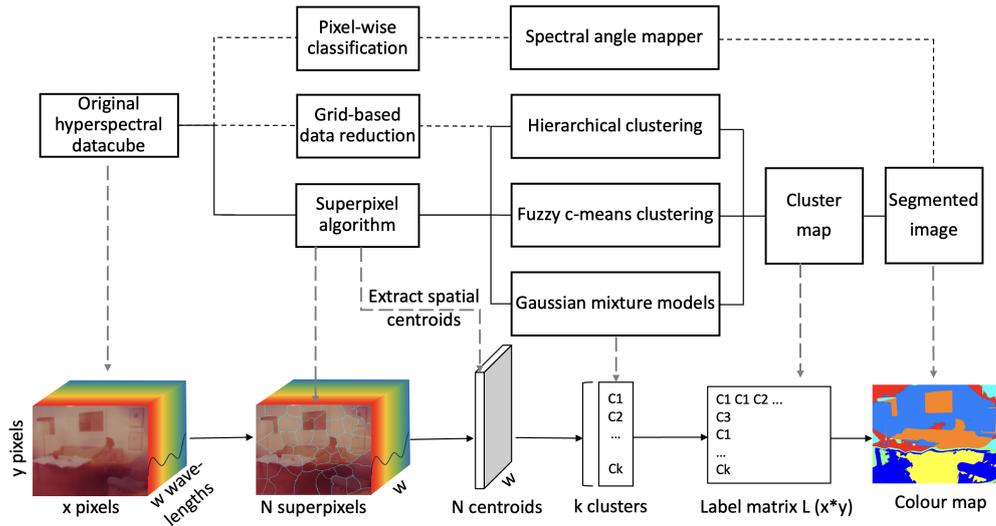
Fig. 2.3 Schematic overview of the data processing methodology employed in this chapter.

2.4.1 Spatial Clustering

SLIC (Simple Linear Iterative Clustering) Superpixel [3] is a widely used algorithm for image segmentation to simplify image processing operations. It divides the image spatially into a number of small and compact regions, within each the superpixel is relatively homogeneous. It has a high potential to reduce the hyperspectral data size, since it can effectively extract the information out of the datacube by grouping similar and repetitive spectra while keeping the full spectral signatures. SLIC works based on the following principles: First, the cluster centres are initialised, i.e. evenly distributed throughout the image, based on the chosen number of superpixels (K). Then, for each cluster centre, the distances between the centre and the pixels in a local neighbourhood are calcu-

lated taking into account both the colour similarity and spatial proximity. Each pixel is assigned to the closest cluster centre. Afterwards, the cluster centres are recalculated and updated by taking the mean colour and spatial coordinates of all pixels assigned to them. These assignment and update steps are iteratively repeated until reach convergence. During the iterations, the boundaries between superpixels are continually refined also according to the compactness parameter, which specifies the regularity of the superpixel, i.e. the lower the compactness, the better adherence to object boundaries and the more irregularly shaped. A regularisation term is also applied to ensure that each superpixel region is continuous and connected. In the end, a label matrix (L_{slic}) is generated that indicates the membership of each pixel to the superpixels, with the real number of superpixels (N) usually slightly smaller than the input K.

In this step, colour images are used as inputs for simplicity and efficiency during calculation. Multi-channel HSI data are projected and converted respectively into 3-dimensional RGB colour space in the film case and L*a*b* colour space in the Signac painting case using the code provided in [128]. In the latter case, the CIELAB coordinates are preferred, as they enable a subsequent characterisation of the distribution of tone contrasts (L*) and hue variations (a* and b*) in Signac's work. Those resulting images are of the same spatial composition as the HSI cube, and the obtained superpixel label grid L_{slic} could be directly transferred to the HSI data. Then, a spatial centroid approach is adopted to prevent the loss of critical information when averaging the spectra within each superpixel. Specifically, the coordinates of pixels within each superpixel are averaged to get the spatial centre, and the spectra at the spatial centres are extracted to form a new spatial centroid matrix in dimensions of N (number of superpixels) x w (channels). This guaranteed the preservation of unmixed spectral signatures and ensured the precision of material identification. In this way, the HSI data is vastly down-sampled without losing the spectral resolution. The centroid matrix is used as the input in the spectral clustering stage.

Alternatively, one grid-based data reduction is provided as a comparison to the superpixel segmentation. The datacubes are decomposed into small sub-

images by taking geometrical grids in the size of 25x25 pixels. The centroids are extracted for each sub-image in the same way to form the input matrix. This approach is only demonstrated for the films using hierarchical clustering.

2.4.2 Hierarchical Clustering

The Hierarchical Clustering Algorithm (HCA) is an unsupervised hard clustering approach used to divide the data into nested levels of clusters [143]. Alternative to the most common practices, such as k-means clustering, HCA does not require the specification of a desired number of clusters beforehand. Instead, it forms a hierarchy of clusters that enables the exploration of different levels of details in the data set. The full process is described step-by-step in the following paragraphs.

Distance metrics First, HCA regards each spectrum in the input dataset as a single cluster and calculates the distances of every spectrum to each other. Three different metrics are available for measuring the distances between every pair of initial clusters: Euclidean, Chebyshev, and Cosine distance. Euclidean and Chebychev distance are specialized forms of Minkowski distance (with $p = 1$ and $p = \infty$ respectively) [81], where the distance d between two spectra s_r and s_t of n channels is given by

$$d = \sqrt[p]{\sum_{i=1}^{n} |s_{ri} - s_{ti}|^p}. \tag{2.1}$$

For Cosine metric, the distance d of two spectra s_r and s_t is instead given by:

$$d = 1 - \frac{s_r s_t'}{\sqrt{(s_r s_r')(s_t s_t')}}. \tag{2.2}$$

The distance matrix is used to determine the similarity and proximity of the spectra, where the closest two spectra are grouped into a new cluster.

Linkage   Then, the newly formed cluster is considered as one large cluster in the following step, and the intra-cluster distances are repeatedly recalculated with a specified linkage method. Four different linkage functions are selected for the tests: Complete, Single, Average, and Weighted. In Complete agglomeration linkage, the distance *d (p,q)* between two clusters *p* and *q* is defined by looking for the farthest distance between objects in two clusters

$$d(p,q) = \max \left(\text{d} \left(s_{pi}, s_{qj} \right) \right), i \in (1, \dots, n_p), j \in \left(1, \dots, n_q \right) \qquad (2.3)$$

where n_p is the number of objects in cluster *p*, n_q is the number of objects in cluster *q*, s_{pi} is the i_{th} spectra in cluster *p*, and s_{qj} is the j_{th} spectra in cluster *q*. On the opposite, the Single linkage method finds the smallest distance between objects in two clusters and is also known as the nearest neighbour. The Average linkage method is defined by averaging the distance between every pair of spectra in two clusters *p* and *q*:

$$d(p,q) = \frac{1}{n_p n_q} \sum_{i=1}^{n_p} \sum_{j=1}^{n_q} \text{dist} \left(s_{pi}, s_{qj} \right), i \in (1, \dots, n_p), j \in \left(1, \dots, n_q \right) \qquad (2.4)$$

where n_p and n_q are the numbers of objects in cluster *p* and *q*, s_{pi} is the i_{th} spectra in cluster *p*, and s_{qj} is the j_{th} spectra in cluster *q*. The Weighted linkage then measures the weighted average of the components that have been linked in the former step and divides the distances of these two components against another cluster by 2.

Clusters   During the iterations, HCA repeatedly merges the pair of clusters with the smallest distance into larger clusters, until it reaches only one cluster that contains every single initial cluster in the dataset. The grouping index and intra-cluster distance information were recorded during each step in the process, and this information forms a hierarchical tree that can be plotted in a dendrogram. The tree-shaped graphics show clearly how the initial clusters are developed in a multilevel hierarchy. The X-axis indicates the original cluster index, and

the Y-axis represents the dissimilarity of clusters at each node. The height of each node is proportional to the intra-cluster distance of the two leaves that are connected. Finally, the cluster assignment of the data points could be obtained by cutting the hierarchical tree at a given level or threshold. The entire dataset is classified into two clusters at the level of node 1, three clusters at the level of node 2, and so forth. Each cluster is a collection of multiple initial clusters; in other words, every initial cluster is assigned to a specific cluster group. The HCA is performed in Matlab using the Statistics and Machine Learning Toolbox™ [195].

2.4.3   Fuzzy C-Means Clustering

Fuzzy C-Means clustering (FCM) is a commonly used soft clustering technique that extends the traditional hard clustering to handle uncertainty in cluster assignment [79]. Unlike hard clustering which divides data into distinct clusters, FCM allows data points to belong to multiple clusters in varying degrees, making it especially suitable for pigment mixtures where the division of data is frequently ambiguous. FCM works based on an iterative operation. The optimal number of clusters (K) is determined using the subtractive clustering function in Matlab [196], which estimates the total variations of data in FCM models [47]. With a specified number of desired clusters (K), the FCM algorithm initialises with a random assignment for cluster centres μ. Then, the membership possibility $p_{i,j}$ for each spectrum s_i to each cluster centre μ_j are calculated based on the fuzzy c-means formula:

$$p_{i,j} = \left(\sum_{k=1}^{K} \left(\frac{\|s_i - \mu_j\|}{\|s_i - \mu_k\|} \right)^{\frac{2}{m-1}} \right)^{-1} \tag{2.5}$$

where m is the fuzziness exponent that defines the overlapping degree of the clusters. Afterwards, the cluster centres μ are updated based on the new membership

degrees $p_{i,j}$

$$\mu_i = \frac{\sum_{j=1}^{N} p_{ij}^m s_i}{\sum_{j=1}^{N} p_{ij}^m}, i \in (1,\ldots,K) \tag{2.6}$$

where N is the total number of spectra in the dataset. By iteratively updating the cluster centres and recalculating the membership degrees, the objective function J_m is gradually minimised:

$$J_m = \sum_{i=1}^{K} \sum_{j=1}^{N} p_{ij}^m D_{ij}^2 \tag{2.7}$$

where $D_{i,j}$ is the distance of the j_{th} spectrum to the i_{th} cluster, employing Euclidean metrics. The algorithm finds the optimal partition of the dataset upon convergence when the objective function J_m improves less than a specified minimum threshold. Finally, FCM outputs the list of cluster centres μ and a fuzzy partition matrix p that specifies the cluster membership grades for each data point. The experiment is performed in Matlab using *fcm* function [194].

2.4.4   Gaussian Mixture Models

Gaussian mixture model (GMM) is further explored as a more advanced soft clustering technique that employs probabilistic models to structure complex data distribution [134]. In GMM, each cluster generated is fitted into a Gaussian distribution component defined by its mean μ and covariance matrix Σ, where a vector of mixing coefficients π defines the mixture of different components. For a specified number of components (K), the initial values for component means μ are initialised by applying the k-means algorithm [10], where the covariance matrices are set to be diagonal and identical and the mixing coefficients are set to be uniform for all components. With those initial parameters, the model is iteratively refined using the Expectation-Maximization (EM) algorithm [134], which consists of two main steps: the E-step (Expectation) and the M-step (Maximization). In the E-step, the posterior probability $p_{i,j}$ for the i_{th} spectrum

s_i belonging to j_{th} Gaussian component is given by

$$p_{i,j} = \frac{\pi_j \mathcal{N}\left(s_i \mid \mu_j, \Sigma_j\right)}{\sum_{k=1}^{K} \pi_k \mathcal{N}\left(s_i \mid \mu_k, \Sigma_k\right)}, \tag{2.8}$$

where $\mathcal{N}\left(s_i \mid \mu_j, \Sigma_j\right)$ is the Gaussian probability density function [134]. In the M-step, the algorithm parameters, including the component means μ, covariance matrices Σ, and mixing proportions π, are updated using the posterior probabilities p as weights:

$$\pi_j = \frac{N_j}{N} \quad \text{with} \quad N_j = \sum_{i=1}^{N} p_{i,j}, \tag{2.9}$$

$$\mu_j = \frac{1}{N_j} \sum_{i=1}^{N} p_{i,j} s_i, \tag{2.10}$$

$$\Sigma_j = \frac{1}{N_j} \sum_{i=1}^{N} p_{i,j} \left(s_i - \mu_j\right) \left(s_i - \mu_j\right)^{\mathsf{T}}, \tag{2.11}$$

where N is the total number of data points. These steps are iteratively repeated until the log-likelihood of all data points reach convergence. After the training stage is completed, the optimised component means μ serve as the cluster centres and the posterior probability matrix p contains the membership possibilities to each component of each data point. Furthermore, the output GMM model could be applied to a new data set to predict the cluster belongings of unseen spectra. The training is performed in Matlab using *fitgmdist* function [195], with regularisation value set to 0.01 and other algorithmic parameters as default.

2.4.5 Spectral Angle Mapper

Spectral angle mapper (SAM) is one of the most widely used spectral classification and mapping techniques in CH field [59], [60]. This is a supervised algorithm working based on the principle of comparing the spectral signature of a targeted spectrum in the HSI cube to reference spectra. In this process, each

spectrum is transformed into a vector in an n-dimensional scatter plot, where n is the number of wavelengths. The angle α between the reference spectra r (endmembers) and the observed spectra t are then calculated using the *arccosinus* function [105]

$$\alpha = \cos^{-1}\left(\frac{\sum_{i=1}^{n} t_i r_i}{\sqrt{\sum_{i=1}^{n} t_i^2}\sqrt{\sum_{i=1}^{n} r_i^2}}\right) \tag{2.12}$$

The smaller the angle, the greater the similarity between two spectra (endmember and observed). An angle threshold is set manually such that the spectrum is classified as similar to an endmember if its angle is smaller than the fixed threshold. The SAM was performed for the painting object solely using the software Spectronon (©Resonon). Ten different endmembers have been selected manually.

2.4.6   Visualisation

After obtaining the clustering results, a post-processing step is performed to visualise the image segmentation. The cluster centres are first extracted and plotted to examine the spectral signatures of each group determined. For HCA, the cluster centre is determined by averaging all the spectra assigned to it, while for FCM and GMM, the cluster centres are inherent components of the algorithm and are automatically generated as part of the output. Then, each cluster label is assigned a specific colour to create the colour map. In the case of the film dataset, the jet colour scale is employed to enhance the differentiation among various degradation groups. Conversely, in the analysis of the Signac painting, the actual visual response, represented by RGB values, is calculated for each centre spectrum by transforming the hyperspectral vector to RGB space using the method described in [128]. This approach provides a realistic and intuitive representation of the pigment groups, where the different colours and mixtures are visually identifiable.

To perform image segmentation based on the clustering results, each pixel in the image is assigned a label corresponding to the cluster of the superpixel's

centroid. In other words, if the centroid of the j_{th} superpixel is assigned to Cluster C_k, all the pixels within that superpixel are labelled with Cluster C_k. For soft clustering algorithms FCM and GMM, where each spectrum belongs to multiple clusters simultaneously, the cluster with the maximum possibility is assigned. This process propagates the cluster assignments from centroids back to individual pixels, resulting in the creation of a label map. Subsequently, each pixel is coloured with a designated colour, leading to image segmentation where each cluster is visually represented by a unique colour.

2.5   Results and Discussion

The segmentation outcomes achieved on both the film and painting datasets are detailed in the following subsections. Within each case study, a thorough discussion of the proposed joint clustering approach is presented, accompanied by a comparative analysis against common practices.

2.5.1   Segmentation on Degraded Films

Spatial clustering results

First, SLIC superpixel segmentation was performed on the four film frames, 21, 22, 23 and 25. Two critical input parameters, the desired number of superpixels (k) and the compactness, were investigated for the impact on segmentation. A high number of superpixels can significantly increase the computational load in subsequent processing steps, while an insufficient number can result in inadequate division of sub-areas, such as improper grouping of incoherent pixels and imprecise adhering to object boundaries. The compactness influences the shape of superpixels, with the higher number leading to more regularly formed contours.

As illustrated in Figure 2.4, presenting frame 22 as an example, various k values ranging from 500 to 2,000 superpixels and the compactness from the lowest to highest were tested. As shown in Figure 2.4a, when using an insufficient

number of superpixels, larger segmented areas are achieved, sacrificing the finer details. For instance, the thin painting frames are not clearly defined and have merged with nearby background pixels on the wall. By increasing the number to 2,000 (Fig. 2.4c), the finest contours of the human figure, the support of the lamp, and the painting frames are successfully captured and it was determined to be sufficient. On the other hand, with the lowest compactness (Fig. 2.4a-c), the generated superpixel grids are irregularly shaped, conforming to the contours of image content. Conversely, when the compactness increased to 20 (Fig. 2.4e), the segmented superpixels were more geometrically shaped and sacrificed the fine details. Thus, the number of 2,000 superpixels and compactness of 1 were chosen for the following tests. Consequently, the four film frames were segmented into 1,822, 1,864, 1,898, and 1,882 superpixels, averaging 659, 644, 632, and 638 pixels per superpixel, respectively. The spectral centroids of each superpixel were extracted and concatenated together to form the input matrix for the following spectra clustering. This reduces the data set size from 4.8 million spectra to 7,466 spectra with variations coming from all degradation areas, maintaining full spectral resolution.

As a comparison, the simple grid-based data reduction is also applied (Fig. 2.4f). Grid size of 25x25 pixels was chosen to obtain a comparable size of divided areas (625 pixels per sub-image). Similarly, the spectral centroid of each sub-image is averaged and extracted to form the input matrix, obtaining 7,680 spectra combining all four frames.

Parameter analysis via grid-based approach

The initial phase of the tests involved assessing the effectiveness of various clustering algorithms on a single film frame. In contrast to paintings, where areas painted with distinctive pigments exhibit unique chemical compositions and distinctive spectral characteristics, all pixels in a film share similar chemical constituents and spectral responses. These constituents consist of yellow, magenta, and cyan dyes, each combined in different proportions at every pixel during exposure to imaging objects. Their visible spectra all share similar broad
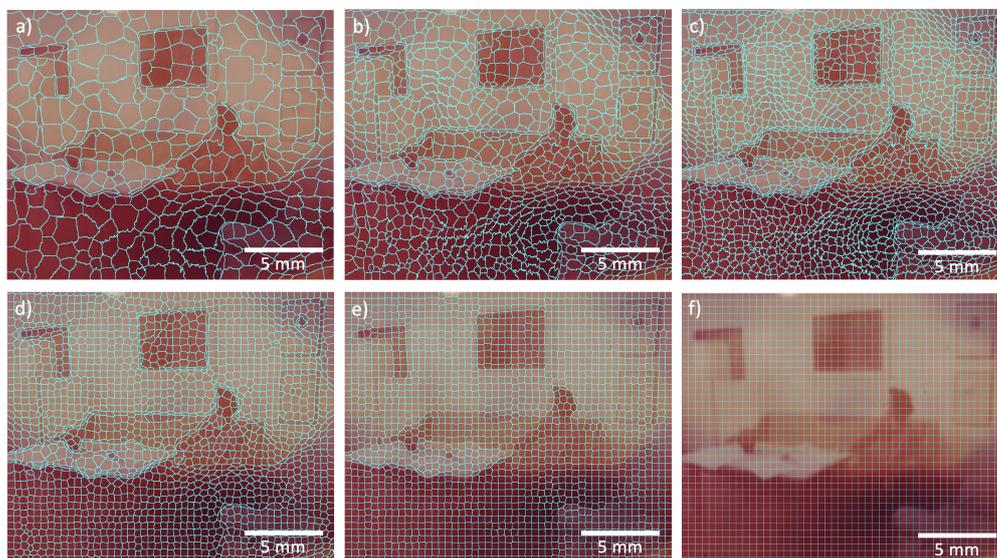
Fig. 2.4 Spatial segmentation results obtained on frame 22: a-c) 500, 1,200, and 2,000 superpixels with the compactness of 1, d-e) 2,000 superpixels with the compactness of 5 and 20, and f) 25x25 geometric grid.

absorption positions, at around 430 nm (corresponds to the yellow dyes), 550 nm (magenta), and 650 nm (cyan), with variations primarily in intensities. However, the inevitable degradation process introduces non-uniform loss of dye content across the image, while also being non-linear across the three dye types. Notably, the residual dye density is a complex interplay between the original image content and the degree of degradation. This means that pixels subjected to the same level of degradation may exhibit different spectral intensities, and pixels with similar density of residual dyes can vary slightly in spectral signatures. Therefore, while film data sets exhibit fewer spectral variations compared to the painting case (presented in section 2.5.2), the magnitude of colour loss is subtle and nuanced to capture. Hence, the effective segmentation of diverse degradation groups is challenging but crucial for the future digital restoration process.

The spectral clustering methodology using HCA was first developed and tested on grid-extracted centroid data (Fig. 2.4f) to explore the best parameter combinations. As introduced in Section 2.4, in HCA, the optimal number of clus-

ters is explored and determined after the hierarchical cluster tree is constructed. Figure 2.5 illustrates the results obtained on film frame 22 with 10 clusters, through different algorithmic parameter combinations. Three metrics (Euclidean, Chebychev, and Cosine distance) and four linkage methods (Complete, Single, Average, and Weighted) were tested. Due to the lack of ground-truth values, quantitative analysis was not able to be conducted. Nevertheless, distinct degradation features serves as crucial indicators, allowing for a qualitative evaluation of the segmentation effectiveness. As marked on Figure 2.5, the identification and separation of the less degraded corners (A1 and A4), the yellowish stains (A2), and the remained magenta area surrounded by the stain (A3) are major considered.

Out of the twelve combinations tested, the majority yielded successful results in segmenting the image based on dye densities, with the exception of Single linkage. Single linkage tends to generate long and thin clusters due to its reliance on the smallest distance between objects in two clusters, favouring the separation of outliers. This characteristic, advantageous for clustering galaxies in astronomy [70], posed crucial limitations in our context. In contrast, Complete linkage, which relies on the largest distance and produces highly compact clusters, demonstrated superior performance. It is the only one that successfully separated the upper-left corner (A1) from the rest of the stained area. The Average and Weighted Average methods allowed segmentation based on major variances but were less sensitive to small variations compared to the Complete method, only successful in detecting A4 area. On the other hand, the choice of metric had a less pronounced impact compared to the linkage functions, which significantly influenced the shape of final clusters. The results among the three metrics are comparable, with Chebychev metric combined with Complete linkage outperformed the rest, clearly delineating the yellowish stain (A2) and separating the surrounded A3 area.

While this initial application of HCA to the film dataset successfully segmented main degradation features, the precision and accuracy of the segmentation should still be improved. The current limitation lies in the use of fixed
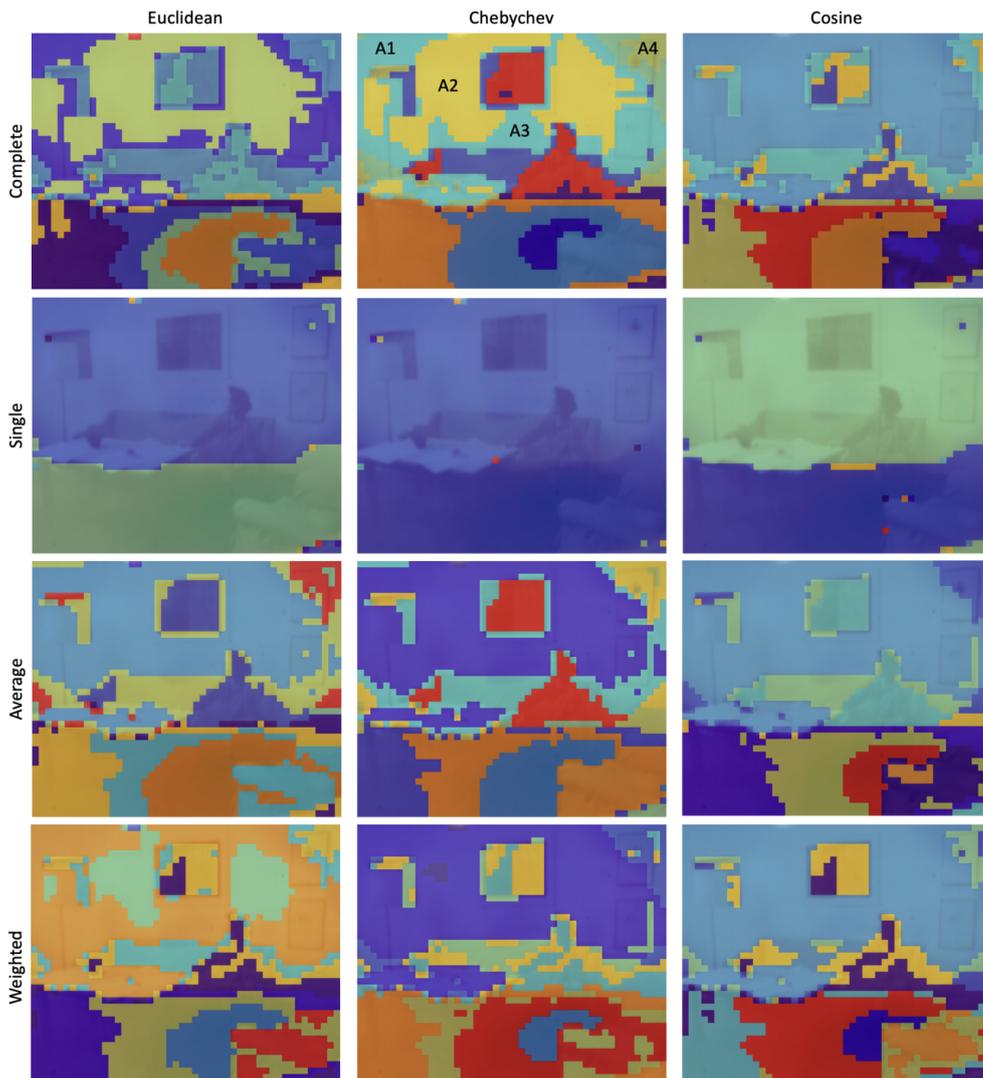
Fig. 2.5 Comparison of the segmentation results obtained on film frame 22 using hierarchical clustering with different combination of algorithmic parameters.

grids for sub-image division, resulting in rigid boundaries between adjacent sub-images and restricting the final image resolution. Given that the restoration requires highly accurate treatment, especially in CH, the spatial relationship of the pixels must be considered in addition to the spectral dimension.

Joint spatial-spectral clustering

The joint spatial-spectral clustering methodology was then tested and applied. The initial phase of the tests involved assessing the effectiveness of various clustering algorithms, first the hard clustering algorithm HCA and then the soft clustering techniques, on a single film frame. After a thorough examination, a threshold of 10 clusters was found to yield the most efficient results. Figure 2.6 a, c, e, and f depict the segmentation results conducted on frame 22, with the 10 clusters (C) mapped back to the original image. An arbitrary colour scale was employed to maximize the visual distinction between clusters. This approach effectively captured overall data variations, distinguishing areas with varying degrees of degradation, including the most extensively degraded yellowish regions (C2, C3, C4, and C9), less degraded magenta sections (C1 and C5), and less degraded darker areas (C6, C7, C8, and C10). Figure 2.6b, d, f, and h illustrate the spectral centres associated with each cluster.

The primary distinctions among clusters mainly relate to overall spectral intensity, featuring a shared broad absorption peak around 550 nm that gradually decreases in reflectance. For the most extensively degraded group, denoted as C2, the peak is nearly completely lost. In contrast, the least degraded group, C10, exhibits another absorption peak around 650 nm indicating a slight preservation of cyan dyes. However, cluster C9, partially covering the most degraded region, also included a small portion of less degraded pixels with very light dye density, as seen on the upper-left corner of the wall and the newspaper on the desk. Similarly, clusters C1 and C5, characterized by less degradation and middle dye density, exhibit overlap with the highly degraded region, such as areas on the painting and the sofa. Notably, the spectral characteristics of C5 and C9 are highly similar (Fig. 2.6b), with their spectral centres almost parallel to each other

and a difference in reflectance intensity of less than 0.1. This small variation challenges HCA's ability to differentiate, often leading to the mislabeling of pixels at the border of these two groups. Furthermore, it is worth noting that while HCA minimizes differences within each cluster, it lacks a mechanism to maximize variations between clusters. Consequently, cluster sizes tend to become imbalanced, with smaller clusters persisting even in the later stages. As we attempt to increase the number of clusters for finer data partitioning, larger clusters divide, but smaller, less meaningful clusters are often extracted which increases the difficulties in interpreting the results.

The results are improved upon transitioning to soft clustering techniques, specifically Fuzzy C-Means (FCM) and Gaussian Mixture Models (GMM). Figure 2.6c provides a visual representation of the segmentation achieved using FCM with 10 clusters. The regions with varying dye densities are effectively partitioned, as the results obtained through HCA. The less degraded darker areas are covered by clusters C2, C7, C8, and C10, with C8 characterising the least degraded region, in a shape similar to C10 of HCA results. For areas with light dye density, FCM successfully distinguished extensively degraded regions (C1, C3, and C9) from the less degraded surroundings (C4, C5, and C6). FCM exhibits greater precision in separating two distinct degradation states, as evident in the clearer distinction between clusters C3 and C6. However, cluster C1 continues to mix pixels of high dye content, encompassing pixels from both the highly degraded areas, such as the painting in the centre of the wall, and the less degraded surroundings, including the upper-right corner.

This ambiguity is also reflected in the spectra centres (Fig. 2.6d) where FCM tends to form cluster centres that differ primarily in intensity but are uniform in spectral characteristics. Specifically, clusters C1 in both HCA and FCM correspond to the less degraded pixels in the upper-right corner. In HCA, C1 retains a spectral curvature similar to the least degraded group C10, distinct from spectra centres with similar intensity. However, in FCM, C1 is smoothed and averaged to fit into a compromised stage, consequently encompassing pixels with distinct spectral signatures.
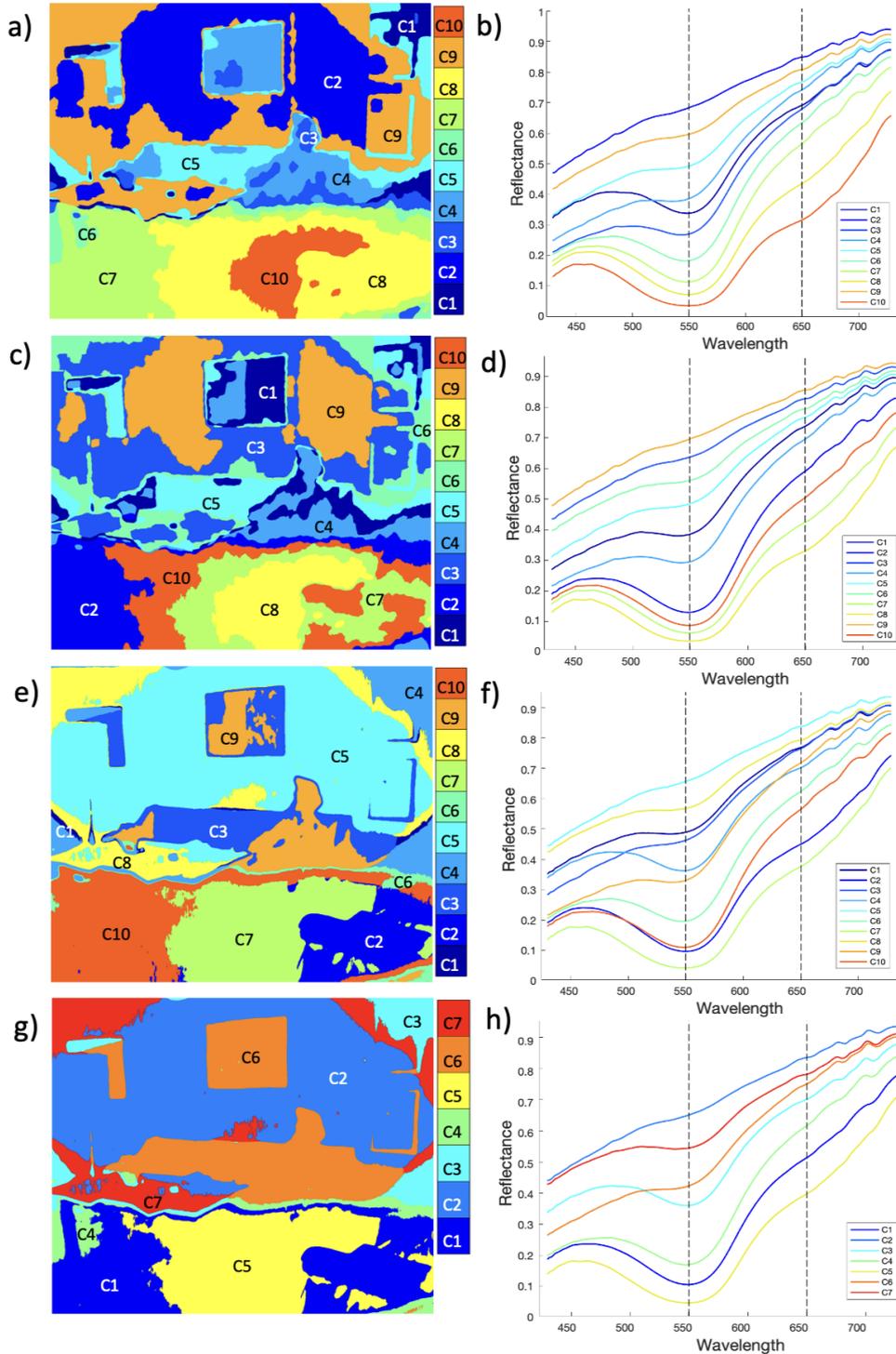
Fig. 2.6 Comparison of the spectral clustering results obtained on film frame 22. First row, the segmented images marked with false colour scale via a) HCA (chebychev metric, complete method, 10 clusters), b) FCM (10 clusters, exponent of 2), c) GMM (10 clusters, regularization = 0.01) and d) GMM (7 clusters, regularization = 0.01) respectively. Second row, the corresponding cluster centres obtained by e) HCA, f) FCM, g) GMM (10 clusters) and h) GMM (7 clusters).

In this context, GMM demonstrated superior performance compared to FCM in efficiently capturing subtle spectral signatures. As opposed to FCM, which relies on distance-based similarity measurements like in hard clustering, GMM assumes that the data patterns within each cluster follow a Gaussian distribution. By computing Gaussian components more sensitive to spectral shapes, GMM succeeds in detecting patterns in ambiguous data, rather than primarily influenced by overall intensities. As depicted in Figure 2.6e, the segmentation achieved using GMM with 10 clusters shows impressive results. The contours of the most degraded yellow stain were accurately captured, with these areas assigned to clusters C5, C3, and C9, while the surrounding magenta regions were appropriately grouped into clusters C1, C4, and C8. Visual examination of the RGB ground truth (Fig. 2.1) confirms that these two degradation stages are well-separated, eliminating the misclassifications observed in HCA and FCM. Specifically, the accurate partition of the newspaper on the table into two degradation groups, C5 and C8, is achieved for the first time, even though they are extremely similar in spectral intensities. Within this partition, C8 represents the less degraded magenta areas, while C5 is within the extensively degraded group, situated primarily on the wall. This clear separation is also observed in the painting hanging on the left of the wall.

The efficient segmentation achieved by GMM is also reflected in the spectral centres, as illustrated in Figure 2.6f. Analyzing the six clusters from the above-mentioned two degradation groups reveals pairs of clusters for comparison: in sequence decreasing in reflectance intensity, C5 and C8 the light dye density pair, C3 and C9 the middle-density pair, and C9 and C4 the higher-density pair. Within each pair, the former cluster, resembling the most degraded yellowish areas, exhibits a less pronounced absorption curvature of around 550 nm, as compared to the latter cluster which has a better preservation state of magenta dyes that absorb mainly around 550 nm. Even when reducing the number of clusters, this fine separation of various degradation degrees was retained, evident in Figure 2.6g and h. GMM efficiently captures the total variances of degradation patterns with a minimum of 7 clusters. Two clusters, C2 and C6, successfully

depicted the yellowish region, while clusters C3 and C7 accurately captured the surrounding magenta areas.

Collective segmentation

Based on its superior performance demonstrated on a single frame, GMM was subsequently applied to perform collective segmentation across the entire dataset. The input data consisted of superpixel centroids concatenated from all four frames (21, 22, 23 and 25), including a broader range of spectral signatures. After the model was trained, it was applied to each individual frame to predict the cluster membership of each pixel. The results of this collective approach are shown in Figure 2.7.
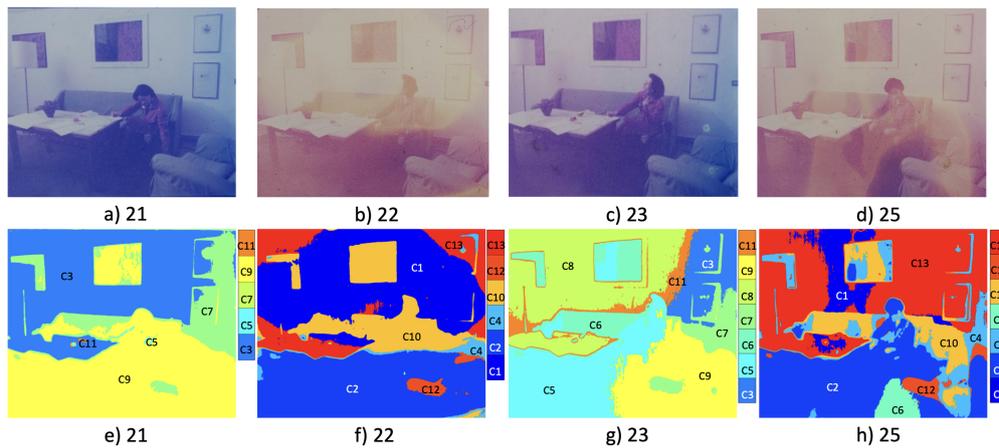
Fig. 2.7 The RGB images of frames 21, 22, 23, and 25 (a-d) and the collective segmentation results marked in false colour using GMM models (e-h).

It was determined that employing a total cluster number of 13 yielded the most efficient segmentation. As a result, the diverse degradation stages present within each frame were effectively separated, while across all frames, areas exhibiting similar degradation signatures were assigned to the same cluster. In more detail, the fine contour of the extensively degraded yellowish region in frame 22 was well captured as achieved in the single frame case. Frame 21,

characterized by minimal degradation, was segmented mainly based on image content into its light whitish part (C3), middle-tone portion (C8), and dark bluish area (C9). Frame 23, presenting a transition from better-preserved bluish colour (on the right) to slightly degraded purplish hues (on the left), was also effectively segmented. Notably, the well-preserved bluish sections in frame 23 shared clusters (C7 and C9) with frame 21, indicating the least degraded state. On the other hand, the slightly degraded purplish sections were labelled with distinct clusters (C8 and C5), reflecting their unique appearance across all frames. In the case of frame 25, which displayed a degradation state comparable to frame 22 but featured different degradation stain patterns, similar clusters with frame 22 were assigned. Both frames 22 and 25 represented cluster C13, designating the less degraded magenta areas, while C1 and C10 corresponded to the extensively degraded yellowish regions. However, the yellowish stain that extended to the darker lower half of frame 25 is not captured, as it merged with a larger cluster C2 of less degraded magenta colour. Besides those clusters shared with frame 22, another cluster C6, corresponding to the darkest pixels in the image, was the same cluster of purplish middle tone on frame 23, indicating a slight degradation stage.

The spectral centres depicted in Figure 2.8 further highlighted the superior ability of GMM to model the spectral variations. Besides modelling the nice spectral response of degradation as in the single frame case, it's also worth noticing that cluster C3, primarily associated with frame 21, was the sole cluster characterized by a higher proportion of cyan dyes than magenta, as indicated by a larger absorption peak at 650 nm than 550 nm. Meanwhile, cluster C8, mainly presented in frame 23, exhibited a pronounced absorption curvature while decreased in the cyan range (650 nm), corresponding to its pinkish appearance with a slight degradation of cyan dyes. Thus, GMM effectively captured both dimensions of data variation, the dye content, and the degradation degree, as represented in the segmentation results and reflected in the defined cluster centres.
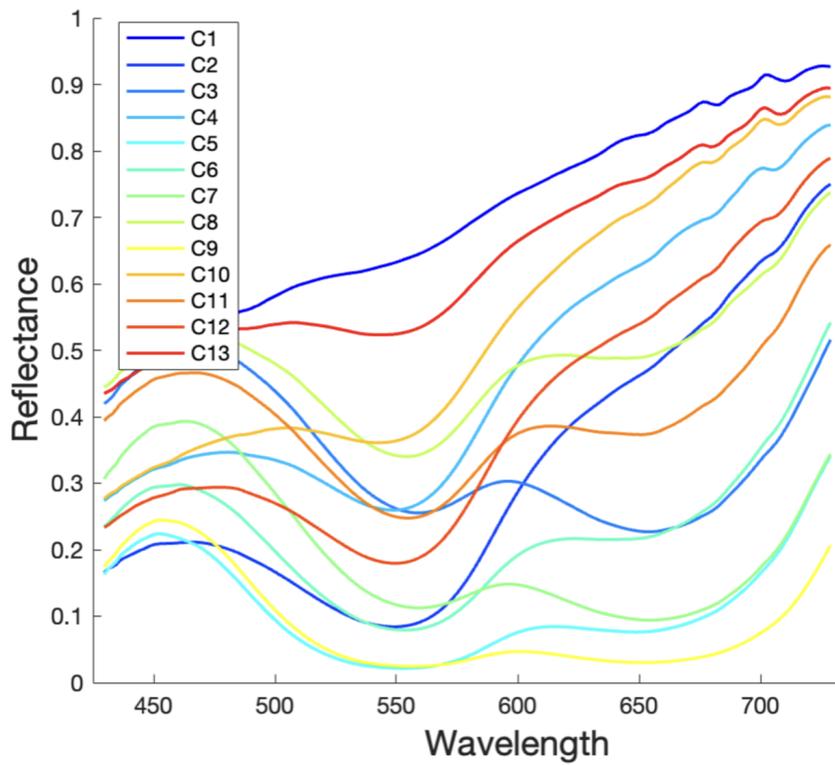
Fig. 2.8 The spectra centres extracted in the collective segmentation of film dataset using GMM models.

In summary, GMM outperformed HCA and FCM at capturing small, distinct spectral variations within the data, making it more effective at identifying and segmenting various degradation areas in films. Furthermore, as the ultimate goal is the digital restoration of the degraded film, it is crucial to achieve a coherent and high-quality image after compensating the diverse degrees of degradation. When employing hard clustering techniques, especially with the grid-based data reduction method, the boundaries between different fading areas are sharp and require additional efforts to smooth these boundaries for a better visual result. This issue is addressed by superpixel spatial segmentation and soft clustering techniques. Given that the algorithms produce a posterior possibility matrix, the corrections applied to each degradation group can be effectively weighted by this matrix. Pixels situated at the border of two groups usually exhibit comparable possibility scores for both groups, resulting in a smooth transition along the border, even without the need for additional filtering.

2.5.2 Segmentation on Neo-impressionist Painting

Spatial clustering results

Like discussed in the previous section, in the case of Signac's painting, 75,000 superpixel was identified as the most efficient input value for k, while enabling the generation of an output matrix size still within our computational capabilities for subsequent steps. The original image, with spatial dimensions of 6,481 x 4,457 pixels (nearly 29 million spectra), was effectively partitioned into 74,697 compact and visually coherent regions (superpixels). These superpixels exhibit an average size of 387 pixels, with the smallest and largest superpixels containing 97 and 1,461 pixels respectively. As evident in Figure 2.9c and d, this number of superpixels facilitated the effective capture of different paint dots, with boundaries conforming to the artistic brush strokes and adjusting to varying sizes.

In further detail, the various shades of blue in the sea juxtaposed with each other (as could be seen in the RGB image Fig. 2.9a) were distinctly separated
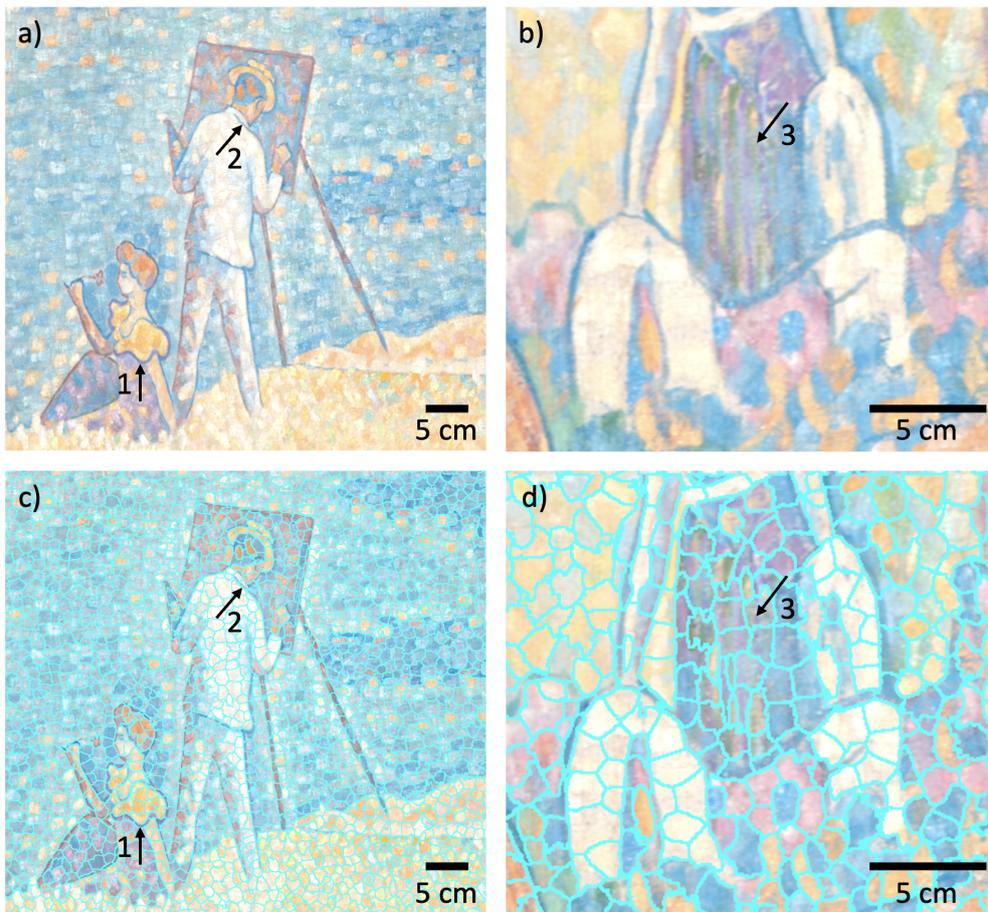
Fig. 2.9 Enlarged areas of the original painting (a and b) and superpixel partition results obtained on Signac's painting (c and d, k=75,000).

into discrete superpixels. Similarly, the irregularly shaped pink, blue, and yellow dots placed next to each other in the lower part of Figure 2.9d were clearly defined. Furthermore, it's noteworthy that the thin outlines are mostly captured as well, exemplified by the female figure in Figure 2.9c (indicated by Arrow 1). However, the algorithm faces limitations in precisely delineating exceptionally thin outlines, such as that on the male figure (indicated by Arrow 2) in Figure 2.9b, which is approximately 3 pixels wide. The superimposing straight lines of blue, yellow, and violet colour in the upper part of figure 2.9b (indicated by Arrow 3) were also beyond the limit of the algorithm to be clearly defined, and compounded by the limited spatial resolution of HSI data.

Spectral Angle Mapper results

The dataset associated with Paul Signac's painting was 6 times larger and more complex than the films studied previously. This dataset exhibited a wider range of variations, including highly diverse spectral signatures and varying intensities, resulting from the different pigment compositions and their numerous mixtures. In this case, the supervised algorithm Spectral Angle Mapper (SAM) was employed as a reference and comparison to the proposed approach. Then, the same methodological approach was systematically tested to classify the VIS-NIR spectra into clusters and generate corresponding mappings, starting from the unsupervised hard clustering techniques HCA (H), to the soft clustering techniques FCM (F), and GMM (G).

The complete pixel-wise identification of the painting using SAM is shown in Figure 2.10. The rough location of the 8 major pure pigments was correctly identified as compared to the original RGB image (Fig. 2.2): two cobalt blue ($CoO.Al_2O_3$), one bright and one dark according to the proportion of white [13], vermilion (HgS) [2], undetermined yellow, viridian green ($Cr_2O_3.2H_2O$)[144], copper and arsenic-based green (Scheele green or emerald green (3 $Cu(AsO_2)_2.Cu(CH_3COO)_2$) [71]) [89, 16], cochineal lacquer [75, 38], undetermined white, cobalt violet ($Mg_{3-x}Co_x(AsO_4)_2 . x H_2O$) [50, 37] and one mixture of vermilion and undetermined yellow. The pure pigments were previously identified using

a database of known pigments and selected as endmembers as introduced in section 2.4.5. Only one pigment mixture (vermilion and yellow) was incorporated as supplementary endmember due to its predominant presence on the canvas. The extracted endmembers are illustrated in Figure 2.11 and the identification associated with each endmember is summarised in Table 2.1.

Fig. 2.10 Complete painting segmentation results generated using SAM (10 clusters).

Although the pixel-wise supervised identification gives a quick and computationally inexpensive result, where each spectrum is compared against the 10 reference endmembers, SAM's results rapidly presented several limitations. First, it is heavily dependent on the choice of representative endmembers, especially when dealing with 29 million spectra. By taking the spectra of the purest pigments, all the minor variations in spectra caused by alteration or pigment mixtures were ignored. Second, given the vast number of spectra, determining a suitable threshold became a delicate balance between accuracy and efficiency.

SAM		
Cluster name	**HSI results**	**Identification**
Endmember 1	Three absorption maxima: 545 nm, 585 nm and 630 nm	Cobalt blue ($CoO.Al_2O_3$)
Endmember 2	Three absorption maxima: 545 nm, 585 and 630 nm	Cobalt blue ($CoO.Al_2O_3$)
Endmember 3	Sigmoid form, maximum of the first derivative: 590 nm	Vermilion (HgS)
Endmember 4	Maximum of reflection at 520 nm, small inflexion at 700 nm	Viridian green ($Cr_2O_3.2H_2O$)
Endmember 5	Two maxima of absorption: 530 nm and 590 nm	Cobalt violet ($Mg_{3-x}Co_x(AsO_4)_2$. x H_2O)
Endmember 6	Two absorption maxima $\log(1/R)$: 522 and 552 nm	Cochineal lake
Endmember 7	Linear reflectance around 1	Undetermined white
Endmember 8	Sigmoid form, maximum of absorption at 450 nm, maximum of reflection reached around 709 nm, maximum of the first derivative: 526 nm	Undetermined yellow
Endmember 9	Maximum of reflection at 525 nm (broad peak), maximum of absorption from 650 nm	Copper and arsenic-based green (Scheele green or emerald green (3 $Cu(AsO_2)$ $_2.Cu(CH_3COO)_2$)
Endmember 10	Inflection at 550 nm, maximum of absorption at 450 nm, maximum of reflection from 709 nm	Vermilion and yellow

Table 2.1 Spectral characteristics and identified pigments for each endmember in SAM.
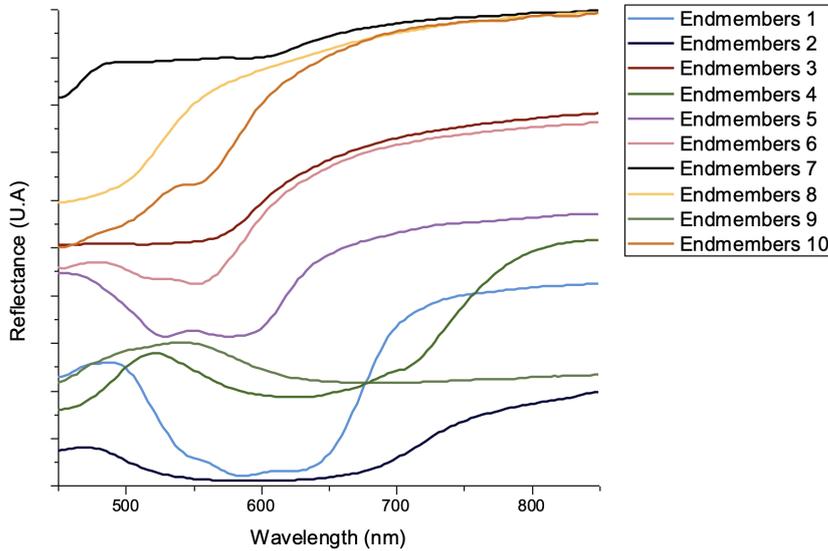
Fig. 2.11 The associated endmembers used in SAM.

A lower threshold value left many pixels unassigned, as marked in black in Figure 2.10, while a higher threshold, aimed at efficiently identifying more pixels, increased the likelihood of misclassification, such as assigning the pigment mixtures not included in the endmembers to the wrong group of single pure pigment. Moreover, the insensitivity to variations in spectral intensities of SAM makes it unable to represent the different shades of blue or green used by Signac to create his contrast. In this context, we found it more promising to reduce the dataset using SLIC and subsequently apply unsupervised clustering algorithms for identifying and mapping a wider range of pigment mixtures and degradations.

Hard clustering results

Unsupervised hard clustering HCA was then applied based on the superpixel centroids. Cosine metric is selected due to its similarity to the SAM algorithm's settings in distinguishing spectral shapes. Figure 2.12 shows the most effective results obtained with a threshold set at 20 clusters, and the centre spectra

of each cluster are illustrated in Figure 2.13. Then, the pigments associated with each cluster were interpreted by comparing the cluster centres to the reference database, and the key spectral features and the identification results are summarised in Table 2.2.

Fig. 2.12 Complete painting segmentation results generated using HCA (20 clusters, Complete linkage method and Cosine metric).

As listed in Table 2.2, this clustering included the same pure pigments, as manually identified previously, except for vermilion which is only found mixed with yellow (C6H). Additionally, the algorithm successfully distinguished a larger number of mixtures that took into account the variation in overall spectral intensity, resulting from an increasing quantity of white pigments [95]. Consequently, 6 different clusters of cobalt blue (C3H, C4H, C7H, C11H, C12H, C18H), each with varying levels of white pigments, were differentiated and mapped (Fig. 2.12). HCA revealed unidentified spectral mixtures characteristic of : i) cochineal lacquer and cobalt blue (C17H - result of superimposition of
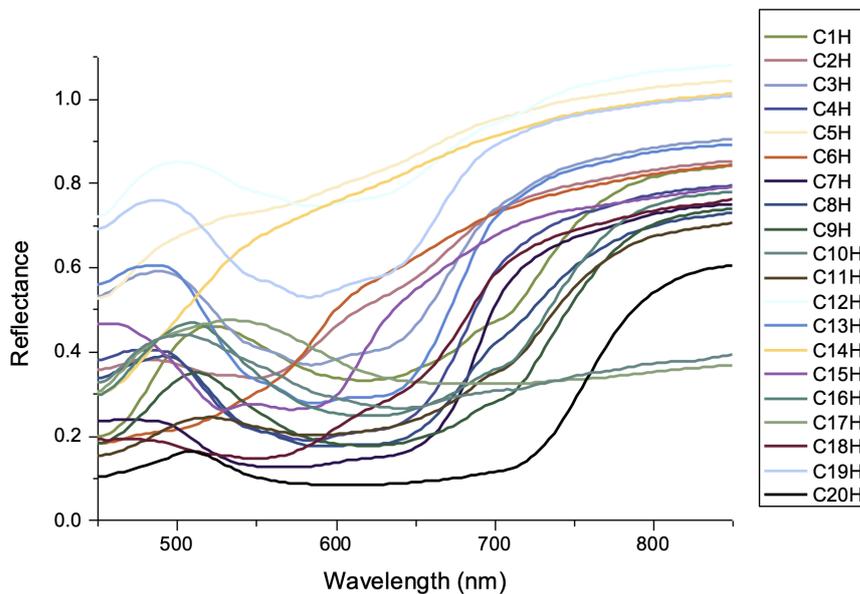
Fig. 2.13 Cluster centres extracted through HCA.

two dots of cochineal lacquer on cobalt blue), ii) viridian green and yellow (C1H, C10H), and iii) mixture of cobalt blue and viridian green (C8H, C19H). The latter combination was employed very occasionally by Signac to paint the darker, almost black dots on the "pétanque" balls, respecting his preferences to avoid all black pigment. In this way, the HCA results offered valuable insights into Signac's palette, particularly with regard to his various mixtures.

However, some recurring problems of this algorithm persist. As previously explained on film samples, the size of certain clusters is not homogeneous. Some clusters are too small and could potentially be combined with another cluster without losing information, i.e. C10H of viridian green is composed of only 98 superpixels, while others are overly large, such as C5H of yellow composed of 14 934 superpixels which includes white dots or paler yellows that should be separated into different clusters.

HCA		
Cluster name	**HSI results**	**Identification**
C3H, C4H, C7H, C12H, C18H	Three maxima at 544 nm, 582 nm and 622 nm	Cobalt blue
C11H	Three maxima at 544 nm, 582 nm and 622 nm. High reflection >0.8	White with some cobalt blue
C8H	Three maxima at 544 nm, 582 nm and 624 nm, small inflexion at 710 nm	Viridian and cobalt
C19H	Maximum of reflection at 510 nm, maximum of absorption from 580 nm to 711 nm, increase of the reflection	Viridian and cobalt
C17H	Maximum of absorption from 517 nm to 572 nm, small inflexion at 625 nm	Cochineal lacquer and cobalt
C6H	Two inflexions at $\sim$ 506 nm and $\sim$565 nm. Maximum of the first derivative at 587 nm, another maximum at 467 nm	Vermilion and yellow
C1H, C10H	Maximum of reflection at 520 nm, small inflexion at 704 nm	Yellow and viridian
C5H	Shape of sigmoid, minimum of reflection at 452 nm and maximum reached around 750 nm, inflexion from 540 to 700 nm, maximum of the first derivative at 471 nm	Yellow
C13H	Shape of sigmoid, minimum of reflection at 452 nm and maximum reached around 750 nm, maximum of the first derivative at 484 nm	Yellow
C14H	Two maxima of absorption at 528 nm and 571 nm	Cobalt violet
C15H, C9H	Maximum of reflection at 510 nm, small inflexion at 704 nm	Viridian green
C2H	Two maxima of absorption at 520 nm and 550 nm	Cochineal lacquer
C16H	Maximum of reflection at 533 nm (broad peak)	Copper and arsenic-based green
C20H	Maximum of reflection at 505 nm (broad peak)	Copper and arsenic-based green

Table 2.2 Spectral characteristics and identified pigments for each cluster in HCA.

Soft clustering results

The soft clustering algorithms, FCM and GMM, were then tested. The most effective results were obtained with 25 clusters as explained in the methodology section 2.4.3. Figure 2.14 and 2.15 present the mapping results corresponding to FCM and GMM clusters respectively. The cluster centres and the identification of each cluster are presented in Figure 2.16 and 2.17 and listed in Table 2.3 and 2.4. Upon general observation, it is evident that the cluster centres produced by FCM exhibit a higher uniformity in spectral features, displaying similar peak positions with varying intensities divided into steps (Fig. 2.16). In contrast, the GMM clusters present a greater diversity in spectral signatures with more intermediate colour mixtures captured (Fig. 2.17).

As a result of careful inspection, FCM and GMM algorithms both distinguished the following pure pigments: cobalt blue, viridian green, cobalt violet, cochineal lacquer, undetermined white, and undetermined yellow like HCA. However, FCM was not successful in distinguishing the Scheele or Emerald green, unlike HCA and GMM algorithms. In the case of more complex pigment mixtures, soft clustering algorithms have provided insights into a wider range of compositions. Five different mixtures of two pigments have been identified inspecting the FCM spectral centres: once more, i) mixtures of vermilion and undetermined yellow (C7F, 18F), ii) mixtures of cochineal lacquer and cobalt blue (C5F), and iii) a mixture of viridian green with cobalt blue (C9F, C19F). GMM further expended the mixtures distinguished, including vermilion and undetermined yellow (C6G, C15G), cobalt blue and undetermined yellow (C8G), as well as mixtures of cochineal lacquer and cobalt blue (C3G, C9G, C17G), and a mixture of viridian green with cobalt blue (C1G, C24G). In total, five different pigment mixtures are distinguished by the FCM, while eleven different mixtures are recognised by the GMM, without taking into account mixtures with white.

An enlarged area of the painting and the main associated clusters are presented in Figure 2.18. The pink dots on the dancer's dress as presented in Figure 2.18 a-c illustrate, however, the difficulties encountered by FCM to distinguish some small variations in the spectrum. GMM demonstrated its ability to effec-

FCM		
Cluster name	**HSI results**	**Identification**
C3F, C6F, C13F, C14F, C16F, C21F, C22F	Three maxima at 544 nm, 582 nm and 622 nm	Cobalt blue
C7F, C18F	Sigmoid shape with two inflection at ~503 nm and ~564 nm. Maximum of reflection of the first derivative at 588 nm, another maximum at 467 nm	Vermilion and yellow
C10F	Maximum of reflection at 514 nm, small inflexion at 704 nm	Viridian green
C9F, C19F	Maximum of reflection at 500 nm, small inflexion at 710 nm	Viridian green and cobalt blue
C2F	Maximum of absorption at 523 nm and 580 nm (broadband), maximum of reflection at 473 nm	Cobalt violet
C12F, C15F	Two maxima of absorption at 520 nm and 555 nm	Cochineal lacquer
C5F	Two maxima of absorption at 520 nm and 570 nm (broadband), small inflexion at ~640 nm	Cobalt blue and cochineal lacquer
C4F, C24F	Quite linear and high reflectance >0.8	Undetermined white
C1F, C8F, C23F	Minimum of reflection at 451 nm, maximum of reflection reached around 740 nm, maximum of reflection of the first derivative: 476 nm and 516 nm (C1G), 521 nm (C8F), 483 and 516 nm (C23F)	Undetermined yellow
C11F, C17F, C20F, C25F	Minimum of reflection at 451 nm, maximum of reflection reached around 740 nm, inflexion around 530-740 nm, maximum of reflection of the first derivative: 480 nm (C11F), 481 nm (C17F), 472 nm (C20F), 469 nm (C25F)	Undetermined yellow

Table 2.3 Spectral characteristics and identified pigments for each cluster in FCM.

GMM		
Cluster name	**HSI results**	**Identification**
C10G, C13G, C22G	Three maxima at 545 nm, 584 nm and 627 nm	Cobalt blue
C6G, C15G	Sigmoid shape with two inflection at ~503 nm and ~562 nm. Maximum of reflection at ~585 nm, two others maximum at ~467 and ~526 nm	Vermilion and yellow
C7G, C11G, C21G	Maximum of reflection ~510 nm, small inflexion at 704 nm	Viridian green
C25G	Two maxima of absorption at 530 nm and 568 nm (broad band) until 590 nm	Cobalt violet
C19G	high reflectance >0.8	Undetermined white
C5G, C16G, C18G	Sigmoid shape, minimum of reflectance at 452 nm and maximum of reflectance reached at 700 nm, maximum of the first derivative: 468 nm (C5G), 473 nm (C16G) and 518 nm (C18G)	Undetermined yellow
C12G, C20G, C23G	Maximum of reflection at 520 nm, minimum of reflection at 606 nm and small inflexion at 705 nm	Undetermined yellow and viridian
C14G	Maximum of reflection at 500 nm, minimum of reflection at 578 nm, small inflexion at 705 nm	Copper and arsenic-based green (+some viridian green)
C1G, C24G	One maximum of reflection at 497 nm, three minima of reflection at ~547 nm, 584 nm and ~631 nm, small inflection at 705 nm	Viridian green and cobalt blue
C2G, C4G	Two maxima of absorption at 520 nm and 555 nm	Cochineal lacquer
C3G, C9G	Two maxima of absorption at ~518 nm and 555 nm, small inflection at ~628 nm	Cochineal lacquer and cobalt blue
C17G	Maximum of absorption at 555 nm (broadband) and inflexion at 627 nm	Cochineal lacquer and cobalt blue
C8G	Two maxima at ~543 nm and ~580 nm, small inflection at 624 nm	Cobalt blue and undetermined yellow

Table 2.4 Spectral characteristics and identified pigments for each cluster in GMM.

Fig. 2.14 Complete painting segmentation results generated using FCM (25 clusters, exponent=2).

tively distinguish this small pigment subgroup, which is correctly represented as pink in Figure 2.18c, while in FCM results these dots were inaccurately represented as yellow (Fig. 2.18b). Cochineal lacquer is characterised by the two minimum reflections at 520 and 555 nm [75], as precisely reproduced by GMM (C2G and C4G) (Fig. 2.18g). On the contrary, FCM failed to distinguish these unique spectral signatures, misclassifying them along with other yellow dots in the grass (C25F) (Fig. 2.18f).

Moreover, as shown by Figure 2.18e, three spectra (C12G, C20G, C23G) of the green clusters generated by GMM stand out from the others. Notably, the spectra of these clusters have an absorption maximum shifted to the highest wavelength at 520 nm, while hydrated chromium oxide (viridian green) alone has a maximum absorption at 510 nm (C7G, C11G, C21G in Figure 2.18e, and C10F in Figure 2.18d) [144]. This type of shift is generally due to variations

Fig. 2.15 Complete painting segmentation results generated using GMM (25 clusters, regularization=0.01).

in the pigment particle size or to the presence of another material [83],[116]. The spectra obtained by GMM corresponded in our case to a mixture of viridian green and yellow at 520 nm. These results particularly highlight the contrast of hues, dear to Signac to represent the vibration of his grass.

However, regarding the pure pigment mixed with white, the largest variety of blue tones are identified with FCM, i.e. 7 with FCM (C3F, C6F, C13F, C14F, C16F, C21F, C22F) against 3 with GMM (C10G, C13G, C22G) (Table 2.3 and 2.4). The different proportions of white mixed in those clusters mainly lead to changes in spectral intensities, without altering the absorption positions [95]. On this point, FCM is more sensitive in capturing the variations in intensities, leading to a better illustration of the contrast of tone used by Signac to represent ripples in the sea (Figure 2.14).
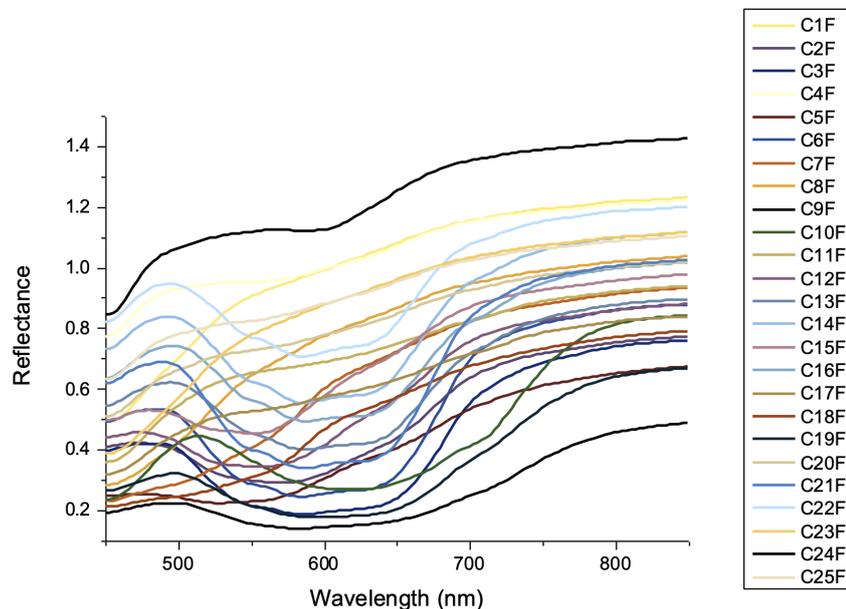
Fig. 2.16 Cluster centres extracted through FCM.

The same results were observed for the yellow dots. FCM identified 7 different shades of yellow (C1F, C8F, C11F, C17F, C20F, C23F, C25F) while GMM obtained only 3 clusters (C5G, C16G, C18G). As marked with a false colour scale in Figure 2.19a (FCM) and Figure 2.19d (GMM), the location of each cluster is clearly mapped. This finer partition of FCM provides a more refined representation of the yellow tones close to the real painting, as depicted with the converted colours from spectral centres in Figure 2.19b, compared to the GMM clusters with less gradient (Figure 2.19e). The spectral centres related to each cluster are plotted in Figure 2.19c and f. Some spectra of FCM showed a decrease in intensity (C1F, C8F, C23F). The diminution of reflectance between each cluster is around 0.1 and can be due to a decreasing quantity of white. The C18G of GMM algorithm has the same sigmoid shape and quite the same intensity as the cluster C23F of FCM algorithm (0.02 of difference). Some other clusters in FCM algorithm showed the same decrease of intensity as well as an inflection in the 530 - 740 nm range (C25F, C20F, C11F, C17F).
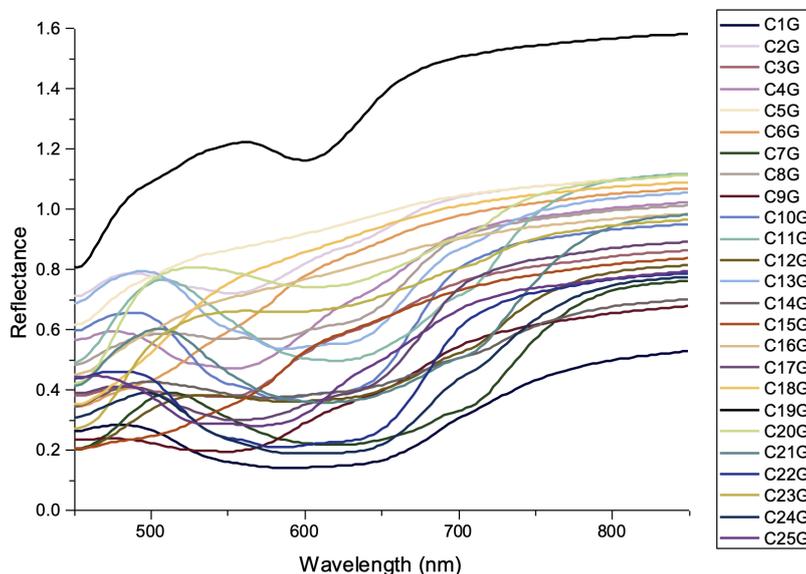
Fig. 2.17 Cluster centres extracted through GMM.

The same inflexion and shape of the sigmoid are observable on the two spectra of GMM: C5G and C16G. This inflexion, like the drop-in intensity, could be a clear sign of the deterioration of certain dots in the canvas. This would explain the general discolouration of the yellow grass. While direct exposure to daylight may have accelerated the process, it's worth noting that Signac himself was already aware of the issue: "the bright cadmium in two days turned into a dirty greenish brown" Paul Signac, Journal, 14 July 1894 [183]. However, the inflexion visible in the 530-740 nm range may suggest a sign of degradation of chromium-based pigments rather than cadmium. Indeed, a similar inflexion appears during the aging of $PbCrO_4$ and $PbCr_{0.2}S_{0.8}O_4$ in linseed oil in [141], whereas this inflexion is not observed nor correlated to the degradation of CdS, cadmium yellow pigment in linseed oil [140]. Signac may have changed the nature of his yellow pigments in response to the problem of colour changes during retouching.
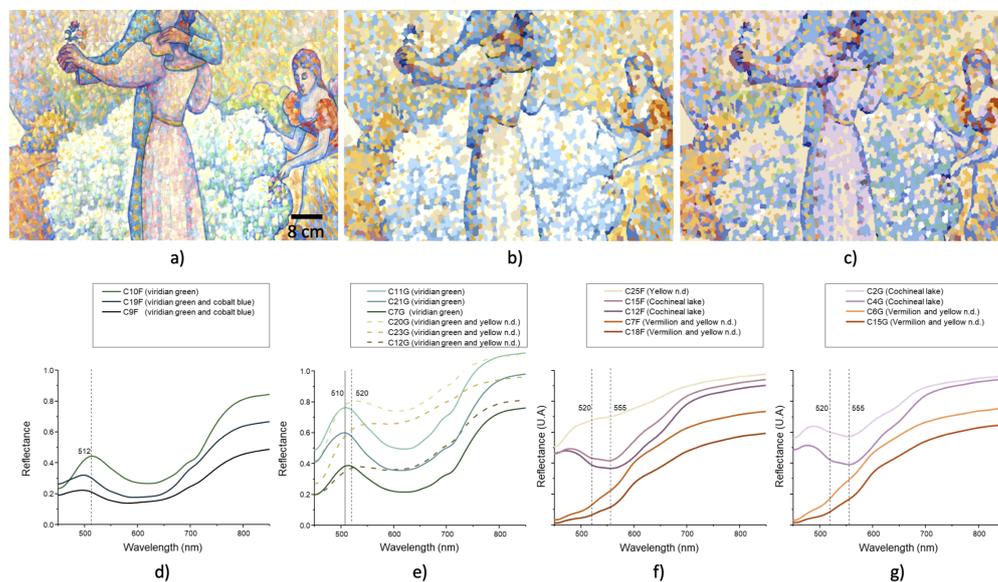
Fig. 2.18 Detail area of the original image (a) and the map obtained by b) FCM (25 clusters), and c) GMM (25 clusters), supported by some spectra of green dots identified by d) FCM and e) GMM, and some others characteristic spectra of f) FCM and g) GMM.

In summary, the results presented above highlight the distinct strengths of FCM and GMM in handling spectral data. FCM excels in its sensitivity to variations in overall intensities, effectively partitioning pure pigment groups based on a range of hues. Conversely, GMM demonstrates its superior precision in capturing subtle spectral changes, enabling the accurate differentiation of diverse materials and their combinations. These differences between FCM and GMM may be attributed to their initialisation strategies [213]. For Signac's dataset, the size of different pigment groups is unevenly distributed. For instance, there are approximately 24% of cobalt blue dots and 16% of undetermined yellow dots on the canvas, while only 2.5% of cochineal lacquer and 0.1% of Scheele or Emerald green. Consequently, since FCM takes initial cluster centres randomly, larger pigment groups associated with more data points are more likely to be over-represented by assigning more cluster centres relating to it, potentially leading to the neglect and merging of smaller groups into larger clusters. On the
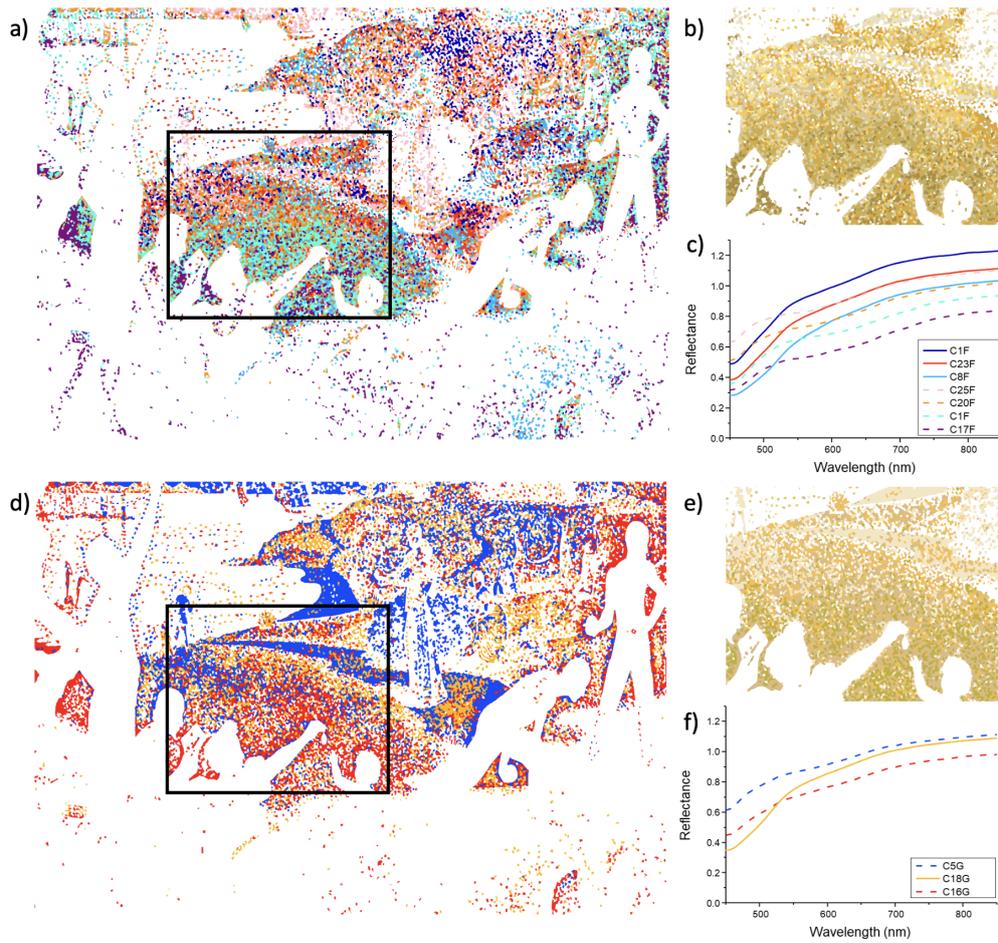
Fig. 2.19 Maps of yellow clusters marked in false colours, real RGB colours, and their associated spectra obtained by FCM (a-c) and GMM (d-f).

contrary, GMM initialises the cluster centres based on a k-means algorithm [10], which initially filtered all data points based on spectral features.

Furthermore, the distance measurement metrics and objective functions employed in the two algorithms could have significantly contributed to the observed differences in results. On one hand, the classical FCM algorithm employed relies on the Euclidean metric as its distance measurement function in the cluster updating stage. This metric naturally tends to yield intensity-based partition and favours the formation of spherical clusters, where data points are distributed more uniformly around a cluster centre. On the other hand, the GMM algorithm models the data points as a combination of multiple Gaussian distribution components with specific mixing coefficients as their weights and optimises the log-likelihood of data points to a component centre. This inherent flexibility allows GMM to form nonspherical and unbalanced clusters that can accurately capture more complex underlying data patterns.

As a result, both the FCM and GMM clusters exhibit an average size of 2,988 spectra, while significantly different in the standard deviations (1,009 for FCM and 2,857 for GMM). GMM produced the largest cluster comprising 10,087 spectra (C22G), associated with cobalt blue along with two other substantial clusters each containing over 7,000 spectra (C10G and C13G). In contrast, FCM's largest cluster contains only 4,800 spectra, also in the blue category, accompanied by six moderately-sized clusters (Table 2.3), each averaging around 3,000 spectra. Conversely, GMM generated very small clusters, including the tiniest one with only 96 spectra dedicated to pure viridian green (C21F), and several other clusters of hundreds of spectra, such as a 137 spectra cluster representing Scheele or Emerald green, a 163 spectra cluster for unmixed white, and a 399 spectra cluster denoting cochineal lacquer. On the other hand, FCM yields just one cluster with fewer than a thousand spectra, C24F, with 686 spectra presenting mixed white. Scheele or Emerald green and cochineal lake are integrated into the larger groups.

Thus, in essence, FCM excels in depicting variations within substantial material categories, while GMM is adept at capturing unbalanced or non-uniform

data patterns. These distinct advantages of soft clustering techniques have not only contributed to a more comprehensive understanding of Signac's palette but also paved the way for future re-colourisation attempts, enabled by the precise localisation of degraded pigments.

2.6   Conclusions and Perspectives

In this chapter, we explored the application of machine learning algorithms to address complex identification and classification problems. We have developed an innovative data processing pipeline combining both spatial and spectral clustering techniques to extract and map spectral signatures, for diverse types of data from both film samples and Paul Signac's painting. The SLIC superpixel algorithm has effectively reduced the data size by grouping similar and repetitive spectra from the original dataset while preserving full spectral signatures and important spatial features. Incorporating spatial information into classification processes is proven to greatly enhance processing efficiency within extensive datasets, as compared with the conventional methods.

Furthermore, the unsupervised clustering techniques have demonstrated their superior ability to reveal underlying patterns of the data. Soft clustering algorithms, specifically Fuzzy C-Means (FCM) and Gaussian Mixture Models (GMM), excelled at capturing subtle spectral variations from overlapped data and offered a more nuanced and detailed perspective of the art materials studied. They succeed in distinguishing various inorganic pigments presented in the painting, identifying complex colourant mixtures in both cases, as well as localising the degradation. These findings have not only shown the potential of machine learning techniques in aiding artwork diagnostics and conservation, but also provided an invaluable foundation for future restoration works. These unsupervised techniques have proven invaluable in uncovering nuanced details that humans cannot determine from the vast amount of data. However, while unsupervised algorithms excel in finding intricate patterns and trends, the human interpretation of the extracted spectral signatures remains indispensable. The

synergy between automated techniques and human expertise is crucial for the holistic and meaningful conservation process.

For the applicability of the model, the proposed process has been tested and proven effective in two entirely diverse cases, indicating its adaptability to different scenarios. In particular, the Gaussian Mixture models constructed using the current training samples are directly transferable to the analysis of artworks with similar palettes or material compositions. Through the expansion of the included spectral dataset, these models can be applied widely and efficiently to numerous similar scenarios. Future research efforts will also explore advanced parameter settings, such as Mahalanobis distance-based, as introduced by Gustafson and Kessel in 1979 [85], which instead assumes non-spherical cluster shapes and may have improved performance on our dataset, given its flexibility in handling clusters with diverse shapes. It can be set on FCM or GMM with different covariance matrices, to further adapt these techniques to diverse data analysis needs.

Chapter 3

Digital Restoration of Colour Films via Vector Quantization Algorithm

> To what extent can machine learning models detect and mitigate the degradation of cultural heritage artefacts over time?
> How transferrable are the machine learning models developed to similar scenarios in cultural heritage conservation and what future applications can be envisioned?
>
> **— RQ-1.2, RQ-1.3**

This chapter proposes a novel vector quantization (VQ) algorithm for restoring movie frames based on the acquisition of VNIR hyperspectral data. The VQ algorithm utilizes what we call a multi-codebook that correlates degraded areas with corresponding non-degraded ones selected from reference frames. The spectral-codebook was compared with professional commercially available film restoration software tested both on RGB and on hyperspectral providing better results in terms of colour reconstruction.

3.1   Introduction

Digital restoration is a rapidly growing methodology in cultural heritage whereby images of art objects are computationally manipulated to visualize their original appearance or reveal hidden information without actual physical intervention [149, 17, 106, 90]. Digital restoration is increasingly playing a role in interpreting and displaying an artwork when it is severely damaged [158, 210] or when it has been stripped of historically significant information [35].

As has been recognized by UNESCO since 1980, moving images are a fundamental part of the world's Cultural Heritage [197]. Throughout the 20th century, films were coloured with light and heat-sensitive dyes incorporated into the emulsion layers. Today, these films often exhibit colour degradation, fading, colour loss, bleaching, and colour change [197], thus necessitating their digital restoration [126, 40, 41]. For motion pictures, the film is commonly restored by scanning using an RGB scanner and manually processed with dedicated software, such as *Photoworks Photo Editor 2021* [151], *DaVinci Resolve 17* by Black Magic [24], and *Paintshop Pro* by Corel [51], to re-balance the colour and adjust the colour saturation and contrast [40]. Conventional digital restoration is laborious, with the resulting appearance reliant upon the restorers' skills and judgments about what looks appropriate.

In response to these subjective approaches, several algorithms have been developed to automatically restore digitised films with minimal intervention [126, 187]. Several of these techniques successfully detect scratches or lacunae, and this missing content is in-painted using standard techniques [90, 52, 166, 112]. However, for faded colour, most existing models assume a homogeneous reduction in colour and hue across the image frame. Only one deep learning algorithm, based on latent space translation, trains with paired synthetic data [206] to compensate for uniform fading. For more severe and inhomogeneous colour loss, the algorithm is prone to failure. Other CNN algorithms focusing on the colourization of black and white films [29, 109, 45, 107] rely on synthetic training data sets that have the same limitation when it comes to uneven fading. In historical films, the degradation of colour usually varies across and within each

frame, so restoration models trained using many homogeneous synthetic images may impose inaccuracies or even colour distortions. Another approach, based on what is known as the *Automatic Colour Equalization Model* [41], imitates the mechanics of the human visual system, optimizing colour contrast, saturation, and balance according to human perception and aesthetics rather than restoring the film to its original appearance. Such methods are generally ineffective when attempting to restore artworks to a historically accurate state, as is the central requirement in the cultural heritage sector.

Practical restoration of differentially degraded colour film thus remains an unsettled problem. Here we propose advanced tools, such as spectral imaging, to face the challenges imposed by the complexity of colour degradation in historical films. Hyperspectral imaging has been increasingly applied to the analysis and conservation of important artefacts [115, 74, 6]. The fine spectral resolution afforded by optical reflection spectroscopy, down to nanometre resolution, enables the capture of degradation phenomena of film at high spatial and spectral resolution, which is otherwise hard to identify with the conventional RGB captures. By combining spectral imaging with advanced machine learning algorithms, the limitations of using synthetic data alone are overcome, given the large amount of spectral data that may serve as the input to the algorithm. In addition, machine learning also handles the challenge of processing large amounts of data which is often a major concern in cultural heritage applications. Such methodology has already been reported in the study of illuminated manuscripts where hyperspectral imaging and a deep neural network were combined to perform the spectral unmixing and quantitative estimation of pigment concentrations [163]. Another important work on the degraded medieval manuscript [154] proposed a codebook algorithm to fuse the hyperspectral data and XRF data that successfully revealed the hidden content through the correlated spectral mapping. Although no application of this approach has been reported on film restoration, those research projects open the door for a novel solution to the colour degradation problem in damaged historical films.

This study proposes a machine learning algorithm that avoids subjective choices in restoring differentially faded film. As described in more detail below, a vector quantization algorithm is proposed that exploits a sparse representation of spectral reflectance data obtained from degraded and non-degraded films. After registration of representative degraded and non-degraded frames, a joint dictionary is learned from these data sets, which calculates a restored representation for the entire film. Spectral data were first processed using a simple codebook approach and further improved by a multi-codebook method capable of restoring frames with different degradation effects. The method proposed here provides more accurate results than those obtained with the currently available restoration software.

The remainder of the chapter is structured as follows. Section 3.2 illustrates the research questions of this study. Then, we describe the dataset used and the methodology of VQ algorithm in Section 3.3. In Section 3.4 we present and discuss the results and comparative analysis. Finally, Section 3.5 concludes this chapter and proposes possible future research perspectives.

3.2   Research Questions

Starting from the background analyzed in the previous section, in this chapter our goal is to respond to the following research questions:

RQ-1.2 *To what extent can machine learning models detect and mitigate the degradation of cultural heritage artefacts over time?*

RQ-1.3 *How transferrable are the machine learning models developed to similar scenarios in cultural heritage conservation and what future applications can be envisioned?*

3.3   Materials and Methods

3.3.1   Materials

Six miscellaneous positive prints (Fig. 3.1) were provided by L'immagine Ritrovata Film Restoration Laboratory in Bologna, taken from a historical film fragment. They belong to the same scene but report different fading effects. Each frame is measured to have a physical dimension of 34 mm x 19 mm (± 0.5 mm). The FUJI N4 edge mark is found on the perforation, identifying the manufacturer as FUJI film. In addition, "N4" indicates the support material to be cellulose acetate, and the code "86-AJ" was used to indicate the manufacture time which corresponds to April-June 1986. The Fuji film set has a thickness of around 130 μm, excluding the protective tapes attached to the surface. The emulsion layer has a trichromatic structure with a thickness of around 15μm, in dye sequences of yellow, magenta, and cyan. The dyes located in different layers are degraded at separate rates through time [67].

Fig. 3.1 Optical RGB images captured by Canon EOS 5D Mark IV camera of the film samples (S1-S6) considered in this chapter.

The best-preserved film S1 is regarded as "the good" reference to form the codebook. From the degraded frames, four different fading effects are recognizable, as indicated in Figure 1 (Type 1-4):

1. A yellowish hue is formed due to the decomposition of cyan and magenta dyes (yellowish upper part in S2).

2. A pinkish hue in S5 and S6 is due to the degradation of the cyan dye.

3. Purplish hue on the left part of S3 and the upper part of S4, probably due to a very mild degradation that leaves an amount of the cyan dye and keeps most of the colour density.

4. Bluish strips on the right of S3 and S6, as well as sample S1, are considered un-faded parts that preserve most of the dyes.

Sample S6, containing pixels presenting all four described fading effects, provides the richest information on the fading and is selected as the fade reference.

3.3.2   Data Acquisition Methods

The frames were scanned with a custom-made VNIR push-broom hyperspectral camera [65], specifically built to have high spatial and spectral resolutions. Spectral images were acquired in a reflectance geometry with a broadband LED light source covering the spectral range from 380 nm 1000 nm [65]. The light source was configured to illuminate the whole sample homogeneously from two sides while using a combined diffusor/polarizer (Bolder Vision Optik, Inc., USA) in front of the LEDs. The samples were positioned on a white Spectralon (Labsphere, USA) and affixed to the white surface using a custom-made frame, with the dyed layers oriented towards the camera. Throughout this case study, the quantity reported and discussed is, for simplicity, the reflectance. It is important to note that this is the reflectance of the film and the white imaging substrate. In terms of the film, it is composed of the true film reflectance and the square of the film transmittance, since light first passes through the film, is subsequently reflected from the substrate, and finally passes the film again before being detected. This process amplifies spectral features, which is desirable for the problem at hand. The simplification is justified in the scope of this case

study since only the different spectral shapes of degraded and well-preserved dyes are of interest. The imaging part of the system included an ImSpector V10E imaging spectrograph (Specim, Spectral Imaging Ltd, Finland), a 50 mm lens (Schneider Kreuznach Xenoplan 2.8/50-0902, Jos. Schneider Optische Werke GmbH, Germany) and 5.0MP monochrome CMOS camera (Blackfly S BFS-U3-51S5M-C, Flir Systems Inc., USA). To mitigate specular reflections, a polarizer (Bolder Vision Optik, Inc., USA) was used in front of the objective in a cross-polarized configuration with the LED polarizers. Images were acquired with the resolution of 2,048 × 2,448 pixels in spectral and spatial dimensions of the spectrograph, respectively. The spectral range considered is from 380 nm to 780 nm. The field of view in the direction perpendicular to the scanning axis was 73 mm. The system's effective spectral and spatial resolutions were 2.9 nm and 100 μm, respectively, as evaluated by a gas discharge tube and spatial grids used for system calibration. In comparison, RGB images were acquired with a Canon EOS 5D Mark IV camera.

3.3.3   Vector Quantization Algorithm

To gain the information needed to restore the degraded film, the proposed algorithm relies on two spectral reference images for training: the best preserved of an individual scene and a representative faded frame of the same scene. First, we map the two frames into a space so they may be compared. To do this the preserved B and degraded frames F are spatially registered pixel-by-pixel. This is conceptually valid as the physical materials in both frames should be made of the same material classes within a given spatial distribution. The only difference is that the degraded frame B has a slightly altered chemistry compared to the degraded frame F. Therefore, if the spectra of the preserved frame B can be successfully clustered, such that the clusters represent the concentration of the photographic dyes, then these clusters should be expected to correlate with the degraded frame F intensities. At the core of our algorithm, we are finding the reflectance spectra of the preserved frame that are best related to the reflectance spectra of the degraded frame.

To find these mappings, the degraded frame F is clustered into K groups to find pixels composed of the most similar spectra. We use a vector quantization method analogous to K-means clustering- except without an update step. Once the initial clustering of F has been performed, we can predict what response for the preserved frame B. The mean of F per cluster should be a strong estimate for the composition of the pixels belonging to those clusters. Based on this, one can estimate an image of a "restored" frame by replacing each cluster of pixels in F with the mean of B response. Thus, the restored frame is estimated from the degraded frame as

$$\hat{X}_{\text{restored}} = \sum_k \varepsilon_{C_k} [X_F] I_{C_k} \tag{3.1}$$

where k is the cluster index and C_k is the k_{th} cluster, and I_{C_k} is the indicator function of cluster C_k. The whole process is schematized in Figure 3.2 and summarized step-by-step here:

Step 1: To map the correlations between the degraded and non-degraded states of the film, two references need to be selected: as previously stated, 1) a relatively best-preserved reference frame B that serves as the source of "good" spectral signatures and 2) another faded frame F of the same scene that provides the degraded spectra. The paired references are then registered pixel-wise using Scale-invariant feature transform (SIFT) correspondence and landmark transformation [169]. Through a pixel-to-pixel correlation, each pixel spectrum on the faded reference F_i has a well-preserved correlated spectrum $B_i = a_i F_i, i = 1, \ldots, N$ where a_i is the transformation coefficient of each pair of spectra. From this point, it is useful to think of these correlated images as a joint codebook C that contains paired information between the unfaded frame and the faded one. Each spectrum thus serves as a codeword, and the paired references form the codebook C of N paired codewords:

$$C = \{(B_i, F_i) \mid i = 1, \ldots, N\}. \tag{3.2}$$
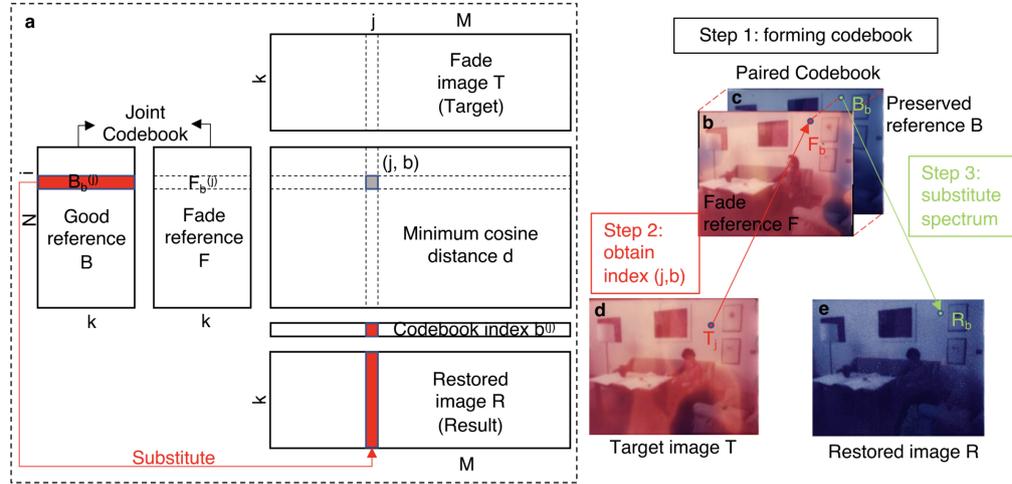
Fig. 3.2 Schematic overview of the vector quantization algorithm. a) k is the spectral wavelength number from the flattened image, N and M represent the total pixel number respectively in the reference image and target image, and (j,b) represents the index of the best representative codeword. b-e) RGB representation of fade reference (b), good reference (c), target frame (d) to be restored, and the restoration result (e).

Since the faded frames used in this study are not artificially simulated but actual historical samples representing degradation behaviours, there are subtle content changes from frame-to-frame, such as human figures with slightly adjusted postures. These shifts cause small non-matched pixel regions around the head area that are masked during the following calculation to avoid mismatches.

Step 2: Looking up the codebook and finding the index. To "translate" a degraded spectrum in the target frame T into a best-preserved spectrum, we need to look it up in the codebook first. Therefore, the spectrum T_j at each pixel position in the target needs to be compared with every element throughout the codebook C to locate the most representative spectrum F_b. The codebook index b is learned by calculating the minimum cosine distance d between the target spectrum T_j and the reference faded spectrum F_i. During this stage, the spectral data cube is unfolded spatially into a flattened image with dimension Nxk, where

N is the total number of pixels and k is the wavelength channels. When using RGB images as input data $k = 3$, and for spectral data cubes $k = 240$. After restoration, the unfolded matrix is reformed into the original spatial structure. For each spectrum, T_j in the target cube T (Fig. 3.2d), the cosine distance d to every spectrum F_i in faded reference F (Fig. 3.2b) is obtained by

$$d\left(T_j, F_i\right) = \frac{F_i \cdot T_j}{\|F_i\| \, \|T_j\|} \tag{3.3}$$

The resulted distance matrix d has the dimension of NxM, where N is the total pixel numbers in reference and M is that in the target. The two images do not necessarily have to be identical in total pixel number and resolution. For the j_{th} spectrum in the target, the codeword $F_b^{(j)}$ that best represents T_j is identified by

$$d\left(T_j, F_b^{(j)}\right) = \min_i d\left(T_j, F_i\right) \tag{3.4}$$

which indicates the best match between the target spectrum and the reference spectrum.

Step 3: Reconstructing the restored image. A target-to-codebook relationship is established via the codebook index $b = \left\{b_1, b_2, \ldots, b_j, \ldots, b_M\right\}$ formed of all elements $F_b^{(j)}$. Then, exploiting the pixel-to-pixel correlation in the codebook, each spectrum T_j in the target image can be substituted by the corresponding good spectrum $B_b^{(j)}$ of the best representation $F_b^{(j)}$ in the codebook. A restored cube R (Fig. 3.2e) is formed where

$$R_j = B_b^{(j)} = a_i \cdot \mathrm{argmin}_i \, d\left(T_j, F_i\right). \tag{3.5}$$

The faded image is reconstructed with the "good" spectral signatures, collected from the good reference, following the calculated index.

VQ has already been applied to many recognition problems associated with language [32]. As an analogy, the principle of this digital restoration strategy is like the process of translating a foreign language. Each faded pixel in a film frame would be an element to be translated into an unfaded pixel. Once

the codebook is built, it may be applied to restore any frame with the same degradation characteristics, not limited by size or resolution.

To overcome the limitations of the VQ methods, a multi-codebook was also created to improve the algorithm's accuracy using spectra hand selected from multiple frames. The multi-codebook is one in which the atoms are hand-selected and concatenated from multiple frames.

3.3.4   Comparative Analysis

Different elaborations were undertaken to compare the application of the novel algorithm on the hyperspectral data. First, RGB-acquired images were digitally restored with the professional software, including *Photoworks Photo Editor 2021* [151], *Paintshop Pro* [51], *SoftOrbits Photo Retoucher*[185], and *DaVinci Resolve 17* [24], following the instructions available on their website. The RGB data were also tested on our developed algorithm following the same pipelines employed for processing the spectral data (Fig. 3.2).

3.4   Results and Discussions

3.4.1   Comparative Analysis: digital restoration of RGB data

Conventional approach

RGB images were acquired with a conventional camera and processed with professional restoration software to evaluate the algorithm's efficiency and reconstruction quality.

For the standard software, the colour correction relies on hand adjusting the parameters until the best possible appearance is obtained according to the operator's skills and aesthetic judgments. Moreover, it is hard to compensate for differential degradation across the frame and achieve a uniform result. As a comparison to our approach, we tested several commercially available restoration software to digitally restore the colour of our faded films. A first attempt was

made using *Photoworks Photo Editor 2021* (Fig. 3.3). Their retouch tool could effectively remove structured defects, such as dirt and cracks, and restore the image to a clean state as done by chemical cleaning. The colour restoration, however, depends on the hand adjustment, which is time-consuming and highly subjective. It is hard to separately treat the more degraded area and remove those stains, bringing the entire image to a uniform appearance. For unskilled non-professions like us, the best we can achieve still presents a large colour difference from the reference.

Fig. 3.3 The RGB image before and after the retouch using *Photoworks* software.

Other more automatic restoration software were also tested (Fig.3.4). The automatic fade correction enabled by *Corel Paintshop Pro* [51], noted to be based on artificial intelligence technology, considers the fading to be homogeneous on the image not only spatially, but also among different types of dyes. The enhanced image is even more unbalanced, and the lost information of the already degraded dyes could not be recovered through this type of practice. On the other hand, the *SoftOrbits Photo Retoucher*[185], designed specifically for restoring old photos, is equipped with an automatic re-colorization function. However,

this program is developed based on black-and-white photos. The restoration result appears yellowish and lost completely the unique colour characteristics.

Fig. 3.4 The RGB image before and after the automatic fade correction using *Corel Paintshop Pro* (left) and re-colourised by *SoftOrbits Photo Retoucher* (right).

Then, the DaVinci Resolve 17 software [24] is also tested, as it is one of the commonly used commercial software by film restorers. In Figure 3.5, the general restoration procedures are demonstrated.

As an example, Figure 3.5a reports the restoration result obtained on frame S2 in comparison to the original state. As shown in Figure 3.5b, the primary colour balance was adjusted on the base of selected points on the background (wall areas). Then, two more correction nodes were added to fine-tune the contrast, saturation, hue, tint, temperature, and RGB curves to restore the overall appearance. After those steps, the darker lower half achieved a comparable visual effect to reference S1. At the same time, the more extensively degraded upper part still had a pronounced yellowish hue, which was finally enhanced by selecting those areas and separately adjusting the parameters. The colour appearance of the result (Fig. 3.5a) appears non-uniform, presenting a colour difference from the reference. It can be confirmed with the histogram graphs, taking the blue channel as an example, shown in Figure 3.5c, that the distribution of the colours in the restored image is shifted mainly from the original S2 and now matching with the reference S1.
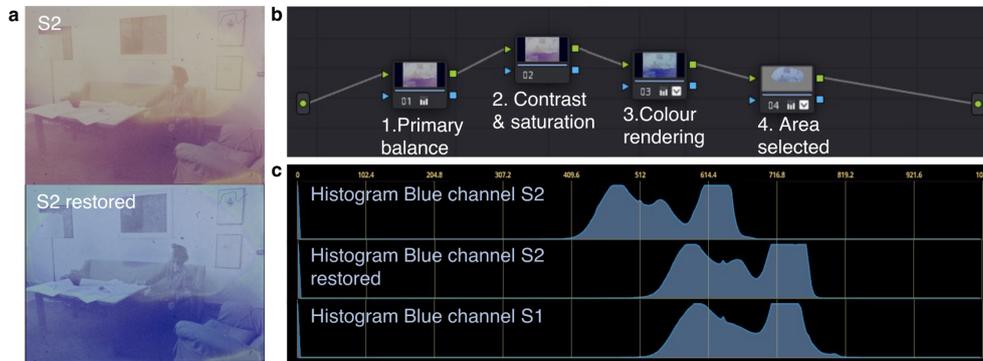
Fig. 3.5 The conventional digital restoration strategy. a) Comparison of RGB images before and after the hand restoration achieved by DaVinci Resolve 17. b) Illustration of the processing pipelines. c) Comparison of the histograms before and after restoration with the reference S1.

The overall balance of the RGB colours is also adjusted closer to the reference, but the restored image still presents inconsistencies that are hard to correct completely. The same procedure is repeated also on frames S3 and S5, and the results are summarised in Figure 3.6d. The problems relating to restoring a single frame persist, while it can also be noticed that a uniform appearance among all three frames is also hard to achieve. Thus, the results achieved by conventional practice are limited by the restorer's subjective choice, personal taste, and proficiency. Moreover, the processing of one single frame takes up to 30 minutes and for historical films with inhomogeneous degradation, the fine-tuning of parameters is inevitable from frame to frame.

VQ algorithm on RGB data

The VQ algorithm was applied to RGB data to evaluate the advantages of processing with hyperspectral data. In particular, a codebook was created using S1 and S6 as paired references (Fig. 3.6a,b). The wavelength channel has a dimension $k = 3$. Then the digital restoration was achieved on target images S2, S3, and S5 by the vector quantization method proposed above, finding the best

a Fade reference
c Original RGB images of the targets
b Good reference
d DaVinci software restoration results
e RGB codebook restoration results
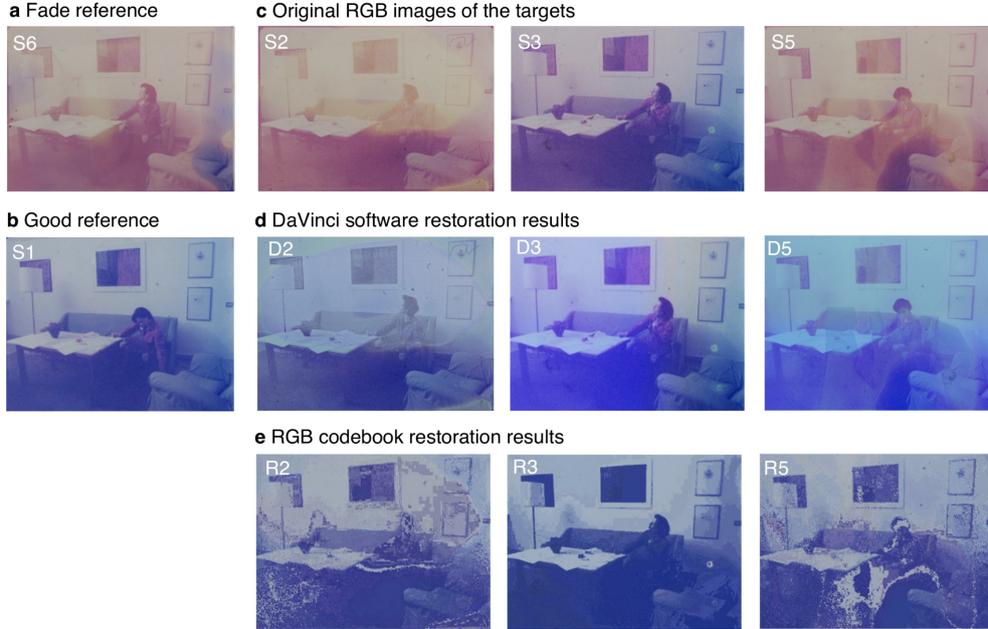
Fig. 3.6 Originals and restoration results obtained via DaVinci Resolve software and RGB-codebook approach. a-b) Optical RGB images of references S6 and S1. c) Optical RGB images of target frames S2, S3, and S5. d) Manually restored frames D2, D3, and D5 using DaVinci Resolve software. e) Restoration results R2, R3, and R5 obtained via RGB codebook approach.

representative codewords in the fade reference S6 and then substituting each pixel with the corresponding good one. The results are reported in Figure 3.6e. A simple evaluation of the restoration performance is based on calculating the colour difference ΔE between the obtained results *R* and reference *S1* per pixel through Euclidean distance:

$$\Delta E = \sqrt{(r_R - r_{S1})^2 + (g_R - g_{S1})^2 + (b_R - b_{S1})^2} \qquad (3.6)$$

where *r, g,* and *b* represent RGB channel values. The resulting grey-scale matrices ΔE are shown as colour maps marked with the colour scale (Fig. 3.7), where higher ΔE values are marked in warm colour, and smaller ΔE values are drawn in blue. As references for the initial level of colour difference, the

Frame No.	Original dE	DaVinci Restored	RGB codebook	Simple codebook	Multi codebook
		Colour difference (dE)			
S2	24.3901	7.9507	7.3894	4.7016	3.8401
S3	6.3129	7.7945	9.0021	6.2975	4.1670
S5	22.6927	8.8564	8.7019	6.7695	4.3850
		PSNR (in dB)			
S2	24.3901	22.1350	21.1825	25.3592	25.8963
S3	6.3129	19.4800	19.3644	22.3881	25.1200
S5	22.6927	19.6584	19.6892	22.5719	25.3910

Table 3.1 Calculation results of averaged colour difference dE and peak signal-to-noise (PSNR) level of frames S2, S3, and S5.

original colour difference maps (Fig. 3.7a) are also obtained by calculating ΔE between the good reference S1 and each target image before restoration. Areas highlighted in red and yellow indicate more significant colour differences with respect to the reference, thus more degraded than the areas marked in blue. However, it is worth mentioning that the rise of ΔE in contours is due to the non-perfect alignment between the RGB image pairs, as is evident in Figure 3.6f S3. Then a quantitative estimation of the overall performance *dE* for each image is obtained by averaging the colour difference ΔE for all pixels:

$$\text{dE} = \frac{1}{M} \left(\sum_{i=0}^{M} \Delta\text{E}_i \right) \tag{3.7}$$

and listed in Table 3.1.

Observing the restored R2 and R5 images in Figure 3.6e, it can be highlighted that even though the basic structure of the images is maintained, many pixels were mismatched, especially in the most extensively degraded areas. This is confirmed by the colour difference map shown in Figure 3.7c. For frames S2 and S5, even though the overall colour differences have decreased, from 24.3901 before restoration to 7.3894 after restoration in S2 and from 22.6927 before restoration to 8.7019 after restoration in S5 (as reported in Table.3.1), there are

several pixels mismatched, mainly in the most extensively degraded areas (Fig. 3.7c). As a comparison, the results obtained by DaVinci manual restoration (Fig. 3.6d) are more uniform in overall appearance without the irregular mismatching pixels (Fig. 3.7b). However, the average colour difference for conventional software method is slightly higher than the RGB codebook results, where for results D2 dE = 7.9507 (dE_{R2} = 7.3894, Table 3.1) and for D5 dE = 8.8564 (dE_{R5} = 8.7019, Table 3.1), resulting from the overall shifts in colour. The accurate colour representation is quite challenging using conventional restoration software.

On the other hand, the original frame S3 (Figure 3.6c S3), which is much less degraded than samples S2 and S5, is more uniform in colour and has more negligible colour difference with respect to reference S1 (dE =6.3129, Table 3.1). However, since the degradation characteristics are very different from those included in the fade reference (Figure 3.6a), the restoration result (Figure 3.7c R3) is less accurate with a higher dE value (9.0021, Table 3.1), especially in the wall painting areas and dark part on the right of the image. For restoration result D3 obtained with DaVinci Resolve (Fig.3.6d), the colour difference dE is also elevated to 7.7945, though still the smallest among all three frames. In this case of the RGB codebook, the matching accuracy is primarily limited by the short spectral vector formed from the RGB triplet values as source data. Nevertheless, given the overall evaluation and quick processing time (seconds) compared to hand restoration, the VQ technique still performs promising.

3.4.2   Vector Quantization: simple codebook approach

The proposed codebook method was performed on the spectral data obtained with the VNIR hyperspectral camera [65]. The collected data cube with dimensions of 2,448 x 1,400 spatial pixels x 2,048 wavelengths was cropped and binned to a spectral range limited to 380 nm to 780 nm (240 channels separated by approximately 1.8 nm) and spatial dimensions of 1,300 x 1,040 pixels as described in the section above and shown in Figure 3.8a. The high-resolution three-dimensional data cube provides richer than the RGB image since it contains
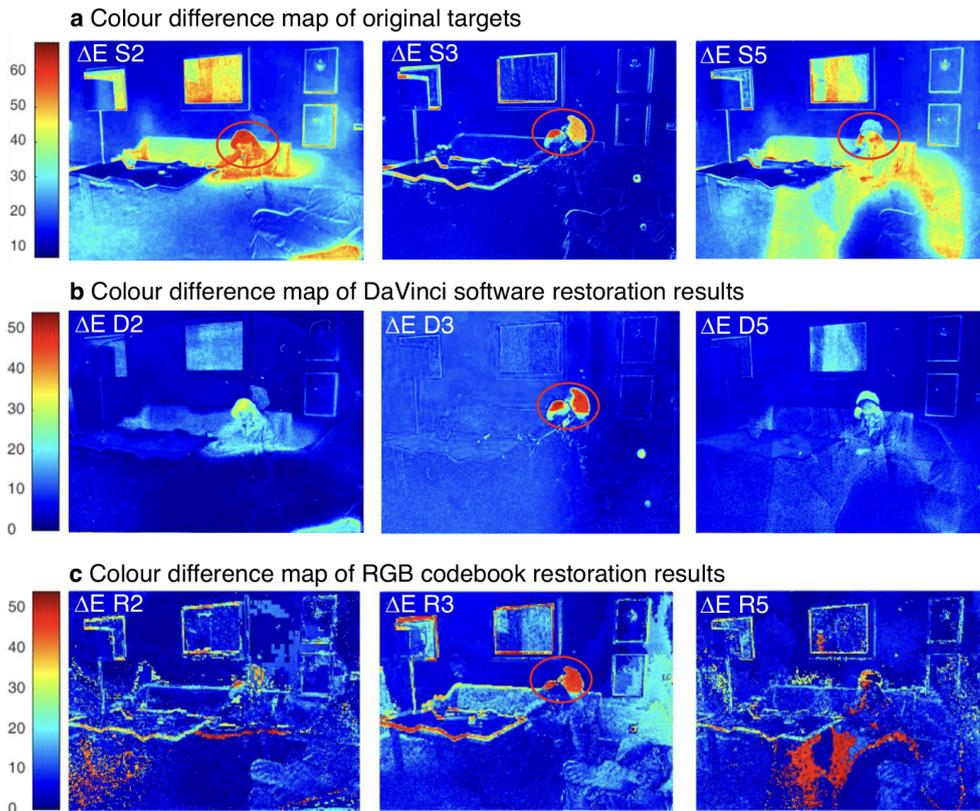
Fig. 3.7 Evaluation of the results obtained via RGB images-based approaches. a) Colour difference (ΔE) map of the original target frames S2, S3, and S5 compared to the reference S1. b) Colour difference (ΔE) map of the restoration results D2, D3 and D5 using DaVinci software. c) Colour difference (ΔE) map of the restoration results R2, R3 and R5 obtained via RGB triplet codebook approach.

spectral features associated with dye molecule deterioration. Reflectance values are affected by two factors. First, every pixel on the film may contain a different level of dye density due to variations in image content. For instance, the brighter wall (pixel W in Figure 3.8) has a lower dye density than the dark floor (pixel F in Figure 3.8), thus a higher reflectance value. Secondly, less light is absorbed as the fading becomes more severe, contributing to the more intense reflectance values observed.
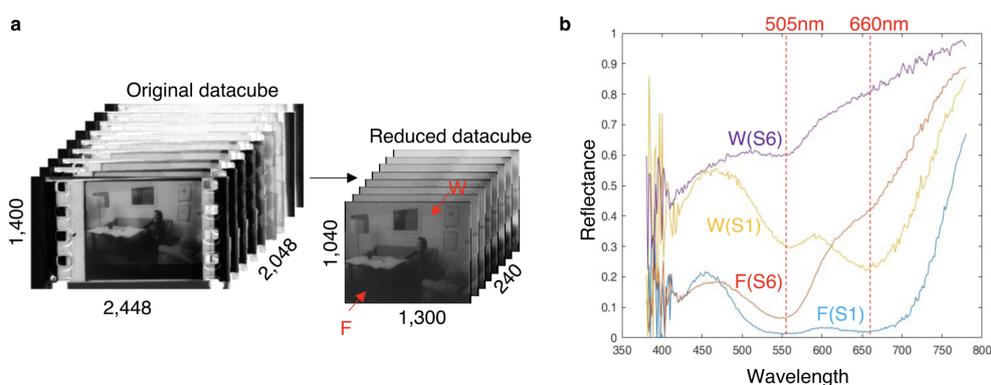
Fig. 3.8 Datacube processing and selected spectra. a) Grey-scale illustration of original datacube before and after initial processing. b) Comparison of spectrum extracted from the same x and y coordinates on S1 and S6, respectively.

Two pairs of spectra, one from the wall (pixel W) and the other from the floor (pixel F), were extracted from corresponding pixels in the better-preserved S1 and the degraded S6. The comparison of spectra (Figure 3.8b) shows that the wall pixel on deteriorated film W(S6) has a significantly higher reflectance across the spectral range than sample S1. Furthermore, the deteriorated film F(S6) drastically loses the absorption band around 660 nm compared to the best-preserved sample F(S1). This feature is related to the degradation of cyan dyes that lead to the overall purplish hue of the faded film. Those subtle spectra variations serve as the fingerprints for looking up the most representative spectrum in the codebook.

For convenience, the restored data cubes are transformed into RGB space, as shown in Figure 3.9a, using wavelength weighting methods. It can be observed that the restoration results (Fig. 3.9a) contain many fewer mismatched pixels than those tested using RGB triplet values (Fig. 3.6e), with overall lower dE values in all three frames (Table 3.1). Employing the vector quantization algorithm, however, the accuracy of the restoration is still dependent on the selected reference spectra. If a spectrum on the target image is not contained in the reference, the perfect match cannot be found, and a rise in noise level and shift in hue would be expected. As already discussed in RGB codebook results (Fig. 3.7c R3), frame S3 restored using a simple spectral codebook (Fig. 3.9a R3) still presents unwanted bluish hues in the background and mismatched shade on the right (marked in the yellow square), though much more limited than what obtained with the RGB approach. The improvement of colour difference (from 9.0021 in RGB codebook to 6.2975 in simple codebook approach, Table 3.1) could also be observed in the colour difference map (Figure 3.9b R3), where pixels with high ΔE value disappeared in most of the areas and decreased in intensity as compared to Figure 3.7c R3. On the other hand, frames S2 and S5 that have similar degradation features with reference S6 achieved better restoration accuracy, with significant mismatches corrected.

3.4.3   Vector Quantization: multiple codebook approach

To improve the representativeness of the spectra and the algorithm's applicability, a multi-codebook was created using spectra hand-selected from multiple frames. Since the data cube is collected with a high spatial resolution, adjacent pixels are mostly similar and highly repeating. Taking every pixel into the codebook would produce a large vector, resulting in a high computational load. As illustrated in Figure 3.10, a selection of pixels from S2, S3, S5, and S6 was employed to create the multi-codebook. S1 was used as an unfaded reference to each faded reference frame, and the pixel-to-pixel correlation was maintained between the pair of references. One out of every five columns for each frame is kept reducing oversampling, preserving the total variance and representativeness of the spectral

a. Reference and results of Simple codebook approach
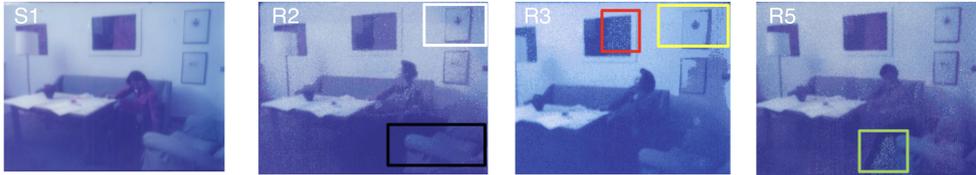
b. Colour difference maps of the results
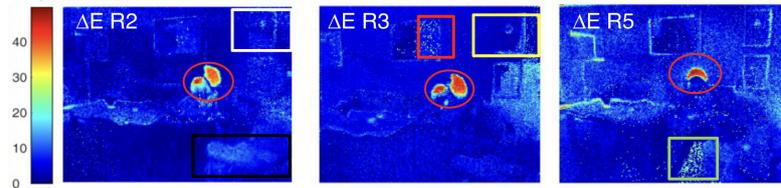
Fig. 3.9 Results and evaluation of simple codebook approach. a) Digital restoration outcomes R2, R3, and R5 were obtained via a codebook approach visualized in RGB format. b) Colour difference map of the above restoration results in a simple codebook approach.

features. Then, the four compressed reference cubes are combined to form a new multi-codebook reference containing spectral information from all instances. The multi-codebook was tested for each frame via the same vector quantization and index substituting process.

The results of the digital unfading via the multi-codebook approach are reported as RGB images in Figure 3.11a. Restoration using the three-codebook approach (Fig.3.6, Fig.3.9, and Fig.3.11,) achieved the most satisfactory restoration accuracy with the lowest dE values (Table 3.1). From the colour difference maps (Fig.3.11b), it could also be confirmed that the unevenly degraded pixels in the original frames (Fig. 3.7a) are recovered, achieving a uniform appearance, and the mismatches occurred in RGB codebook, and simple codebook approaches are now corrected. For frame S3, the noticeable shift in hue and the mismatching on the right (highlighted in yellow square in Figure 3.9 R3) disappeared when the more representative multi-codebook was used (in Fig.3.11 R3). Furthermore, the noise of the painting hanging on the wall (highlighted in red square in Fig.3.9 R3) is removed, and the colour is restored closer to the
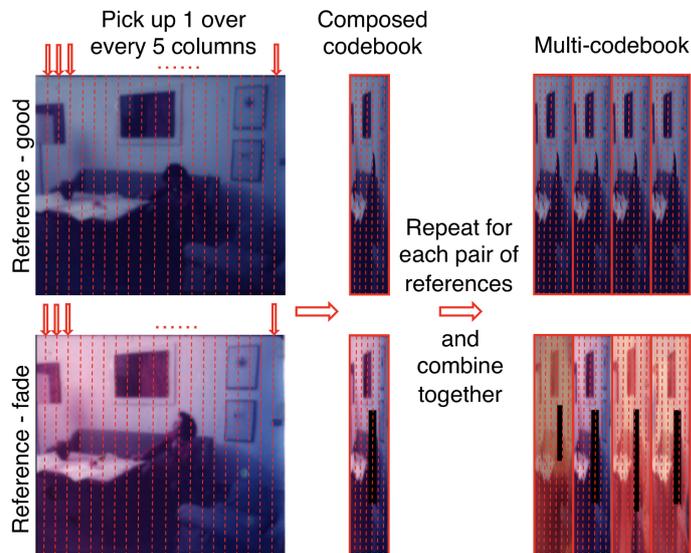
Fig. 3.10 Schematic overview of constructing multi-codebook, using spectra hand-selected and combined from multiple samples S2, S3, S5, and S6.

reference when using a multi-codebook (Fig. 3.11 R3). For frame S2, the lighter desk and the brightness and contrast of the sofa (highlighted in a black square in Fig. 3.9 R2) are corrected and improved when using multi-codebook (Fig. 3.11 R2) and are now closer to the reference image. The restoration results obtained through simple and multiple codebooks are comparable for sample S5, which shares the most similar degradation features with reference S6 (Figure 3.6a). At the same time, a small noisy area around the sofa (highlighted in the green square in Fig. 3.9 R5) is successfully corrected with a multi-codebook (in Fig. 3.11 R5). The pixels of significant colour difference concentrated in the head area (marked in the red circle in Figure 3.11b) are not restoration errors, as could be checked with Figure 3.11a, but due to the non-overlapping of the human figure (as already noticed in Figure 3.7a) in the target frames with the reference S1.

a. Reference and results of Multi-codebook approach

b. Colour difference maps of the results
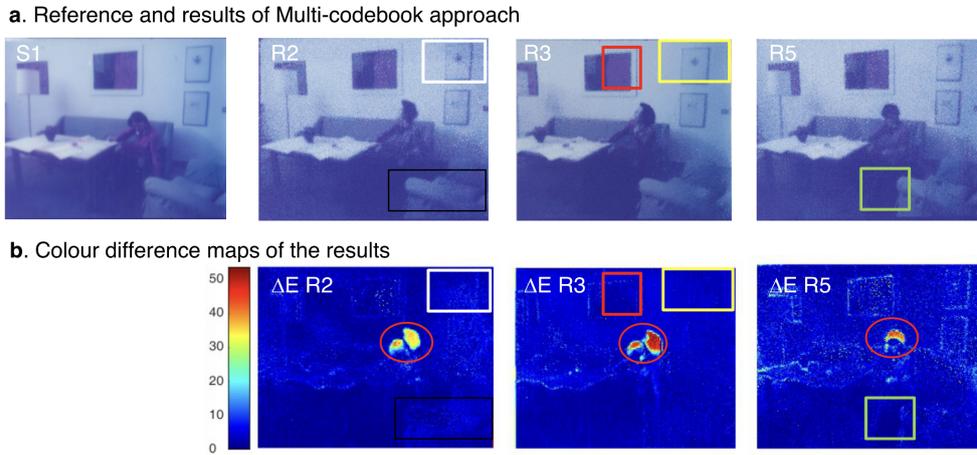
Fig. 3.11 Results and evaluation of the multi-codebook approach. a) Digital restoration outcomes R2, R3, and R5 were obtained via a multi-codebook approach visualized in RGB format. b) Colour difference map of the above restoration results in a simple codebook approach.

3.4.4   Comparison of Different Restoration Methods

To better evaluate the quality of the reconstruction, peak signal-to-noise ratio is also estimated for all the results (Table 3.1). The Root Mean Square Error (RMSE) between the restoration result R and the reference B is first calculated. Then, the PSNR value (in dB) of the reconstructed image R is estimated by:

$$PSNR = 20\log_{10}\left(\frac{MAX_B}{RMSE}\right) \tag{3.8}$$

where MAX_B is the maximum signal value in the referential ground truth image B. This measurement is conducted both on the restoration results obtained from the conventional restoration software and on the transformed RGB representation from the spectral results achieved via codebook approaches. The image format is double decibel in our process, so the MAX_B here is 1. Following the definition of PSNR, the higher the value, the better the quality of the degraded image being reconstructed.

Observing the results in Table 3.1, the results achieved with RGB triplet codebook are comparable with those obtained with the conventional restoration software DaVinci Resolve 17, while the spectral codebook approaches tend to have significantly higher PSNR levels, indicating a better reconstruction quality. The general tendency of increasing PSNR level when including more elements in the codebook is also observed. The multi-codebook outperformed all other tested methods, obtaining the highest PSNR level while achieving the best colour difference metrics. Even though the image quality is improved the most with the multi-codebook approach, there is still noticeable noise in the reconstruction results. This noise is presumably connected to the still over-abundant sampling rate when constructing the codebook. Other data reduction methods, such as clustering and segmentation techniques to remove the non-correlated elements and only keep the representative centroids of each cluster, may help in further reducing the noise.

3.5   Conclusions and Perspectives

In this chapter, we focus on the application of machine learning to address degradation issues, presenting a case study of the digital restoration of degraded cinematic films. This chapter demonstrates that imaging spectroscopy combined with digital unfading machine learning technique successfully restores historic motion pictures with inhomogeneous fading, obtaining a result which is hard to achieve with conventional methods. Our vector quantization method has been positively tested with a pipeline of data processing techniques to restore faded cinematic film, mainly because of the high-resolution spectral features that capture the minimal but essential differences among pixels.

In terms of transferability, the constructed multi-codebook exhibits promising potential for the restoration of deteriorated films of the same type. Additionally, the spectra bank collected in the codebook could be further expanded to include various types of samples and degradation effects, enabling its application to a broader spectrum of damaged films. This advancement may facilitate the

simultaneous automatic restoration of multiple images from the same movie. However, the effectiveness of the proposed method relies upon the availability of a reference non-degraded frame to achieve authentic and subjectively accurate restoration results. The method's applicability is notably constrained by the absence of ground-truth images, limiting its scope. Another drawback lies in the large size of the high-resolution spectral data files, potentially leading to prolonged computational processing times.

As a future perspective, the method can be further improved by combining with advanced clustering techniques to include only the centroids of each obtained cluster in the spectra bank. Forming a more compact multi-codebook, the computational load will be further lightened, and the noise level is expected to be reduced. The proposed method could also be transformed into a robust dictionary learning program to complete the tasks.

Chapter 4

Neural Networks for Hyperspectral Imaging of Historical Paintings

> Can machine learning techniques address complex identification and classification problems in CH? How transferrable are the machine learning models developed to similar scenarios in cultural heritage conservation and what future applications can be envisioned?
>
> **— RQ-1.1, RQ-1.3**

This chapter provides an exhaustive analysis of the literature related to Neural Networks (NNs) applied for hyperspectral data in the Cultural Heritage field. We outline the existing data processing workflows and propose a comprehensive comparison of the applications and limitations of the various input dataset preparation methods and NN architectures. By leveraging NN strategies in CH, this chapter contributes to a wider and more systematic application of this novel data analysis method.

4.1   Introduction

Hyperspectral imaging (HSI) has become widely used in the field of cultural heritage (CH) for painting analyses supporting the identification of original and degraded paint compounds, and revealing underdrawings [6, 74, 115, 152]. The non-invasive and non-destructive analysis tool can be applied in situ in either reflectance, transmittance, or fluorescence mode [108, 157, 179]. It registers a spatial map and collects a spectrum at each pixel position, in the wavelength range varying from infrared to X-ray. The contiguous set of 2D images collected through the energy range of interest produces a 3D data cube containing millions of spectra, which poses an analytical and computational challenge. Thus, the effective processing of heavy spectral datasets remains an active research area with the primary challenge being to efficiently process data and extract rich information from the inherently complex and delicate painting materials [57, 58].

A range of cutting-edge data processing techniques has particularly been applied in the CH domain to reduce the dimensionality of the dataset, classify spectral signatures and finally unmix spectral signature to map paint components. Thus, numerous approaches were developed, starting from the conventional multivariate analysis and statistical methods (e.g., spectral angle mapper (SAM) [14, 55, 63, 68], fully constrained least square (FCLS) [78, 82, 162], principal component analyses (PCA) [172, 150, 69, 161], minimum noise fraction transform (MNF) [39, 44], and k-means clustering [15, 165, 193]), to the more advanced machine learning algorithms (support vector machine (SVM) [80, 153], hierarchical clustering [122], embedding techniques [7, 199, 156, 202], MaxD [15, 14, 100], dictionary learning [121, 201]) with a growing interest for neural network algorithms (NNs) [72]. NN-based models first gained a tremendous rise in digital image classification due to their superior ability in feature extraction and pattern recognition [148, 170, 171, 198]. Then, the application of NNs for spectral imaging data drastically expanded over the last 5 years for performing advanced mapping of artistic painting materials (Figure 4.1) [49]. NNs' main advantages reside in the extreme flexibility of the types of data they can process, with superior ability to extract hidden and sophisticated structures (both

linear and nonlinear) contained in the raw spectral data [1]. This relies on the large flexibility of the NN architectures available (number of neurons, type of layers, and depth), making NNs adaptive for a wide range of tasks and input data [148]. Unlike other machine learning algorithms, the learning of NNs is fully automated, without the need for extensive manual tuning of the underline functions. This powerful and efficient learning ability makes them suitable for complicated tasks and complex painting systems, where the large nonlinear data manifold becomes a merit.
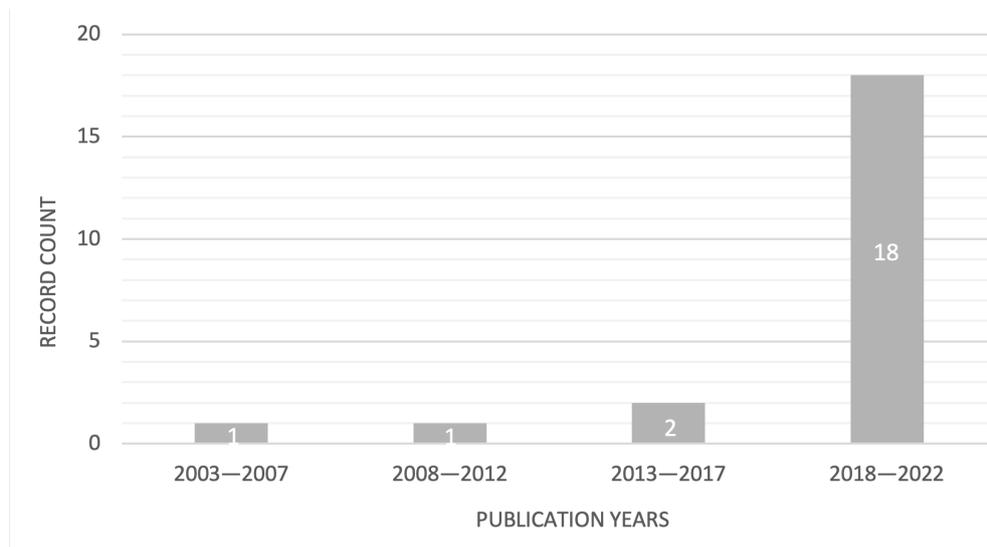
Fig. 4.1 Record count of the scientific publications related to the use of NNs to process hyperspectral datasets from CH paint-based materials over the last twenty years (Web of Science, December 2022 [49], using keywords "paint", "spectral imaging" and "neural network").

In this circumstance, the rapid rise of NN-based strategies has outlined a variety of data analysis workflows available to the community. However, the flexibility in the architecture designs and the automated learning process (often referred to as "black-box") cause difficulties in application, and very rare studies have documented their choice on the architecture and appropriately described the use of specific features. In this context, the present article proposes a comprehensive overview of the applications of NNs for HSI of historical

paintings. A systematic and exhaustive search was conducted in the time frame up to December 2022 [21, 42, 43, 53, 77, 80, 97, 101, 102, 117, 118, 146, 155, 160, 159, 163, 164, 178, 180, 188, 190, 189, 212], which included database searches on Web of Science [49] and manual searches of relevant journals, conference proceedings, and PhD theses. The inclusion criteria were set on implementing NN techniques for processing HSI data in the CH field, with the main focus on solving paint component mapping problems.

The remainder of the chapter is structured as follows. Section 4.2 illustrates the research questions of this chapter. The basic terminology and working principle related to NN is introduced in Section 4.3. Then, we describe the two main applications involving the use of NN for the analyses of data acquired on historical paintings in Section 4.4. In Sections 4.5 and 4.6 we present adaptations of the dataset and NNs architecture to the specific constraints of ancient and historical materials. Finally, Section 4.7 discusses and questions the prominent trends that may lead to the future novel implementation of NN algorithms for historical paintings.

4.2   Research Questions

Starting from the background introduced in the previous section, in this chapter we aim to respond to the following research questions:

RQ-1.1 *Can machine learning techniques address complex identification and classification problems in CH?*

RQ-1.3 *How transferrable are the machine learning models developed to similar scenarios in cultural heritage conservation and what future applications can be envisioned?*

4.3   NN Working Principle

NNs are a group of computational algorithms that receive and process informa-
tion in a way that mimics the neural system of a brain. Most applications in the
CH field have adopted the feed-forward structured NN schematically presented
in Figure 4.2 [192]; hence, in the following, we concentrate on this type of NN.
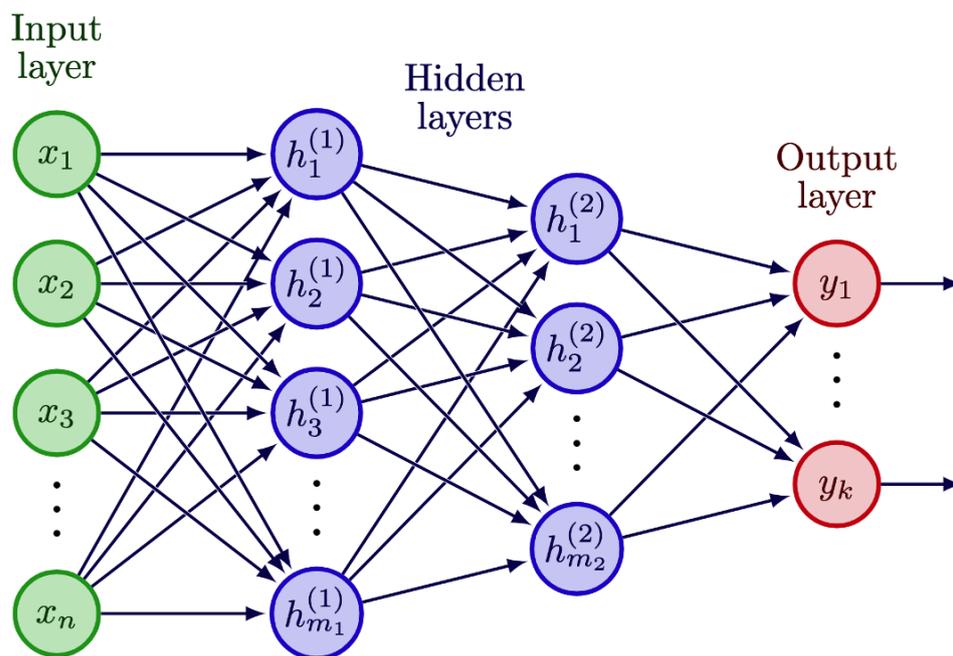
Fig. 4.2 A schematic illustration of an example of a fully connected feed-forward
neural network with two hidden layers. x, h, and y represent the neurons in the
input, hidden, and output layers, respectively, while n, m1, m2, and k are the
total number of neurons in each layer.

The basic units of a neural network are called "neurons". Each neuron
receives an input, performs a computation, and passes the output to other neurons.
The neurons are aggregated in layers of different functionality. The input layer
directly receives the input data, its output is passed to the hidden layers which
perform the computation, and their output is passed to the output layer which
produces the final results. The output of a given neuron is obtained as a weighted

sum of the output values of the neurons from the previous layer plus a bias
[175, 181]. The obtained value is then usually passed through a non-linear
activation function, thus producing the neuron output. The ensemble of weights
and biases is the model parameters of the NN, and is automatically optimised
via learning. A NN with more than one hidden layer is also known as a deep
learning (DL) model. An additional set of parameters, called hyperparameters,
controls the speed and the quality of the learning process [64]. These parameters
need to be defined before the learning process begins and are further tuned in
order to obtain the optimal model for a given dataset. Below we list some of the
most common hyperparameters encountered in the CH literature:

- *Number of hidden layers and neurons in each layer*. The number of layers
  defines the depth of the network and is at the heart of the architecture
  affecting the performance of a NN [4].

- *Activation functions* define how the inputs to a neuron are transformed into
  an output to be fed into the next layer, where sigmoid, ReLU, and softmax
  functions are frequently used [98].

- *Loss functions*, such as mean squared error (MSE) and cross-entropy loss,
  are used to estimate the error between the ground truth and predicted
  output values [214].

- *Learning rate* defines how quickly a network updates its parameters to-
  wards convergence [86].

- *Number of epochs and batch size*. Epochs are the number of times the
  training data appears to the network and batch size determines after how
  many input sub-sets the network should update its parameters [86].

In a feed-forward NN, data flows through the network in only one direction,
from the input layer to the output layer, without looping back, as in recurrent
neural networks (RNNs) [99, 119] or generative adversarial networks (GANs)
[56, 207]. There are different types of hidden layers used according to the

specific architecture, such as fully connected layers, convolutional layers, or pooling layers [5, 145]. The various NN architectures with different choices of hidden layer types are detailed in Section 4.6.

After defining the NN architecture and hyperparameters, the datasets are usually divided into a training, validation, and test set. During the training process, the model is presented with a training set, which is a set of samples that the model uses to learn the relationships between the input features and the target outputs. The model parameters (weights and biases) are optimised during this step. The system auto-updates itself through the feeding of input data, without the need to hand-adjust up to millions of parameters in the network. A validation set is used during training to adjust the hyperparameters in order to improve the performance of the model and prevent possible overfitting on the data. In the most common case of feed-forward NNs, the optimisation is accomplished through a process called backpropagation that involves two steps. First, during the forward propagation, the input data is introduced to the input layer, and is transmitted through each hidden layer in sequence until an output is generated. By comparing the output result with the ground truth, an error signal is computed for each neuron in the output layer, indicating the direction and magnitude to adjust. Second, during the backpropagation, the error signals are transmitted backwards from the output layer to each node in the last hidden layer, and repeatedly propagate backward layer by layer. After the model has been trained, it is tested on a previously unseen test set to evaluate its performance and the ability to generalise to unseen data.

There are several learning strategies that have so far been applied in the field of CH: supervised learning, unsupervised learning, and transfer learning [84, 192, 216]. In the first case, the available dataset consists of data points and their labels, and the NN algorithm learns a function which maps the input to the output by learning from available labelled examples, as observed in most cases [42, 43, 53, 77, 80, 101, 118, 146, 155, 160, 159, 163, 164, 178, 188, 190, 189]. The second type of learning, unsupervised learning, is used to analyse and cluster unlabelled data [102, 117, 180]. In the case of transfer learning, the model is pre-

trained with a large dataset and then fine-tuned with a separate smaller dataset [21, 97, 212]. This type of learning is discussed in detail in Section 4.6.3.

4.4   NN Application to the Study of Artistic Painting

The ability of NNs to efficiently learn high-level features from vast data has gradually attracted the interest of CH scholars in assisting with spectral data processing.

The preliminary applications of the NN models for CH were implemented on multispectral imaging (MSI) datasets twenty years ago, and focused on the improvement of spectral data resolution and the digital reconstruction of reflectance curves [42, 146, 178]. Super-resolution spectral reconstruction was first studied in the computer vision field with primary attention on the device-independent digital image reproduction using spectral reflectance as the intermediate space for colour management [9, 172]. The very first example of using NNs on multispectral data in CH applications was conducted by Ribés and Schmitt in 2003 [53, 160, 159], to achieve high fidelity and illuminant invariant colour reproduction of fine art paintings. They developed a non-linear model to learn the spectral reflectance curves from multispectral images. Osorio-Gomez et al. later proposed an alternative method for similar tasks on oil canvas paintings [146]. Recently, a similar approach by Shi et al. [178] further optimised faithful print reproduction of oil paintings to model the bidirectional mapping between the spectral reflectance and the 3D printer ink stack layouts. In luminescence MSI development, NNs have also been applied by Chane et al. to learn the radiometrically calibrated emission spectra with resolution in the nanometre scale from 15-band multispectral data [42].

With the significant development and spread of HSI instruments over the last decades, the fast and high-resolution acquisition of the spectral data on artworks has overcome the limitation of multi-spectral sampling. Moreover, in the last five years, the research focus has shifted from spectral super resolution to more

practical conservation and restoration related questions. To date, the objectives of implementing NN on the spectral data analysis in CH are mainly focused on paint component identification and the relative spatial visualisation of their distribution. Conventionally, the material identification is performed through the time-consuming interpretations of measured spectral features by CH experts, which is also challenging when associated with complex material mixtures and layered structures. By implementing NN, the pixel-wise identification process could be efficiently achieved using automated data-driven solutions with minimal human intervention. More specifically, the paint component mappings are generally classified into the following categories, further developed below:

1. *Paint component identification* determines the presence or absence of a pigment based on a given spectrum—classification task;

2. *Paint component unmixing* refers to quantitatively decomposing a given spectrum to its base constituents—regression task.

4.4.1 Paint Component Identification

The majority of the research is focused on the first category of tasks that generates a qualitative map of paint components through NN. The simplest case of such application covers the single pigment identification without consideration of pigment mixtures or binding media. Chen et al. developed a model to automatically identify pure pigment areas of heritage paintings based on visible reflectance (VIS-RIS) data in favour of the authentication of artworks [43]. Jones et al. have recently focused on X-ray fluorescence (XRF) data to automatically classify historical pigments used in Renaissance paintings to help understand the artist's materials and technique [97]. Additionally, a two-step mapping approach has been proposed by Kogou et al. for the material identification of a Peruvian watercolour painting based on macro-XRF data [102]. In this approach, the spectra are first classified into areas with similar features and then the pigments present are hand identified. The proposed strategy has been later applied on visible and

near-infrared (VNIR-RIS) data for bronze corrosion products' identification in a large museum collection [117].

As opposed to the simplified model assumption of a single layer of pure pigments, historical paintings usually feature a more complex mixture of pigments or even have a multi-layered structure. Therefore, a multi-labelling model is necessary for a robust solution to correctly identify the multiple pigments present within every single pixel. Kleynhans et al. have developed a non-linear model to produce a qualitative multi-label map of the intimate pigment mixtures in a single step [101]. Their training dataset is based on VNIR-RIS spectra extracted from well characterised 14th century illuminated manuscripts to make accurate predictions of the materials used by the same art school. In addition to the one-step solution exploiting the hand-labelled database, Grabowski et al. have developed alternative data processing pipelines for a fully automatic pigment classification task [80]. They have broadened the applicability of the algorithms to paintings with unknown material compositions, as they first obtain the labels through automatic identification with several pigment libraries, and then use NNs to classify the entire painting and create a multi-label map.

NNs have also been applied to detect the degradation areas of murals. Mural paintings often suffer from several diseases, such as flaking, deep loss, mud besmirch and sootiness, in various degrees. The documentation of those damaged areas and the identification of their composition is essential for mural conservation and restoration. Sun et al. have applied NN in combination with the dimensionality reduction method, PCA, to identify the degree of flaking in murals at Mogao Grottoes based on spectra in the near-infrared (NIR-RIS) range [189]. Lin et al. adopted a similar approach, combining NN with MNF, to classify the VNIR mural images into various types of damaged regions [118].

4.4.2   Paint Component Unmixing

Several works performed quantitative unmixing of complex paint systems that aim at estimating the concentration of each paint component on a per-pixel basis. Rohani et al. have proposed a quantitative non-linear unmixing strategy

of hyperspectral data that involves two steps: first, the NN decomposes the reflectance into its constituent pure pigments members and then the two-constant Kubelka–Munk model is applied to accurately estimate the concentration of pigments in mixtures [163]. The same authors later developed a semi-quantitative pigment unmixing algorithm that relies on training two identical NNs simultaneously [164]. The first NN performs multi-label pigment classification and assigns multiple pigment classes to every spectrum. The second NN tackles the spectral unmixing problem and outputs coefficient maps, thus providing a semi-quantitative mapping of the abundance of the pigment classes. Pouyet et al. adopted the same architecture and applied it to datasets acquired in the short-wave infrared (SWIR-RIS) domain [155]. Another study by Fukumoto et al. applied the encoder–decoder (ENDEC) NN model (fully described in Section 4.6) that estimates the pigment concentration as an intermediate output and includes a further spectral remix step to reconstruct the predicted spectrum from the estimated concentration [77].

Apart from paint component mapping, NN implementation has also been explored to solve other spectral unmixing problems related to the layered structure of paintings, such as thickness estimation and paint segmentation. Xu et al. have targeted the multi-layered structure of the painting that also applied a NN to fully automate the pigment identification process [21]. Their work is based on XRF data and has simulated XRF spectra with three-layered pigments to train the model. Shitomi et al. have developed a physics-based ENDEC model (fully described in Section 4.6), where the decoder part is based on the Kubelka–Munk (KM) model, to estimate the thickness and mixing ratio of pigments [180]. Research conducted by Striova et al. applied the NN to improve the visibility of the pentimenti and underdrawing style based on VNIR multispectral data acquired on a pair of homonymous paintings by Manet and Titian, respectively [188]. Furthermore, in assisting the study of historical painting styles, Zhang et al. have developed a strategy to extract the sketches of damaged or degraded paintings that also exploits spatial features [212]. In a recent advancement, Sun et al. adopted a pre-trained model originally designed for photo restoration to virtually

repair the scratched mural paintings that also extend the applicability of NN to digital restoration [190].

4.5   Datasets Preparation Methods

The spectral input sources, data types, pre-processing steps, and labelling methods are crucial for the dataset preparation. Figure 4.3 summarises the general workflow across the literature and all the alternatives in this preparation process. After collecting and pre-processing the available data, the obtained dataset is split into two or three subsets: training, validation, and/or test set. A common data splitting ratio is 80% for the training and 20% for the validation and/or test [77, 155, 188, 189]. The alternatives are 64% training, 16% validation, and 20% test according to [163], or 70% training and 30% validation as in [43].
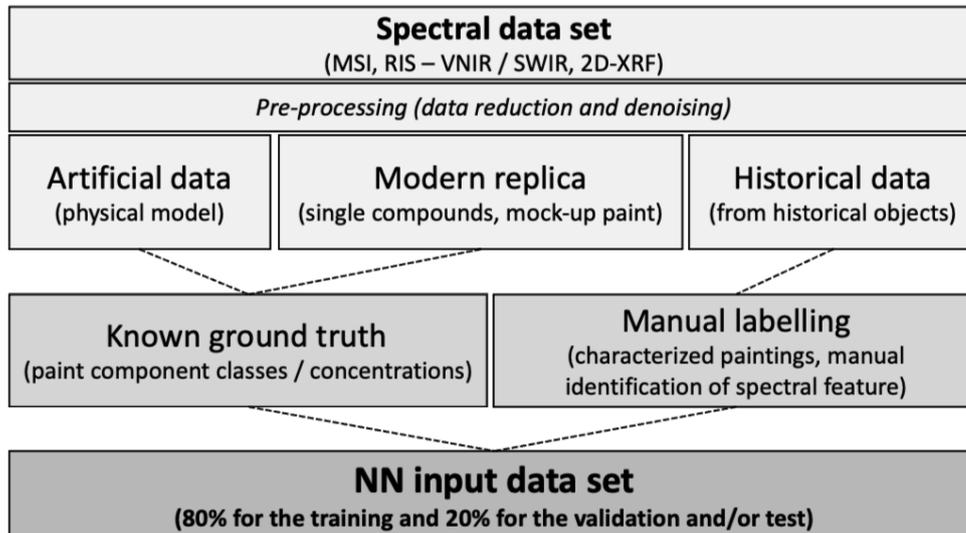
Fig. 4.3 Summary of the dataset preparation workflow.

4.5.1 Spectral Inputs

As introduced in Section 4.4, the early studies were mainly conducted on MSI data for spectral reflectance curves reconstruction. These data were mostly acquired or simulated in reflectance mode, with one exception utilising a 15-band luminescence acquisition in the 450–740 nm range [42]. The reflectance data mainly fall into the VIS domain, either collected with seven narrow-band interferential filters in the range of 480–650 nm [146], or 31 bands with high-dynamic range (HDR) multispectral capture within the spectral range of 420–720 nm [178]. Multispectral data extended to the SWIR range was also inspected, as presented by Striova et al. in 2018, who acquired 32 narrow-band images with spectral coverage from 390 up to 2500 nm [188]. Furthermore, Ribés et al. proposed using artificial data that simulate seven-channel camera responses over the 400–760 nm range [53, 159].

In more recent studies, the hyperspectral data were widely used as inputs to the neural networks. The spectral range of the input data varies from RIS to macro-XRF, and the data samples are characterised with improved spectral resolution down to the nanometre scale and increased wavelength channels up to hundreds or even thousands. Some of the studies are based on XRF data, acquired in various experimental conditions (power source, acquisition time, and beam size) [21, 97, 102]. The RIS data cover several wavelength domains, from the most conventional VIS range (380–750 nm [77], 400–700 nm [180], 400–720 nm [43]), extended to NIR (383–893 nm [163, 164], 377–1033 nm [190], 377–1037 nm [212], 400–950 nm [101], 400–1000 nm [117, 118], 822–1719 nm [189]), to the recently emerging SWIR range (930–2500 nm [117], 1000–2500 nm [80, 155]).

4.5.2 Data Types and Labelling Methods

Artificial data

In the shortage of adequately labelled datasets, physical models can be applied to generate a library of synthetic spectra with a sufficient amount of training data.

Two approaches have been used in accordance with different spectral inputs: the fundamental parameters (FP) method for XRF spectra and the Kubelka–Munk (KM) theory for RIS data. For these data, the ground truth is known, favouring not only the multi-label classification, but also quantitative estimation or even unmixing in multi-layer structures. The FP method is based on Sherman's equations that describe the theoretical relationship between the measured XRF signal intensity of the element and its concentration in the sample [177]. With defined pigments and concentrations, simulated spectra responses are generated. Jones et al. applied this method to create a training dataset containing 3000 spectra of 15 pigment classes [97], modelled as a simplified case that only accounted for the primary fluorescence of single-layered samples. Xu et al. [21] have employed a more complex model that simulates XRF spectra for three-layer structures in mixtures of 12 different pigment classes, resulting in 16,224 spectra in total. Both studies have adopted a transfer learning strategy that further fine-tunes the NN model with a smaller quantity of experimental data, after the initial training with the large synthetic dataset.

For the reflectance data, Kubelka–Munk (KM) theory is widely used to estimate the interaction between the incident light and paint layers [19, 139]. This physical model measures a sequence of absorption and scattering coefficients to predict the reflectance spectra from the composition of pigment mixtures. The KM theory can be applied to decompose the measured spectra and thus estimate the pigment concentrations via spectral unmixing. Moreover, it can be used for mixing models that generate many patterns of synthetic spectral data, with given pigment types and concentrations. Rohani et al. have applied a nonlinear KM mixing function to generate simulated spectra of pigment mixtures [163, 164]. All possible two/three mixtures out of 11 pure pigments [164] and 12 pigments [163] were modelled and 500 random combinations of coefficients/concentration values were selected for each mixture, obtaining 110,000 spectra and 143,000 spectra in total , respectively. Fukumoto et al. have used a similar approach to build the synthetic dataset based on 19 reference pigments [77]. Two datasets were prepared: the first one that served as a trial to validate the method contained

1,771 spectra of only one combination and the second one contained 1,445,136 spectra which included all combinations of 4 colours out of 19, including white with 5% mixing steps. Furthermore, Shitomi et al. synthesized the spectral responses of layered pigments with a given layer thickness and mixing ratio [180]. Three pigments of primary colours are used for the simulation with thickness varying from 0 to 3.2 μm, generating in total 35,700 spectra.

Chen et al. used an alternative method to create an artificial spectral database [43]. They generated augmented pure pigment samples by adding random noise and amplitude offset to the reference pigment database, and synthesised simulated colour mixtures using an ideal subtractive mixing function. All possible mixtures of 3 pigments out of a selected set of 16 pigments were generated with increments of 10%, yielding a total of 21,240 spectra.

Modern data

Even though artificial data is advantageous regarding the large amount of data and the known ground truth, it remains challenging to fully simulate the instrumental and environmental noise, and the real physical processes. To address this issue, many cases included the data acquired on modern samples in the training process. These samples are prepared prior to analysis, including single pigment powders or pellets, and mock-up paintings. The true concentration, or at least the known composition of paint layers, can be used to automatically label the training data.

A few studies have measured the spectral response directly on dry pigment samples as either pigment powders or pellets of a single pigment, with small data sizes up to hundreds of spectra [80, 97, 117, 155]. Most generally, mock-up samples are prepared to imitate the historical painting materiality (binder, pigment(s), preparation layer, and support). The pigments selected to create the mock-ups are commonly based on analysis of the historical artworks and the range of pigment fractions, binder ratios, and layer thicknesses are consistent with that in the historical objects [21, 80, 155, 180]. The spectral data acquired on the mock-ups covered both pure pigment areas and two to three pigment

mixtures, with also variances in paint thickness in some cases. The input datasets have an average size of 80 thousand spectra.

In some cases, the mock-ups were used to test the performance of the NNs previously trained on artificial datasets. Rohani et al. created mock-up paintings with selected combinations of mixtures and pure pigment examples to test their model [163, 164]. Fukumoto et al. applied their NN on 33 mock-up paintings made by previous work [77], and Xu et al. created a set of mock-up paintings, of which 20% of the obtained spectra were used to fine-tune the model and 80% served as the test case [21].

Historical data

Historical data are those acquired directly on historical objects. Those data are either included in the training process by forming the input dataset or serving as a test case to verify the performance of the proposed model. The types of art objects and the sizes of the dataset used present a large variability. Within the historical set, both labelled and unlabelled data are used.

Kogou et al. used an unlabelled dataset composed of 41,327 XRF spectra acquired on a Peruvian watercolour painting (c. 1860) by Francisco Pancho Fierro [102]. With the aid of an unsupervised NN model SOM, the large number of XRF spectra was reduced into 13 distinct clusters of unlabelled material compositions. Liggins et al. applied the same model to two excavated bronze fragments, classifying the SWIR and VNIR images with hundreds of thousand spectra (exact size unreported) into three corrosion areas [117]. They also built a reference spectra database composed of 17 samples of powdered and fragmented ancient Chinese bronze extensively characterised by other analytical instruments. The identification was performed by manually comparing the obtained averaged spectra in each cluster with the reference.

On the contrary, supervised networks require labelled input data, which is quite challenging for historical objects that are usually unknown and heterogeneous materials. The main labelling method proposed in the literature relies on the visual inspection of the spectral features, and manual annotation to mark areas

with a certain presence of pigments or degradation features. To confirm this manual identification, complementary analyses can be used, such as RIS, XRF, Fibre Optics Reflectance Spectroscopy (FORS), X-ray Diffraction (XRD), Raman and Fourier Transform Infrared (FTIR) spectroscopies [80, 102, 117, 118, 189]. However, this labelling method is very time-consuming. Thus, usually only small inherent areas with limited pixels are selected, resulting in a much smaller dataset size. Moreover, the obtained labels are not suitable for the quantitative assessment of concentrations.

In some cases, the historical objects in question were extensively studied in previous reports and have been well characterised for all chemical components, available as the ground truth maps. Kleynhans et al. used selected areas on four paintings from the illuminated manuscript Laudario of Sant'Agnese (c. 1340) to form their training dataset [101]. In total, 25 classes of paints were identified, and 16,683 individual spectra were collected across all four paintings. Two out of the four paintings were also used to test the performance of the model. Pouyet et al. also used selected areas of the historical object, a Tibetan thangka dated from the beginning of the 19th century, as part of the training data [155]. They have built their input sets from multiple sources, including single-pigment pellets, pigment mixture mock-ups, and the historical thangka, altogether presenting 12,000 spectra. Striova et al. applied the network based on an oil painting Madonna of the Rabbit by É. Manet (c.1856) to improve the visibility of pentimenti and underdrawing style [188]. 50k randomly selected pixels from the analytical data acquired formed their input dataset.

The above-mentioned data types and labelling methods are mix-used in some studies to build a more reliable training dataset. Both historical data and modern replica data were used to build the input datasets in [155]. A large simulated dataset and a smaller mock-up dataset were used in different learning stages in the transfer learning approaches [21, 97].

However, historical data are more frequently used for the test of the developed network as the ultimate goal usually is to study the material composition and distribution of historical artworks: one illuminated folio from the 15th century

Book of Hours in [163]; two impressionistic paintings, the Poèmes Barbares by Paul Gauguin (1896) and The Bathers by Paul Cezanne (1899–1904) in [21]; a set of 11 paintings by the late Portuguese artist Amadeo de Souza-Cardoso in [43]; and an early Renaissance painting, Saint Michael Triumphs over the Devil by Bartolomé Bermejo [97]. In [190], a pre-trained model was tested on damaged murals of a Buddhist temple (c. 1392) to repair the scratches.

4.5.3   Pre-Processing

Several data pre-processing workflows were proposed to improve the data quality before feeding into the NNs. Other than the common calibration process that calibrates the signal with respect to the background and the instrumental environment, the pre-processing mainly focuses on denoising and data reduction.

First, for RIS data, the low wavelength channels (usually the first 15 bands) or the last bands in the infrared range are usually quite noisy and removed as noted in [117, 118, 163, 164, 189]. For XRF data, taking a spectral range starting from 2.80 keV was suggested in [21] to avoid the peak overlaps in the low energy range that often confuses the model and affects the efficiency.

Then, the smoothing of the signal is performed by binning and applying various filters to further improve the signal-to-noise ratio. For RIS data, a spectrally moving average filter that reduced the wavelength channels by a factor of 4 was used in [164], and a Savitzky–Golay filter was applied in [43, 163] to reduce the high-frequency noise. For the XRF dataset, spatial median filtering with a 3×3 pixels kernel and spectral binning on 5 were employed [102]. On the other hand, noise was introduced to the artificial XRF spectra to increase the robustness of the dataset [21].

Data dimensionality reduction techniques, including PCA and MNF, are applied in some cases to centralise the information in fewer bands [118, 146, 190, 189]. An inverse PCA transformation was applied in [189] after processing the first component by high-pass filtering. A pipeline of statistical algorithms, including t-SNE (t-stochastic distributed neighbourhood embedding), DBSCAN clustering, HySime (hyperspectral signal subspace identification by minimum

error), SISAL (simplex identification via split augmented Lagrangian), and SAM (spectral angle mapper), was built in [80] to fully automate the identification process and the results were fed into the NN. Data reduction in relation to balancing the number of samples in each pigment class was also reported in [101]. This task was accomplished by iteratively removing similar spectra (based on measuring Euclidean distance for the similarity) in the over-abundance classes until the samples per class are more evenly distributed.

As suggested by Fukumoto et al., selecting candidate pigments can also be considered a pre-processing step [77]. Even though the detail of pigments considered in each study is not listed and compared in this chapter, it presents a large variability from case to case.

4.6 Model Architectures and Their Evaluation

A wide range of NN architectures has been explored in response to the various research questions, from spectral super-resolution to single or multi-class classification, and from quantitative concentration estimation to spatial feature extraction. In general, as the research problems become more complex, more complicated NNs are implemented to fully learn the structure contained in the input data. In this section, we review the NN models implemented in CH in the time frame from 2003 up to early 2023, roughly in the sequence from the simplest architecture to the more complex ones. We first introduce the unsupervised learning networks with the example of a self-organizing map (Section 4.6.1), then mainly focus on the supervised approaches, such as multilayer perceptron (Section 4.6.2) and convolution neural networks (Section 4.6.3). In the end, the most complex hybrid approaches are presented, including multi-branches networks that mix-used the above basic types, encoder–decoder networks (Section 4.6.4), and deep belief networks (Section 4.6.5). The detailed configurations, structures, choice of hyperparameters, and evaluation methods of the mentioned architectures are summarised.

4.6.1   Self-Organizing MAP (SOM)

A self-organizing map (SOM) is a type of unsupervised learning algorithm, i.e.,
it does not require a labelled dataset. The algorithm takes the pixel-level spectra
as input and maps them onto a 2D layer, the SOM [103]. SOMs thus contain
only a single layer of neurons, in which every neuron is associated with a weight
vector, with the same length as that of the input vectors. The training takes on the
competitive learning strategy, where each neuron competes with its neighbours
to be the "winner" for a given input, with the winning neuron and its neighbours
adjusting their weights to better match the input. The SOM NN has a single input
parameter—the number of neurons, which should be larger than the number
of clusters. This type of clustering method naturally allows for a visualisation
of the high-dimensional spectral data in 2D space in the form of single-cluster
maps. With each cluster representing a distinct group of similar spectra, the
SOM clustering analysis is able to reduce the large number of spectra to a few
clusters of similar spectra, thus capturing variations in the density of the paint
and the relative concentration of materials [102, 117].

SOM has been applied for two different clustering tasks using multispectral
XRF data [102], and SWIR and VNIR data [117]. In the former case, the number
of final clusters was set to 13. In [117], the authors do not provide information
on the number of clusters chosen. In their work, SOM was applied to all spectra
from each of the two image cubes. After initial clustering, the number of data
points was reduced, and the mean spectrum for each cluster was calculated.
Then, a secondary clustering was applied again using SOM, grouping the spectra
by shape but not by intensity. In order to evaluate the clusters, the quantisation
error is computed for each cluster and if it is above a given threshold, the cluster
is disintegrated, and the data is subjected to a new round of clustering. The
threshold is increased stepwise until all data samples are clustered.

4.6.2 Multilayer Perceptron (MLP)

A multilayer perceptron (MLP) is a fully connected class of feed-forward NNs in which each neuron in one layer is connected to every other neuron in the next layer [173], as illustrated in Figure 4.2. An MLP typically has three or more layers: an input layer, one or more hidden layers, and an output layer. An MLP with a single hidden layer is considered a shallow NN, while those with two or more hidden layers are regarded as deep learning models which are frequently denoted as deep MLP or deep neural networks (DNNs). The output layer produces the final prediction, which can be a continuous value (for regression tasks) or a class label (for classification tasks), depending on the activation functions used [98].

Shallow MLP

As an early application, a shallow MLP was applied in [146] to predict reflectance spectral curves. The input datasets are reflectance spectra reconstructed using linear techniques, such as pseudoinverse, PCA, and cubic-spline interpolation. The neural network was composed of an input layer of size $3 \times N$ (where N is the number of channels), a hidden layer of 85 neurons per channel, and an output layer of a single neuron per channel. Linear activation functions were used in the hidden and output layers. The root mean squared error (RMSE) was the metric employed. The neural network parameters were optimised with the aid of the Levenberg–Marquardt algorithm [111, 130]. The training was set to stop when RMSE fell below a certain threshold, which was fixed to a small arbitrary value.

A shallow MLP was also applied in [118] to identify mural disease. A three-layer model of 6, 8, and 5 neurons in each layer was developed. The VNIR spectra of 1,040 channels were pre-processed by MNF to reduce the spectral dimension and minimise the noise. The first six bands of the transformation results were selected to be the input. The output layer corresponds to the five predicted classes, including diseased and non-diseased areas. A single hidden layer of eight neurons was used. No information on activation and loss function

was provided. The number of epochs was set to 1,000, the minimum error of the training target was 0.0001, and the learning rate was 0.01.

A shallow MLP classifier was implemented for automatic pigment identification in [80]. They used an off-the-shelf model from the *scikit-learn* library [174], with hyperparameters set to default values: the model has a single hidden layer of 100 neurons, RELU activation function, and Adam weight optimizer. The input to the NN is the 256-channel SWIR spectra with pigment labels generated from the previous steps using t-SNE, HySime, SISAL and SAM, and the output returns one of the five pigment classes.

Deep MLP

Pigment identification and spectral unmixing using deep MLP networks were exploited in a series of publications [155, 163, 164]. In [163], a multi-label classification (the number of labels equals the number of pure pigments) DNN with four fully connected layers of 256, 128, 64, and 32 neurons and an output layer of 16 neurons was developed. The sigmoid activation function was used for all layers, whose output is a vector of 0 and 1. The loss function employed in this model was the average binary cross-entropy loss. In [164] and [155], the authors expanded the first approach from pigment identification to pigment unmixing in the VIS energy range with the aid of a two-branch DNN model.

As schematically presented in Figure 4.4, the first branch of the NN performs pigment identification as a multi-label classification problem as in [163]. The second branch of the DNN model aims at performing spectral unmixing. It is a fully connected network also consisting of four layers of 256, 128, 64, and 32 neurons, with relu/sigmoid activation functions. The input layer represents a reflectance spectrum of 256 bands. The output layer contains 11 neurons, whose values correspond to pigment concentrations. To correctly predict the latter, the softmax activation function was employed in the output layer, such that the output values sum up to 1. Adam optimizer with default parameters is used. The batch size was set to 64. The training was performed for 200 epochs. In order to

prevent overfitting, early stopping with a patience of 10 was used. The model made use of late fusion, which was shown to give optimal results.
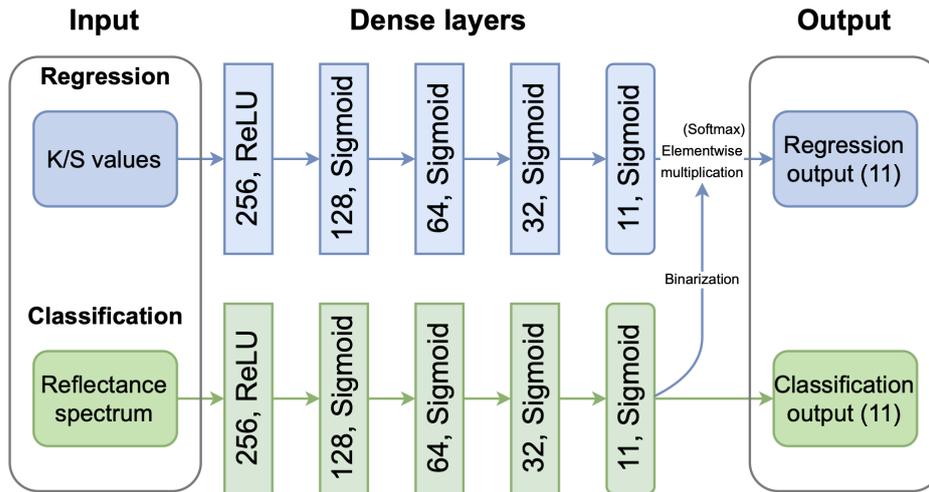
Fig. 4.4 Schematic representation of the two-branch MLP model used in [155, 164].

In [188], Striova et al. used a deep MLP composed of five layers to extrapolate NIR reflectance spectra from VIS data. The proposed NN takes visible reflectance spectra from 400 to 750 nm as input and predicts spectra in the NIR range from 1200 to 1700 nm. The input and output spectra both have 16 channels. The neural network is composed of four hidden layers of sizes 15, 15, 15, and 1 neuron(s). The sigmoid activation function was used in each layer, except for the last one where a linear activation function was employed [26]. The model was trained using the scaled conjugate gradient algorithm. The MLP-extrapolated NIR image contains information related to the VIS bands only. Therefore, in order to obtain the pure NIR spectra without the visible contribution, the computed extrapolated spectra were extracted from the measured NIR spectra, and the resulting intensity values were scaled between 0 and 1.

In [178], accurate painting reproduction via 3D printing was achieved by optimising the multi-layer composition of different inks with the aid of a bidirec-

tional model. The model consists of two fully connected neural networks. The first neural network has four hidden layers of 300 neurons each. The hidden layers as well as the final layer use the ReLU activation function. The loss function is the Euclidean distance between the prediction and the measurement (scaled by the square root of the number of wavelength bands/channels). The second neural network takes the 31-band reflectance spectrum as input and predicts the ink layout (an 11-dimensional vector). This neural network consists of eight hidden layers, each layer containing 160 neurons. The ReLU activation function is used for all hidden layers, and a softmax activation function is used in the output layer. This NN is deeper as it attempts to learn a more complicated dependence, namely the prediction of ink layout from reflectance spectra. Two additional loss functions were employed which were shown to produce more stable and consistent training results, namely perceptual loss and thickness loss. Finally, the backward function is trained to minimise the total loss, which is a sum of the spectral, perceptual and layer thickness loss functions. A soft-quantisation layer is applied at the end of the layout prediction layer in order to force the neural network to predict close-to-integer values for the number of ink layers.

4.6.3   Convolutional Neural Networks (CNNs)

A convolutional neural network (CNN) is a specialised type of deep neural networks that uses convolution operation in at least one of the hidden layers to down-sample the input data and extract features. A convolutional layer is made up of a set of kernels (the set is also known as a filter) which are matrices of weights, optimised during training. The filters convolve with the input to produce a feature map. In the case of CNNs the neurons are sparsely connected as opposed to the full connectivity in MLPs, which means each neuron only receives input from a restricted area around it. Frequently, one or more convolutional layers followed by a pooling layer form a convolutional block. There are usually multiple convolutional blocks followed by a fully connected block to produce the final outputs. Different types of blocks can be connected in various ways to form a customised architecture for the designed task [8, 191].

One-Dimensional-CNN

One-dimensional (1D)-CNN is a CNN that processes one-dimensional input data. Input data for CNNs are often two-dimensional, or an image, taking advantage of CNNs' capacity to learn the spatial correlations. However, in the CH literature, spectral input is commonly used, i.e., series of spectra serve as the inputs to the network without maintaining the spatial relationship.

Kleynhans et al. trained a typical 1D-CNN model for multi-class pigment classification based on VNIR data [101]. As summarised in Figure 4.5, the proposed model consists of one input layer, four hidden layers, and one output layer. The input layer has one dimension with 209 neurons, which is the wavelength number of the input spectra. The first two hidden layers are 1D convolutional layers with, respectively 64 and 32 filters and kernel sizes of 5×5 and 3×3. This is followed by a max pooling layer, which down-samples the previous hidden layer by 2 to reduce the dimensionality, keeping the maximum value of every second neuron. Then, two fully connected layers of sizes 100 and 25 form the last two layers. The output consists of 25 classes corresponding to the labels of the input spectra, including both pure pigments and pigment mixtures. The rectified linear unit (ReLU) activation function is used as activation function for each hidden layer, while softmax is used for the final layer, which gives a probability value between 0 and 1 for each class, summing up to 1. The final class is assigned by defining a threshold; at 0.99, a number of pixels remain unclassified. By decreasing the threshold from 0.99 to 0.85, part of the unclassified areas was assigned a class with rising in false positive identifications.

The performance of the model was measured with a categorical cross-entropy loss function, while the stochastic gradient descent optimizer was applied to minimise the loss function. For the learning process, the model was trained with batch sizes of 50 and epochs of 30, and evaluated on a validation set of 10% of the training data. The training started with a learning rate of 0.01, which was decreased if after four epochs (cycle through full training dataset) the validation loss did not decrease.
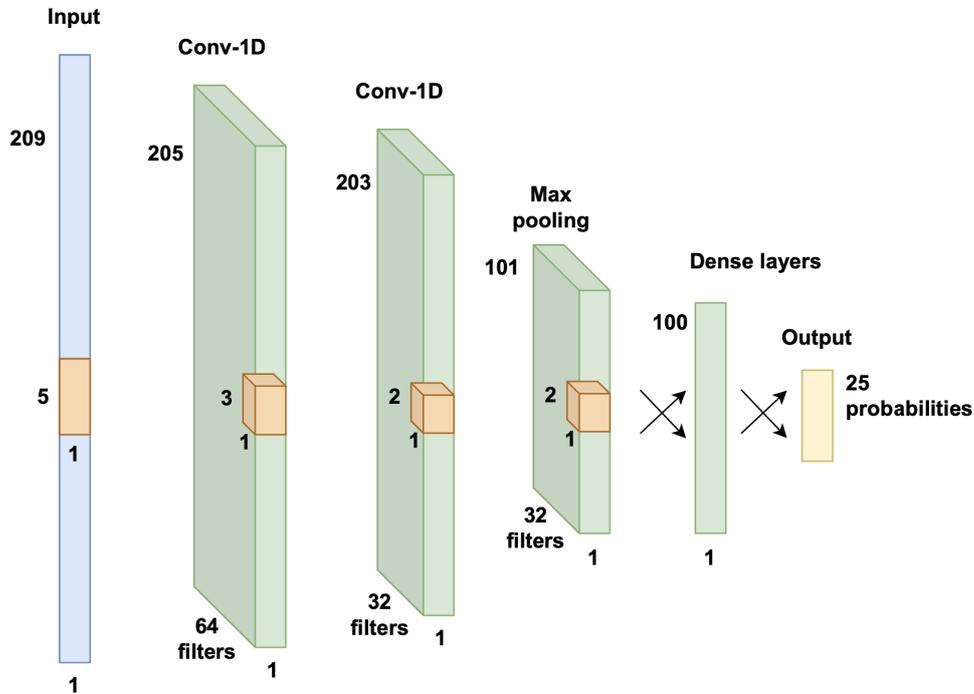
Fig. 4.5 Schematic 1D-CNN network architecture used in [101]. The input layer (shown in blue) takes as input the VNIR spectrum, followed by a series of hidden layers (shown in green) that perform feature extraction using convolutional kernels (shown in orange). The output layer (shown in yellow) produces the final classification results.

CNN with Transfer Learning

Transfer learning (TL) is a learning strategy in which a model trained on one set of data is reused as the starting point to train a new model on data with slightly different properties. Both models adopt the same architecture but possess different parameters. The low-level features learned by the pre-trained model are maintained when training a new model by freezing all parameters on certain layers, commonly the earliest layers which tend to learn more general structures from the input. The model, thus, only adjusts a smaller portion of parameters to adapt to the new data, making it possible to achieve high accuracy with much

less training data and computational resources. TL is especially useful when the available data is limited while the NN has a rather complex architecture, as the model can leverage the knowledge it gained from the larger dataset to the new task. In CH, two works have adopted the TL strategy to exploit large artificial datasets [21, 97].

In [97], a CNN model was applied to classify XRF spectra into pigment classes (15 classes were used in this study). The network was composed of three convolutional layers with max-pooling layers in between, and the output layer predicts the label for the corresponding pigment. The ReLU activation function was used for each layer except the final layer, which used the softmax activation function. The model is pre-trained on synthetic XRF data, and then fine-tuned with spectra acquired on real pigment samples, by freezing the weights in all layers in the network except the final layer. The performance of the model was evaluated before and after the fine-tuning, and the model with transfer learning showed higher accuracy.

A more complex CNN model was developed in [21] to identify layered pigments. The inputs were XRF spectra with 3815 channels. The model consists of five convolutional blocks, each made up of a 1D convolutional layer, a batch normalisation layer, and a max-pooling layer. The number and size of the kernels of each 1D convolutional layer were set at 64,64,64,64,128 and 5,3,3,3,3, respectively. The model is then followed by a post-convolutional layer with 128 kernels with a size of 3, a flatten layer, a dropout layer and a fully connected layer. The activation function LeakyReLU is used for every layer other than the final layer, while the output layer uses a sigmoid activation function. The output predictions were in 11 classes, outputting the probabilities of each class (each pigment) between 0 to 1. The loss was calculated to optimise the performance of the model by averaging the binary cross entropy of each predicted class.

The model was pre-trained with the artificial dataset using randomly initialised weights. It was further fine-tuned with data acquired on mock-ups, in two training steps: first, the pre-trained weights were fixed in all convolutional

layers, whereas only the fully connected layers were fine-tuned. Next, all layers were trainable and were further fine-tuned with the mock-up dataset.

Multi-Branches CNN

Two different types of complex CNN architectures have so far been employed to analyse hyperspectral data from painted artworks: the first one to perform sketch extraction [212], and the second one to identify pure pigments [43].

In the first case, an efficient edge-detection algorithm, consisting of a combination of two NNs, a bi-directional cascade network (BDCN) [87] and a U-net [167], was employed. BDCN consists of five Incremental Detection (ID) blocks, which contain convolutional layers and a Scale Enhancement Module (SEM). Each block is connected to the next one by a pooling layer. A carefully designed loss function making use of a distinction between edge and non-edge pixels was employed. To account for the imbalanced distribution of edge and non-edge pixels, a class-balanced cross-entropy loss was employed [212]. Due to insufficient training data, the authors resorted to transfer learning by pre-training the BDCN model on a publicly available dataset of natural scenes. The network parameters were then fine-tuned on a target cultural relics dataset.

In [43], Chen et al. achieved pure pigment identification with a multi-class classification problem using a three-branch deep-learning (DL) model. Unlike the 1D-CNN where a single spectrum is used as input, in this case, the first two branches of the NN take as input a small sub-cube around a central pixel, thus allowing for a combined spectral and spatial investigation. The first branch of the neural network consists of five sets of 3D convolutional filters with ReLU activation functions. It takes a 9×9 sub-cube of the hyperspectral cube around a central pixel as input. The architecture of the second branch is identical to that of the first one using the derivative of the same sub-cube as input.

The third branch of the model takes the reflectance spectrum of the central pixel as input and computes the spectral correlation map between the pixel and a reference pigment database. This result is fed to a shallow fully connected

feed-forward NN, which analyses the error between the reflectance spectrum of the pixel and the spectral signatures in the reference pigment database.

The outputs of the three branches are flattened, concatenated in a single vector, and fed into a fully connected feed-forward deep NN of five hidden layers with ReLU activation functions. One of these layers is a dropout layer, intended to minimise overfitting. The output layer consists of 17 neurons—16 for the pure pigments, and 1 neuron which predicts whether the reflectance spectrum belongs to a mixture or not. A threshold of 0.9 is applied to the output to discriminate between the spectra of pure pigments and pigment mixtures. The categorical cross-entropy loss function was employed. The NN parameters were optimised with the aid of the Adam optimizer with a learning rate of 0.0001. The training was performed for 20 epochs.

4.6.4 Encoder–Decoder (ENDEC)

An encoder–decoder (ENDEC) is a neural network model which consists of two parts—an encoder and a decoder. The two parts usually have inversed structures, with the encoder working in the normal direction that compresses the input data into a lower-dimensional representation (latent), and the decoder reconstructing the input from the latent, thus working in the opposite direction [104, 137]. The typical structure of an ENDEC model is illustrated in Figure 4.6. The hidden layers can be either fully connected, convolutional or even recurrent, based on the specific architecture [94]. The encoder and decoder are trained together by minimising the difference between the reconstructed outputs and the inputs. The training can be performed through either a supervised or unsupervised learning strategy as the ground truth is already provided in the input data; thus, in some cases, no explicit labels are needed.

A fully connected ENDEC was used in [77] in order to learn the dependence of reflectance spectra on pigment concentrations. The input and output of the ENDEC are reflectance spectra, whereas the intermediate layer represents pigment concentrations. The encoder part consists of eight layers of 300 neurons each, and an output layer of 19 neurons (for 19 pigments). The ReLU activation
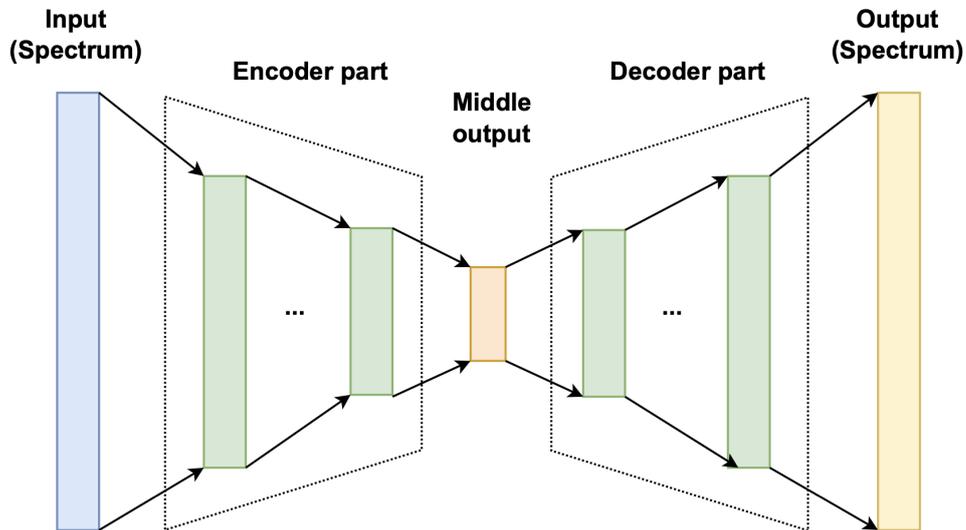
Fig. 4.6 Typical architecture of an encoder–decoder neural network. The input (shown in blue) is first compressed by the encoder part (shown in green) into a lower-dimensional representation (shown in orange). The decoder part (also marked in green) then performs up-sampling from the middle output and produces a reconstruct version of the input (shown in yellow).

function was used in all layers, except for the output layer, where the softmax activation function was employed. The mean absolute error (MAE) between the ground truth and the predicted concentrations was the loss function in the encoder part. The decoder part is composed of four hidden layers of 500 neurons each; the ReLU activation function was used for all layers. The mean squared error was the loss function of the decoder part. The Adam optimisation algorithm with default parameters was employed. In order to avoid overfitting, L2 regularisation was applied, with a weight decay value of 1e-5 . The learning was first performed only in the decoder unit, then the decoder weights were fixed, and the encoder part was trained.

Another type of encoder–decoder architecture, based on CNN and known as U-net [167], was used in [212] to refine and denoise sketches extracted by

a BDCN. The encoder NN consisted of four blocks each composed of three convolutional layers and a pooling layer. The decoder part of U-net consisted of four blocks composed of three convolutional layers. Weight decay regularisation was employed to reduce overfitting.

An unsupervised autoencoder model is proposed in [180] for paint layer thickness and pigments mixing ratio estimation. It is a type of ENDEC network that utilises unlabelled data for training, with the goal to learn a compact representation of the input data (middle output). Since the ground truth of the reconstructed spectrum is the input data itself, the model is also considered self-supervised. In this case, the input of the encoder is spectral data of layered surface objects and is decomposed into latent variables in the middle layer. The decoder part is based on the Kubelka–Munk theory; as such, the latent variables are physically interpretable as pigment thickness and pigment mixing ratio. The encoder consists of seven fully connected layers with ReLU activation function, and the number of neurons in each layer is 300. For the middle output layer, a ReLU function is used to force the thickness to be positive and a softmax function is used for pigment mixing ratios. A special loss function is designed, and the Adam algorithm is used for optimisation. The training was repeated in three cases with different types of inputs: artificial data, data measured on mock-ups of tomb mural pigments, and that of watercolour pigments.

A variation autoencoder (VAE)-based network was applied in [190] for recovering large areas of scratch damage in mural paintings. The model adopted is pre-trained in a previous study [205] designed for old photo restoration; thus, no training process is involved and the murals serve as test cases to the pre-trained model. Since the VAEs are able to learn compact representations of input data (encoder part) and generate new data following the probabilistic distribution (the decoder part is also known as the generator), the model is capable of repairing the lost contents in murals that suffer from similar damages of photos. The input to the NN is a true colour image synthesised from VNIR datacube, which is pre-processed and enhanced to assist the restoration. The model consists of two identical VAEs and an adversarial mapping network (GAN). The two VAEs take

the damaged images and the related ground truth images as inputs, respectively, where the encoder transforms the input into a compressed representation in latent space (middle output), and the decoder part generates a reconstructed version of the input image. Then, the middle outputs of the first VAE (damaged images) are translated into that of the repaired ones of the second VAE through the mapping network. The encoder part and decoder part are both composed of three convolutional/deconvolutional layers and four residual blocks, and the mapping network has six convolutional layers, six residual blocks, and one partial non-local block.

4.6.5 Deep Belief Network (DBN)

Deep belief networks (DBNs) are a type of NN that consist of multiple layers. In a DBN, the lower layers consist of the so-called restricted Boltzmann machines [73, 211], typically used for unsupervised learning and designed to learn the underlying probability distribution of a dataset, and the top layer is usually fully connected. The training process of DBN is different from MLP and CNN, which involves pre-training the lower layers using unsupervised learning and then fine-tuning the top layers using supervised learning through backpropagation [120]. DBN can be used to solve unsupervised learning tasks to reduce the dimensionality of features, as well as for supervised learning tasks to build classification or regression models.

In [189], a DBN-initialised neural network was used to predict the degree of flaking. The NN is composed of two hidden layers, with the number of neurons configured as 200 and 150, respectively. The weights of the DBNs are pre-trained layer-by-layer, from the first hidden layer to the outermost layer; the learning rate was set to 0.01 for 100 epochs, and the momentum was set to 0.4. No information on activation functions and output layer is provided.

4.6.6 Model Performance Evaluation

The testing of the generated output on an unseen dataset is accompanied by different evaluation criteria. For the test sets, the ground truth is already known by the methods described in Section 4.2 and the quantitative evaluation of the prediction accuracy is performed [21, 77, 80, 101, 118, 163, 189, 212]. As summarised in table 4.1, it is measured by calculating sensitivity, overall accuracy (OA), average accuracy (AA), standard deviation (STD), R-squared coefficient of determination, root mean square error (RMSE), or using the individual per-class accuracy. For those tested on a new set of historical paintings, the performance is evaluated more qualitatively. The output results of the networks are commonly cross-validated with other analytical instruments, involving the manual check of spectral features of selected points using XRF, FORS, FTIR or micro-Raman, and optical microscopy [43, 97, 102, 155, 163, 188].

The performance of the NNs is also frequently compared with other methods and algorithms. In [101], it is compared with the conventional two-step approaches that first perform a spectral clustering and then label the present pigments based on additional information. In addition, the performance of the NN models are particularly compared with other linear unmixing algorithms, including support vector machine (SVM), spectral angle mapper (SAM), and fully constrained least squares (FCLS) [43, 101, 155, 163]. In [101], the 1D-CNN is proven to outperform the SAM by 2.1% and the SVM by 6.4%, respectively, in identification accuracy. In [155], the DNN classification model outperformed SAM by 5.1% in accuracy, with the latter performing poorly in identifying pigments present in mixtures or from areas where the endmembers are not priorly defined. FCLS is tested in [163] for decomposing the given spectrum to a linear combination of pure pigment spectra; however, the accuracy of estimated mixing coefficients largely underperformed the DNN model proposed. SAM and FCLS are also tested in [43] for identifying pure pigment areas by setting a threshold, while the performances are far from the multi-branch CNN model developed.

In some cases, the performances of different NN architectures are also evaluated and compared. The CNN-based model is compared with the unsupervised

Metrics Type	Formula	Ref.
Sensitivity	$\text{Sensitivity} = \dfrac{\text{True positive}}{\text{True positive} + \text{False negative}}$	[21]
Overall accuracy (OA)	$OA = \dfrac{\text{No. of correctly classified examples}}{\text{Total No. of examples}}$	[21, 80, 101, 118]
Average accuracy (AA)	$AA = \dfrac{1}{M}\sum_{i=1}^{m} OA_i,$ $M - $ number of classes, $OA_i - $ overall per class accuracy	[80]
Standard deviation (STD)	$STD = \sqrt{\dfrac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2},$ N-total No. of examples, $\mu = \dfrac{1}{N}\sum_{j=1}^{N} x_j$	[80]
Coefficient of determination (R^2)	$R^2 = 1 - \dfrac{RSS}{TSS}$ $RSS = \Sigma_i \left(y_i - y_i^{pred}\right)^2$ -sum of squared residuals $TSS = \sum_i (y_i - \bar{y})^2$ -total sum of squares	[189]
Root mean square error (RMSE)	$RMSE = \sqrt{\dfrac{\sum_{i=1}^{N}\left(y_i - y_i^{\text{pred}}\right)^2}{N}}$	[77, 163, 189, 212]

Table 4.1 Most commonly encountered metrics for the quantitative evaluation of prediction accuracy in CH.

SOM in [61], the 1D-CNN model is compared with MLP and outperforms the latter in [101], and the unsupervised autoencoder is tested to be superior to the supervised model in [180]. Moreover, in developing the algorithms, NN models with different algorithmic choices are tested and evaluated. The DBN is compared with an artificial NN-based approach in [189]; two MLPs performing multi-class and multi-label classification, respectively, are compared in [163]; and four DNN architectures with different fusion stages are tested in [164].

4.7   Discussions and Conclusion

The promising adoption of NNs within CH is evident throughout the literature. The NN models that learn and adapt on their own are relatively easy to build and, once the model is properly trained, the application to real problems is very fast, i.e., it takes only tens of seconds.

As summarised in Table 4.2, we have demonstrated the wide applications of NNs in solving paint component mapping problems, from pure pigments to complex mixtures, from classification to concentration estimation, and from multi-class to multi-layer systems. Different architectures were chosen according to multiple factors, from fully connected networks to the more advanced convolutional networks or encoder–decoders, from shallow single-layer networks to complex multi-branch NN models, taking unsupervised learning to supervised or transfer learning strategies. Furthermore, the training datasets also present a large complexity, varying in spectral ranges, input data types, and preparation methods, from artificial data to historical data, and from modern model paintings to historical artworks of different origins.

Model	Spectral Input	Input Type *	Dataset Size	Task	Application	Ref.
SOM	XRF	III	41,327	Clustering	Peruvian watercolour painting (c.1860)	[102]
	RIS (SWIR)	II, III	NA	Clustering	Excavated bronze fragments (B.C.)	[117]
MLP	RIS (VNIR)	III	NA	Classification	Mural paintings (17th–20th century)	[118]
	RIS (VNIR)	I	143	Classification	II, illuminated folio (15th century)	[163]
	RIS (VNIR)	I	110	Unmixing	II	[164]
	RIS (SWIR)	II, III	12	Unmixing	Tibetan thangka (19th century)	[155]
	RIS (SWIR)	II	NA	Classification	II	[80]
CNN	XRF	I	16,224	Unmixing	II, impressionistic paintings (c.1900)	[21]
		II	1320			
	XRF	I	3000	Classification	Early Renaissance painting (1468)	[97]

Model	Spectral Input	Input Type *	Dataset Size	Task	Application	Ref.
		II	75			
	RIS (VIS)	I	21,24	Classification	Late Portuguese paintings (c.1910)	[43]
	RIS (VNIR)	III	16,683	Classification	Illuminated manuscript (c.1340)	[101]
ENDEC	RIS (VIS)	I	1,445,136	Unmixing	II	[77]
	RIS (VIS) **	I	35,7	Unmixing	II	[180]
		II	213		Decorated tomb murals	
	RIS (VIS)	II	171,6		Watercolours	
	RIS (VNIR)	III	NA ***	Restoration	Mural paintings (c.1392)	[190]

Model	Spectral Input	Input Type *	Dataset Size	Task	Application	Ref.
CNN, EN-DEC	RIS (VNIR)	III ****	500 public images and 41 relics images	Edge detection	Painted cultural relics	[212]
DBN	RIS (NIR)	III	12	Classification	Mural paintings (7th–9th century)	[189]

* Input type I, II, and III refers to artificial data, modern data, and historical data described in Section 4.5. ** The three different input sets serve as separate training cases in ref.[180] . *** Ref.[190] used a pre-trained model developed by a previous study; thus, the application does not include the training process. **** Ref.[212] used a publicly available dataset containing 500 images to pre-train their model, which is the only case and does not belong to our defined categories.

Table 4.2 A summary of the neural network models analysed in this chapter. The distinct characteristics of the NN models are presented, such as the spectral input utilised, the type and size of the dataset employed, the assigned task, and their applications to historical objects.

Over the last few years, various algorithms were developed with the aim of capturing complex features from HSI data for an in-depth analysis of the rich spectral features [114]. However, as observed in the remote sensing field, the scale of the HSI training data is limited due to the cost, complexity, and labelling constraints in their development [176]. That, in turn, causes a sub-optimal learning of NN with large numbers of parameters. The complexity of the architecture and the variability of the data is often unbalanced, which frequently causes poor applicability to real-world problems. In this circumstance, a specific discussion, concerning the datasets and the NNs architectures, is addressed, for a more systematic use of NN for HSI in CH.

Supervised NN models represent a major part of the models developed for classification or unmixing purposes (85.7%). However, due to the costly and labour-intense labelling of HSI data, the supervised models suffer from very limited labelled input datasets that constrain the effective learning of the massive spectral parameters required for accurate classification or regression results [30, 129]. Extensive use of those models will soon represent a bottleneck that will significantly inhibit their practical application for wider HSI data analysis. Building new HSI benchmark datasets to meet the training needs of the NN model is expensive and time-consuming [138]. Several HSI datasets have been built over the last few years for diverse HSI applications, and multiple open-source databases of diffuse reflectance references flourished simultaneously [25, 34, 54].

However, the currently labelled HSI samples are still far below the demands of the current NN models, as well as verifying the performance of evolving algorithms. Limited training samples cause model overfitting during training which significantly affects the model performance. Overfitting can be desirable when the application is specific to a set of data that is mostly similar to the training set, as the model can make very accurate predictions. However, in most cases, it is good practice to avoid overfitting to increase the applicability of the model and make it more resistant to noise and other types of variations. The generation of artificial training data, simulated through theoretical approximations, has been

proposed to help generate new training samples to mitigate the cost of preparing multiple modern replicas [21, 43, 77, 97, 163, 164, 180]. The available training data size is largely expanded: the averaged data sizes for modern replicas and historical data are around 80 k and 12 k, respectively, while the artificial sets generated have on average 250 k spectra (Table 4.2). Similarly, several works address this issue by expanding the dataset through an augmented dataset approach [43], and a mixed use of various dataset types (artificial, modern, historical) [21, 97, 117, 155].

In CH, a general trend has been observed with NN algorithms becoming more powerful and widely applied, leading to increasingly complex architectures to tackle more difficult tasks. This trend is driven by the increasing complexity of data and the improvement in computing power. However, it is important to balance the complexity of the network with the size of the dataset in order to achieve the best performance and efficiency. For large and complex datasets, a more advanced network with more hidden layers and complex structures is preferred, while a simpler network is sufficient for smaller datasets and simpler tasks. This practice is observed in the CH literature where relatively simpler models (SOM and MLP) have an averaged training set size of around 80 k spectra, while for more complex architectures (CNN, ENDEC and DBN), the averaged size is 215 k spectra (Table 4.2). For the NNs of the same category, refs. [97] both adopted a CNN architecture and transfer learning strategy, respectively, while using the same input range and data type. As the model used in [21] is much more complex than in [97], the size of the training data is also significantly increased. It is generally recommended to start with a simple network configuration and gradually increase complexity to find the optimal architecture for a given dataset, and to reach the optimal model performance in terms of accuracy and computing power.

In conclusion, spectral imaging instruments have become a common tool in CH studies accompanied with a natural extension of dataset and data range. The use of neural networks (NNs) is likely to become increasingly important for assisting the interpretation and information extraction of hyperspectral dataset.

As such, the approach will soon grow and provide more examples, available to the community. In this context potential research axes arise from this chapter. Many researchers have indeed recognized the limitation of current practices in terms of poor generalisation ability and applicability to other datasets, highlighting the need for new, larger and more complex reference hyperspectral datasets. One way to address this issue is to continue to increase the dataset by including a wider range of data from larger pigment databases and more variable painting structures. In order to facilitate this process and increase model transferability, a move towards open-source algorithms and publicly available datasets is indispensable. Similarly, the combination of NN models with complex physical models (four-flux KM approximation [18, 200], or the radiation transfer equation [203, 200]) needs to be explored, providing larger and more diverse artificial datasets. In parallel to data augmentation and synthetic data generation that has started to be addressed in the field, neural network models that can efficiently utilise a limited number of labelled samples have to be tested. As such, transfer learning or weakly supervised methods should be further explored to mitigate the demand for training samples [204].

Chapter 5

Conclusions

Cultural heritage conservation and restoration stand at the crossroads of art, history, science, and technological progress. Traditional conservation and restoration methods, while invaluable, are often constrained by their ability to adapt to the complexity of cultural heritage artefacts. In the contemporary landscape, analytical instruments and computational technology are profoundly shaping the interdisciplinary field. Machine learning and data analysis, with their capacity to unravel intricate patterns and trends within extensive datasets, offer a promising avenue for enhancing conservation and restoration practices. In this thesis, we have focused on *how machine learning and data analysis techniques can contribute to the conservation and restoration of cultural heritage objects*. Through various case studies presented, we have investigated the potential and limitations of machine-learning techniques in processing extensive hyperspectral data, in supporting tasks ranging from material diagnostics and classification to mapping and digital restoration. This chapter summarises the contributions of this thesis, highlighting key findings for each research question we presented in Chapter 1.

RQ-1.1 *Can machine learning techniques address complex identification and classification problems in CH?*

This research question is investigated through a case study faced in Chapter 2 and reviewed through the literature in Chapter 4. For unsupervised algo-

rithms, as presented in Chapter 2, the application of joined spatial-spectral clustering algorithms to address complex identification and classification problems is demonstrated. The developed data processing pipeline has exhibited great adaptability to the Cultural Heritage context: the SLIC superpixel algorithm is especially suitable for pointillism paintings for maintaining the important spatial details while enhancing the processing efficiency, and the soft clustering algorithms excel in extracting patterns from complex overlapping data. These algorithms succeed in distinguishing various inorganic pigments in artworks, identifying complex colourant mixtures, as well as localising degradation. However, while these techniques have proven invaluable in uncovering nuanced details that humans cannot determine from vast amounts of data, human interpretation remains indispensable for a meaningful result. The fine-tuning of hyperparameters to achieve a reliable result is highly dependent on human expertise to decompose the chemical meaning behind the subtle spectral variations captured in the clusters.

On the other hand, the application of supervised approaches, such as those discussed in the neural network literature in Chapter 4, involves models learning from labelled training data. The wide adaptability of Neural Networks (NNs) in solving paint component mapping problems is demonstrated, from pure pigments to complex mixtures, from classification to quantitative concentration estimation, and from multi-class to multi-layer systems. These supervised models, once properly trained, offer fast and efficient applications to real problems. However, the success of supervised approaches relies on the accessibility of accurately labelled training data, making the preparation of datasets a critical aspect. In the CH field, where the availability of large labelled datasets is often limited, human input and expertise are crucial in the creation of these labelled datasets, by generating synthetic data from physical models, preparing mock-ups, or hand labelling the data from historical artworks. Therefore, both unsupervised and supervised approaches highlight the interconnected

roles of automated techniques and human efforts, requiring a balanced and collaborative approach to effective cultural heritage conservation practices.

RQ-1.2 *To what extent can machine learning models detect and mitigate the degradation of cultural heritage artefacts over time?*

This research question is explored through case studies in Chapter 2 and Chapter 3. In Chapter 2, the developed approach has succeeded in capturing and localising the degradation of yellow pigments in Signac's painting, suggesting chromium-based pigments rather than cadmium, providing valuable insights into the artist's palette. Furthermore, the algorithm is employed to segment degraded film into differential fading areas. By revealing subtle patterns indicative of degradation, the machine learning algorithms not only provide a nuanced perspective on the degradation process but also lay a foundation for future restoration works.

Chapter 3 delves deeper into the potential of machine learning to address degradation issues, focusing on the digital restoration of degraded cinematic films. The proposed vector quantization algorithm, constructing a codebook that stores the paired spectral signatures between degraded and non-degraded references, achieves subjective restoration by capturing minimal yet essential differences among pixels. The positive results obtained highlight the potential of machine learning in mitigating the effects of degradation, particularly in scenarios with inhomogeneous fading patterns, a challenge hard to treat with conventional methods. However, while this automated technique minimises human intervention in the restoration process, the accuracy and reliability of restoration results depend crucially on the availability of a ground-truth reference. Furthermore, the high computational load arising from the high-resolution spectral data may potentially limit its application to large-scale film stock.

RQ-1.3 *How transferrable are the machine learning models developed to similar scenarios in cultural heritage conservation and what future applications can be envisioned?*

The adaptability and transferability of machine learning algorithms, as an essential aspect of the technique, are discussed throughout the thesis. In Chapter 2, the adaptability of the proposed method is demonstrated through the two distinct cases. The painting dataset exhibits a larger variability of distinctive spectral characteristics rising from the unique chemical compositions of each pigment, while the film dataset is smaller in size with all pixels sharing similar chemical constituents and peak positions, varying mainly in subtle degradation loss. Nevertheless, the same methodology has achieved remarkable results in both cases. In particular, the Gaussian Mixture models built through the current training samples are directly applicable to the analysis of artworks with similar palettes, material compositions or degradation types. By expanding the included spectral features, these models can find wide and efficient applications in numerous similar scenarios.

In Chapter 3, the proposed approach is initially applied to a single frame and subsequently extended for the collective restoration of entire datasets. This extension involves constructing a multi-codebook that incorporates spectral features from multiple frames and exhibits promising potential for the restoration of deteriorated films of the same type. Additionally, the spectra bank collected in the codebook could be expanded to encompass various types of samples and degradation effects, enhancing its applicability to a broader spectrum of damaged films.

In Chapter 4, the superior adaptability of Neural Networks is demonstrated through their wide application with various architectures to tackle more difficult tasks. However, due to the labour-intensive labelling of hyperspectral data, the supervised models suffer from limited labelled input datasets for the accurate updating of the massive parameters. Limited training samples frequently lead to overfitting during training and poor generalisation ability to other datasets. While overfitting can be beneficial for specific cases, balancing the complexity of the architecture and dataset variability is crucial to achieving a robust solution for real-world appli-

cability. This limitation highlights the need for larger and more diverse reference hyperspectral datasets, such as pigment databases, extended spectral ranges, and more variable painting structures. To facilitate this process and increase model transferability, a move towards open-source algorithms and publicly available datasets is indispensable.

For future research directions, numerous aspects within the application of machine learning techniques in cultural heritage context merit exploration. While this thesis has been focused on Hyperspectral data, other types of data, such as macro-XRF and Raman spectroscopy, and the integration of multimodal data sources still need investigation. Dealing with the data challenge offers an opportunity to delve into semi-supervised learning approaches to reduce the dependency on large labelled datasets. Concerning more complex conservation problems, predictive modelling emerges as a potent avenue, considering also the temporal dimension in the degradation pathway. Additionally, there is a critical need to enhance the explainability of machine learning models in the context of cultural heritage conservation, facilitating better collaboration between machine learning algorithms and conservation experts.

References

[1] Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., and Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11):e00938.

[2] Aceto, M., Agostino, A., Fenoglio, G., Idone, A., Gulmini, M., Picollo, M., Ricciardi, P., and Delaney, J. K. (2014). Characterisation of colourants on illuminated manuscripts by portable fibre optic uv-visible-nir reflectance spectrophotometry. *Analytical methods*, 6:1488–1500.

[3] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). Slic superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11):2274–2282.

[4] Adil, M., Ullah, R., Noor, S., and Gohar, N. (2022). Effect of number of neurons and layers in an artificial neural network for generalized concrete mix design. *Neural Computing and Applications*, 34(11):8355–8363.

[5] Albawi, S., Mohammed, T. A., and Al-Zawi, S. (2017). Understanding of a convolutional neural network. pages 1–6.

[6] Alfeld, M. and de Viguerie, L. (2017). Recent developments in spectroscopic imaging techniques for historical paintings-a review. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 136:81–105.

[7] Alfeld, M., Pedetti, S., Martinez, P., and Walter, P. (2018). 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*, 19(7):625–635.

[8] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., and Farhan, L. (2021). Review of deep learning: concepts, cnn architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1):53.

[9] Arai, Y., Nakauchi, S., and Usui, S. (1996). Color correction method based on the spectral reflectance estimation using a neural network. volume 1996, page 5–9. Society for Imaging Science and Technology. issue: 1.

[10] Arthur, D. and Vassilvitskii, S. (2007). K-means++ the advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1027–1035.

[11] Artioli, G. (2010). *Scientific methods and cultural heritage: an introduction to the application of materials science to archaeometry and conservation science*. OUP Oxford.

[12] Ayres, L. B., Gomez, F. J., Linton, J. R., Silva, M. F., and Garcia, C. D. (2021). Taking the leap between analytical chemistry and artificial intelligence: A tutorial review. *Analytica Chimica Acta*, 1161:338403.

[13] Bacci, M. and Picollo, M. (1996). Non-destructive spectroscopic detection of cobalt (ii) in paintings and glass. *Studies in conservation*, 41(3):136–144.

[14] Bai, D., Messinger, D. W., and Howell, D. (2017). A pigment analysis tool for hyperspectral images of cultural heritage artifacts. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXIII*, volume 10198, pages 429–443. SPIE.

[15] Bai, D., Messinger, D. W., and Howell, D. (2019). A hyperspectral imaging spectral unmixing and classification approach to pigment mapping in the gough selden maps. *Journal of the American Institute for Conservation*, 58(1-2):69–89.

[16] Barnes, N. F. (1939). Color characteristics of artists' pigments. *J.O.S.A*, 29.

[17] Barni, M., Pelagotti, A., and Piva, A. (2005). Image processing for the analysis and conservation of paintings: Opportunities and challenges. *IEEE Signal Processing Magazine*, 22(5):141–144.

[18] Barrios, D., Vergaz, R., Sánchez-Pena, J. M., García-Cámara, B., Granqvist, C. G., and Niklasson, G. A. (2015). Simulation of the thickness dependence of the optical properties of suspended particle devices. *Solar Energy Materials and Solar Cells*, 143:613–622.

[19] Berns, R. S. and Mohammadi, M. (2007). Evaluating single- and two-constant kubelka-munk turbid media theory for instrumental-based inpainting. *Studies in Conservation*, 52(4):299–314. publisher: Routledge $_eprint : https : //doi.org/10.1179/sic.2007.52.4.299$.

[20] Bertrand, L., Thoury, M., Gueriau, P., Anheim, É., and Cohen, S. (2021). Deciphering the chemistry of cultural heritage: targeting material properties by coupling spectral imaging with image analysis. *Accounts of Chemical Research*, 54(13):2823–2832.

[21] Bingjie, Xu, Wu, Y., Hao, P., Vermeulen, M., McGeachy, A., Smith, K., Eremin, K., Rayner, G., Verri, G., Willomitzer, F., Alfeld, M., Tumblin, J., Katsaggelos, A., and Walton, M. (2022). Can deep learning assist automatic identification of layered pigments from xrf data? arXiv:2207.12651 [cs, eess].

[22] Biron, C., Mounier, A., Le Bourdon, G., Servant, L., Chapoulie, R., and Daniel, F. (2020). Revealing the colours of ukiyo-e prints by short wave infrared range hyperspectral imaging (SWIR). *Microchemical Journal*, 155:104782.

[23] Bitossi, G., Giorgi, R., Mauro, M., Salvadori, B., and Dei, L. (2005). Spectroscopic techniques in cultural heritage conservation: a survey. *Applied Spectroscopy Reviews*, 40(3):187–228.

[24] Blackmagic Design (2021). Davinci resolve 17. https://www.blackmagicdesign.com/products/davinciresolve/color.

[25] Blažek, J., Soukup, J., Zitová, B., Flusser, J., Hradilová, J., Hradil, D., and Tichý, T. (2014). M3art: A database of models of canvas paintings. Lecture Notes in Computer Science, pages 176–185, Cham. Springer International Publishing.

[26] Blažek, J., Striová, J., Fontana, R., and Zitová, B. (2017). Improvement of the visibility of concealed features in artwork nir reflectograms by information separation. *Digital Signal Processing*, 60:140–151.

[27] Borg, B., Dunn, M., Ang, A., and Villis, C. (2020). The application of state-of-the-art technologies to support artwork conservation: Literature review. *Journal of Cultural Heritage*, 44:239–259.

[28] Borzov, S. and Potaturkin, O. (2018). Spectral-spatial methods for hyperspectral image classification. review. *Optoelectronics, instrumentation and data processing*, 54:582–599.

[29] Boutarfass, S. and Besserer, B. (2020). Improving CNN-based colorization of B amp;W photographs. In *2020 IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS)*, pages 96–101.

[30] Brigato, L. and Iocchi, L. (2021). A close look at deep learning with small data. pages 2490–2497. ISSN: 1051-4651.

[31] Brunetti, B., Miliani, C., Rosi, F., Doherty, B., Monico, L., Romani, A., and Sgamellotti, A. (2017). Non-invasive investigations of paintings by portable instrumentation: the molab experience. *Analytical Chemistry for Cultural Heritage*, pages 41–75.

[32] Burton, D., Shore, J., and Buck, J. (1983). A generalization of isolated word recognition using vector quantization. In *ICASSP '83. IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 8, pages 1021–1024.

[33] Cachin, F. (2000). *Signac, catalogue raisonné de l'oeuvre peint*. Gallimard.

[34] Caggiani, M. C., Cosentino, A., and Mangone, A. (2016). Pigments checker version 3.0, a handy set for conservation scientists: A free online raman spectra database. *Microchemical Journal*, 129:123–132.

[35] Calatroni, L., d'Autume, M., Hocking, R., Panayotova, S., Parisotto, S., Ricciardi, P., and Schönlieb, C.-B. (2018). Unveiling the invisible: Mathematical methods for restoring and interpreting illuminated manuscripts. *Heritage Science*, 6(1):56.

[36] Capobianco, G., Bracciale, M. P., Sali, D., Sbardella, F., Belloni, P., Bonifazi, G., Serranti, S., Santarelli, M. L., and Cestelli Guidi, M. (2017). Chemometrics approach to FT-IR hyperspectral imaging analysis of degradation products in artwork cross-section. *Microchemical Journal*, 132:69–76.

[37] Casadio, F., Bezúr, A., Fiedler, I., Muir, K., Trad, T., and Maccagnola, S. (2012). Pablo picasso to jasper johns: a raman study of cobalt-based synthetic inorganic pigments. *Journal of Raman Spectroscopy*, 43(11):1761–1771.

[38] Casini, A., Cucci, C., Picollo, M., Stefani, L., and Vitorino, T. (2015). Creation of a hyper-spetral imaging reference database of red lake pigments. *COSCH e-bulletin*.

[39] Catelli, E., Randeberg, L. L., Alsberg, B. K., Gebremariam, K. F., and Bracci, S. (2017). An explorative chemometric approach applied to hyperspectral images for the study of illuminated manuscripts. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 177:69–78.

[40] Chambah, M. and Besserer, B. (2000). Digital color restoration of faded motion pictures. In *Computer Graphics and Image Processing Conference CGIP2000, Saint-Etienne, France*, pages 338–342.

[41] Chambah, M., Rizzi, A., Gatta, C., Besserer, B., and Marini, D. (2003). Perceptual approach for unsupervised digital color restoration of cinematographic archives. In *Color Imaging VIII: Processing, Hardcopy, and Applications*, volume 5008, pages 138–149. SPIE.

[42] Chane, C. S., Thoury, M., Tournié, A., and Echard, J.-P. (2015). Implementation of a neural network for multispectral luminescence imaging of lake pigment paints. *Applied Spectroscopy*, 69(4):430–441.

[43] Chen, A., Jesus, R., and Vilarigues, M. (2021). Convolutional neural network-based pure paint pigment identification using hyperspectral images. pages 1–7, Gold Coast Australia. ACM. [Online; accessed 2022-10-20].

[44] Chen, Y., Liu, X., Lyu, S., Wu, W., and Wang, R. (2022). Method of hidden strip information extraction from hyperspectral images of ancient paintings. *Sensors and Materials*, 34(12):4463.

[45] Chen, Y., Luo, Y., Ding, Y., and Yu, B. (2018). Automatic Colorization of Images from Chinese Black and White Films Based on CNN. In *2018 International Conference on Audio, Language and Image Processing (ICALIP)*, pages 97–102.

[46] Chevreul, E. (1839). *De la loi du contraste simultané des couleurs et de l'assortiment des objets colorés considérés d'après cette loi dans ses rapports avec la peinture, les tapisseries...* Pitois-Levrault et Cie, Paris.

[47] Chiu, S. L. (1994). Fuzzy model identification based on cluster estimation. *Journal of Intelligent & fuzzy systems*, 2(3):267–278.

[48] Ciortan, I. M., Poulsson, T. G., George, S., and Hardeberg, J. Y. (2023). Tensor decomposition for painting analysis. part 2: spatio-temporal simulation. *Heritage Science*, 11(1):84.

[49] Clarivate, PLC (2022). Web of science. https://www.webofscience.com/wos/. Accessed: 2022-10-28.

[50] Corbeil, M.-C., Charland, J.-P., and Moffatt, E. A. (2002). The characterization of cobalt violet pigments. *Studies in Conservation*, 47(4):237–249.

[51] Corel Paintshop Pro (2021). How to restore photos. https://www.paintshoppro.com/en/tips/corrections/restore-photos/.

[52] Cornelis, B., Ružić, T., Gezels, E., Dooms, A., Pižurica, A., Platiša, L., Cornelis, J., Martens, M., De Mey, M., and Daubechies, I. (2013). Crack detection and inpainting for virtual restoration of paintings: The case of the Ghent Altarpiece. *Signal Processing*, 93(3):605–619.

[53] Cortes, A. R. (2003). *Multispectral Analysis and spectral Reflectance Reconstruction of Art Paintings*. PhD thesis. [Online; accessed 2022-11-15].

[54] Cosentino, A. (2014). Fors spectral database of historical pigments in different binders. *E Conserv. J*, 2:54–65.

[55] Costanzo, A., Ebolese, D., Ruffolo, S. A., Falcone, S., la Piana, C., La Russa, M. F., Musacchio, M., and Buongiorno, M. F. (2021). Detection of the tio2 concentration in the protective coatings for the cultural heritage by means of hyperspectral data. *Sustainability*, 13(1):92. number: 1 publisher: Multidisciplinary Digital Publishing Institute.

[56] Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., and Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, 35(1):53–65. event-title: IEEE Signal Processing Magazine.

[57] Cucci, C. and Casini, A. (2019). Chapter 3.8 - hyperspectral imaging for artworks investigation. In Amigo, J. M., editor, *Data Handling in Science and Technology*, volume 32 of *Hyperspectral Imaging*, pages 583–604. Elsevier. DOI: 10.1016/B978-0-444-63977-6.00023-7.

[58] Cucci, C., Delaney, J. K., and Picollo, M. (2016). Reflectance hyperspectral imaging for investigation of works of art: old master paintings and illuminated manuscripts. *Accounts of chemical research*, 49(10):2070–2079.

[59] Cucci, C., Picollo, M., Chiarantini, L., Uda, G., Fiori, L., De Nigris, B., and Osanna, M. (2020). Remote-sensing hyperspectral imaging for applications in archaeological areas: Non-invasive investigations on wall paintings and on mural inscriptions in the pompeii site. *Microchemical Journal*, 158:105082.

[60] Daniel, F., Mounier, A., Pérez-Arantegui, J., Pardos, C., Prieto-Taboada, N., de Vallejuelo, S. F.-O., and Castro, K. (2016). Hyperpsectral imaging applied to the analysis of goya paintings in the museum of zaragoza (spain). *Microchemical Journal*, 126:113–120.

[61] Daveri, A., Paziani, S., Marmion, M., Harju, H., Vidman, A., Azzarelli, M., and Vagnini, M. (2018). New perspectives in the non-invasive, in situ identification of painting materials: The advanced MWIR hyperspectral imaging. *TrAC Trends in Analytical Chemistry*, 98:143–148.

[62] Deborah, H., Ulfarsson, M. O., and Sigurdsson, J. (2021). Fully constrained least squares linear spectral unmixing of the scream (verso, 1893). In *2021 11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–5. IEEE.

[63] Delaney, J. K., Zeibel, J. G., Thoury, M., Littleton, R., Palmer, M., Morales, K. M., de La Rie, E. R., and Hoenigswald, A. (2010). Visible and infrared imaging spectroscopy of picasso's harlequin musician: mapping and identification of artist materials in situ. *Applied spectroscopy*, 64(6):584–594.

[64] Diaz, G. I., Fokoue-Nkoutche, A., Nannicini, G., and Samulowitz, H. (2017). An effective algorithm for hyperparameter optimization of neural networks. *IBM Journal of Research and Development*, 61(4/5):9:1–9:11. event-title: IBM Journal of Research and Development.

[65] Dolenec, R., Rogelj, L., Stergar, J., Milanic, M., and Milanic, M. (2019). Modular multi-wavelength LED based light source for hyperspectral imaging. In *Clinical and Preclinical Optical Diagnostics II (2019), Paper 11075_56*, page 11075_56. Optica Publishing Group.

[66] Dooley, K. A., Lomax, S., Zeibel, J. G., Miliani, C., Ricciardi, P., Hoenigswald, A., Loew, M., and Delaney, J. K. (2013). Mapping of egg yolk and animal skin glue paint binders in Early Renaissance paintings using near infrared reflectance imaging spectroscopy. page 11.

[67] Egerton, G. and Morgan, A. (1970). The photochemistry of dyes ii—some aspects of the fading process. *Journal of the Society of Dyers and Colourists*, 86(6):242–249.

[68] Fan, C., Zhang, P., Wang, S., and Hu, B. (2018). A study on classification of mineral pigments based on spectral angle mapper and decision tree. In *Tenth International Conference on Digital Image Processing (ICDIP 2018)*, volume 10806, pages 1639–1643. SPIE.

[69] Farrell, M. and Mersereau, R. (2005). On the impact of pca dimension reduction for hyperspectral detection of difficult targets. *IEEE Geoscience and Remote Sensing Letters*, 2(2):192–195. event-title: IEEE Geoscience and Remote Sensing Letters.

[70] Feigelson, E. (2012). Classification in astronomy: Past and present. *Advances in Machine Learning and Data Mining for Astronomy*, pages 3–10.

[71] Fiedler, I. and A.Bayard, M. (1997). *Artists' pigments A Handbook of their history and characteristics*, volume 3.

[72] Fiorucci, M., Khoroshiltseva, M., Pontil, M., Traviglia, A., Del Bue, A., and James, S. (2020). Machine learning for cultural heritage: A survey. *Pattern Recognition Letters*, 133:102–108.

[73] Fischer, A. and Igel, C. (2012). An introduction to restricted boltzmann machines. Lecture Notes in Computer Science, pages 14–36, Berlin, Heidelberg. Springer.

[74] Fischer, C. and Kakoulli, I. (2006). Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications. *Studies in Conservation*, 51(sup1):3–16.

[75] Fonseca, B., Schmidt Patterson, C., Ganio, M., MacLennan, D., and Trentelman, K. (2019). Seeing red: towards an improved protocol for the identification of madder- and cochineal-based pigments by fiber optics reflectance spectroscopy (fors). *Heritage Science*, 7(1):92.

[76] France, F. G. (2011). Advanced spectral imaging for noninvasive microanalysis of cultural heritage materials: review of application to documents in the us library of congress. *Applied spectroscopy*, 65(6):565–574.

[77] Fukumoto, K., Tsumura, N., and Berns, R. (2020). Estimating pigment concentrations from spectral images using an encoderx2010;decoder neural network. *Journal of Imaging Science and Technology*, 64(3):30502–1–30502–15.

[78] Gao, Z., Du, M., Cao, N., Hou, M., Wang, W., and Lyu, S. (2023). Application of hyperspectral imaging technology to digitally protect murals in the qutan temple. *Heritage Science*, 11(1):1–16.

[79] Gath, I. and Geva, A. B. (1989). Unsupervised optimal fuzzy clustering. *IEEE Transactions on pattern analysis and machine intelligence*, 11(7):773–780.

[80] Grabowski, B., Masarczyk, W., Głomb, P., and Mendys, A. (2018). Automatic pigment identification from hyperspectral data. *Journal of Cultural Heritage*, 31:1–12.

[81] Grabusts, P. (2011). The choice of metrics for clustering algorithms. In *ENVIRONMENT. TECHNOLOGIES. RESOURCES. Proceedings of the International Scientific and Practical Conference*, volume 2, pages 70–76.

[82] Guan, Q., Xu, T., Feng, S., Yu, F., and Song, K. (2022). Optimal segmentation and improved abundance estimation for superpixel-based hyperspectral unmixing. *European Journal of Remote Sensing*, 55(1):485–506. publisher: Taylor Francis $_eprint : https : //doi.org/10.1080/22797254.2022.2125447$.

[83] Gueli, A. M., Bonfiglio, G., Pasquale, S., and Troja, S. O. (2016). Effect of particle size on pigments colour. *ColorResearch and application*, 42:236–243.

[84] Gurney, K. (2018). *An introduction to neural networks*. CRC press.

[85] Gustafson, D. E. and Kessel, W. C. (1979). Fuzzy clustering with a fuzzy covariance matrix. In *1978 IEEE conference on decision and control including the 17th symposium on adaptive processes*, pages 761–766. IEEE.

[86] He, F., Liu, T., and Tao, D. (2019a). Control batch size and learning rate to generalize well: Theoretical and empirical evidence. volume 32. Curran Associates, Inc. [Online; accessed 2022-11-14].

[87] He, J., Zhang, S., Yang, M., Shan, Y., and Huang, T. (2019b). Bi-directional cascade network for perceptual edge detection. arXiv:1902.10903 [cs].

[88] He, L., Li, J., Liu, C., and Li, S. (2017). Recent advances on spectral–spatial hyperspectral image classification: An overview and new guidelines. *IEEE Transactions on Geoscience and Remote Sensing*, 56(3):1579–1597.

[89] Herm, C. (2019). Emerald green versus scheele's green: evidence and occurrence. In *Proceedings of the 7th interdisciplinary ALMA conference, Bratislava, Slovakia*, pages 189–202.

[90] Hou, M., Zhou, P., Lv, S., Hu, Y., Zhao, X., Wu, W., He, H., Li, S., and Tan, L. (2018). Virtual restoration of stains on ancient paintings with maximum noise fraction transformation based on the hyperspectral imaging. *Journal of Cultural Heritage*, 34:136–144.

[91] Hu, C., Huang, X., Xia, G., Wang, Y., Liu, X., and Ma, X. (2022). High precision automatic extraction of cultural relic diseases based on improved slic and ap clustering. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43:801–807.

[92] Imani, M. and Ghassemian, H. (2020). An overview on spectral and spatial information fusion for hyperspectral image classification: Current trends and challenges. *Information fusion*, 59:59–83.

[93] Intharah, T. and Khiripet, N. (2012). Muralcut: automatic character segmentation from mural images. In *2012 9th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, pages 1–4. IEEE.

[94] Ji, Y., Zhang, H., Zhang, Z., and Liu, M. (2021). Cnn-based encoder-decoder networks for salient object detection: A comprehensive review and recent advances. *Information Sciences*, 546:835–857.

[95] Johnston-Feller, R. (2001). *Color science in the examination of museum objects: nondestructive procedures*. The Getty Conservation Institute.

[96] Jokilehto, J. (1999). A century of heritage conservation. *Journal of Architectural conservation*, 5(3):14–33.

[97] Jones, C., Daly, N. S., Higgitt, C., and Rodrigues, M. R. D. (2022). Neural network-based classification of x-ray fluorescence spectra of artists' pigments: an approach leveraging a synthetic dataset created using the fundamental parameters method. *Heritage Science*, 10(1):88.

[98] Karlik, B. and Olgac, A. V. (2011). Performance analysis of various activation functions in generalized mlp architectures of neural networks. *International Journal of Artificial Intelligence and Expert Systems*, 1(4):111–122.

[99] Karpathy, A., Johnson, J., and Fei-Fei, L. (2015). Visualizing and understanding recurrent networks. arXiv:1506.02078 [cs].

[100] Kleynhans, T., Messinger, D. W., and Delaney, J. K. (2020a). Towards automatic classification of diffuse reflectance image cubes from paintings collected with hyperspectral cameras. *Microchemical Journal*, 157:104934.

[101] Kleynhans, T., Schmidt Patterson, C. M., Dooley, K. A., Messinger, D. W., and Delaney, J. K. (2020b). An alternative approach to mapping pigments in paintings with hyperspectral reflectance image cubes using artificial intelligence. *Heritage Science*, 8(1):84.

[102] Kogou, S., Lee, L., Shahtahmassebi, G., and Liang, H. (2021). A new approach to the interpretation ofxrfspectral imaging data using neural networks. *X-Ray Spectrometry*, 50(4):310–319. publisher-place: Hoboken publisher: Wiley WOS:000557372200001.

[103] Kohonen, T. (2001). Self-organizing maps: Third extended addition, v. 30: Springer series in information services.

[104] Kriegeskorte, N. and Douglas, P. K. (2019). Interpreting encoding and decoding models. *Current Opinion in Neurobiology*, 55:167–179.

[105] Kruse, F. A., Lefkoff, A., Boardman, J., Heidebrecht, K., Shapiro, A., Barloon, P., and Goetz, A. (1993). The spectral image processing system (sips)-interactive visualization and analysis of imaging spectrometer data). *Remote Sensing of Environment 44*, pages 145–163.

[106] Lanitis, A., Stylianou, G., and Voutounos, C. (2012). Virtual restoration of faces appearing in byzantine icons. *Journal of Cultural Heritage*, 13(4):404–412.

[107] Lavvafi, M. R., Monadjemi, S. A., and Moallem, P. (2010). Film Colorization, Using Artificial Neural Networks and Laws Filters. *J. Comput.*, 5(7):1094–1099.

[108] Legrand, S., Vanmeert, F., Van der Snickt, G., Alfeld, M., De Nolf, W., Dik, J., and Janssens, K. (2014). Examination of historical paintings by state-of-the-art hyperspectral imaging methods: from scanning infra-red spectroscopy to computed x-ray laminography. *Heritage Science*, 2(1):13.

[109] Lei, C. and Chen, Q. (2019). Fully Automatic Video Colorization With Self-Regularization and Diversity. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3753–3761.

[110] Lei, T., Jia, X., Zhang, Y., Liu, S., Meng, H., and Nandi, A. K. (2018). Superpixel-based fast fuzzy c-means clustering for color image segmentation. *IEEE Transactions on Fuzzy Systems*, 27(9):1753–1766.

[111] Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, 2(2):164–168.

[112] Li, F., Ding, Y., Yu, B., and Dong, S. (2018). Video Inpainting for Vintage Film with Continuous Large Damaged Regions. In *2018 3rd International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, pages 619–622.

[113] Li, J., Xie, D., Li, M., Liu, S., and Wei, C. (2020). Pigment identification of ancient wall paintings based on a visible spectral image. *Journal of Spectroscopy*, 2020:1–8.

[114] Li, S., Song, W., Fang, L., Chen, Y., Ghamisi, P., and Benediktsson, J. A. (2019). Deep learning for hyperspectral image classification: An overview. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9):6690–6709. event-title: IEEE Transactions on Geoscience and Remote Sensing.

[115] Liang, H. (2012a). Advances in multispectral and hyperspectral imaging for archaeology and art conservation. *Applied Physics A*, 106(2):309–323.

[116] Liang, H. (2012b). Advances in multispectral and hyperspectralimaging for archaeology and art conservation. *Appl. Phys. Mater*, 106:309–323.

[117] Liggins, F., Vichi, A., Liu, W., Hogg, A., Kogou, S., Chen, J., and Liang, H. (2022). Hyperspectral imaging solutions for the non-invasive detection and automated mapping of copper trihydroxychlorides in ancient bronze. *Heritage Science*, 10(1):142.

[118] Lin, Y., Xu, C., and Lyu, S. (2019). Disease regions recognition on mural hyperspectral images combined by mnf and bp neural network. In *Journal of Physics: Conference Series*, volume 1325, page 012095. IOP Publishing.

[119] Lipton, Z. C., Berkowitz, J., and Elkan, C. (2015). A critical review of recurrent neural networks for sequence learning. arXiv:1506.00019 [cs].

[120] Liu, H. (2021). Chapter 5 - single-point wind forecasting methods based on reinforcement learning. In Liu, H., editor, *Wind Forecasting in Railway Engineering*, pages 177–214. Elsevier. DOI: 10.1016/B978-0-12-823706-9.00005-3.

[121] Liu, L., Catelli, E., Katsaggelos, A., Sciutto, G., Mazzeo, R., Milanic, M., Stergar, J., Prati, S., and Walton, M. (2022a). Digital restoration of colour cinematic films using imaging spectroscopy and machine learning. *Scientific Reports*, 12(1):21982.

[122] Liu, L., Delnevo, G., and Mirri, S. (2022b). Hierarchical clustering as an unsurpervised machine learning algorithm for hyperspectral image segmentation of films. GoodIT '22, page 397–402, New York, NY, USA. Association for Computing Machinery. [Online; accessed 2023-02-13].

[123] Liu, L., Miteva, T., Delnevo, G., Mirri, S., Walter, P., de Viguerie, L., and Pouyet, E. (2023). Neural networks for hyperspectral imaging of historical paintings: A practical review. *Sensors*, 23(5):2419.

[124] Lorusso, S., Barone, V., Colizzi, L., and Fonseca, C. D. (2014). Analytical-diagnostic and computing technologies for the attribution, authentication and economic evaluation of art works. *Conservation Science in Cultural Heritage*, 14:169–188.

[125] Lyu, S., Mao, J., and Hou, M. (2021). Mural linear feature enhancement algorithm combining semi-supervised image segmentation and dimensionality reduction. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 46:423–427.

[126] Machidon, O.-M. and Ivanovici, M. (2018). Digital color restoration for the preservation of reversal film heritage. *Journal of Cultural Heritage*, 33:181–190.

[127] Madariaga, J. M. (2015). Analytical chemistry in the field of cultural heritage. *Analytical Methods*, 7(12):4848–4876.

[128] Magnusson, M., Sigurdsson, J., Armansson, S. E., Ulfarsson, M. O., Deborah, H., and Sveinsson, J. R. (2020). Creating rgb images from hyperspectral images using a color matching function. In *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, pages 2045–2048.

[129] Makantasis, K., Karantzalos, K., Doulamis, A., and Doulamis, N. (2015). Deep supervised learning for hyperspectral data classification through convolutional neural networks. pages 4959–4962. ISSN: 2153-7003.

[130] Mardquardt, D. W. (1963). An algorithm for least square estimation of parameters. *J. Soc. Ind. Appl. Math*, 11:431–441.

[131] Mazzeo, R. (2017). *Analytical chemistry for cultural heritage*. Springer.

[132] Mazzeo, R., Joseph, E., et al. (2004). Micro-destructive analytical investigation for conservation and restoration. *Monumenti in bronzo all'aperto. Esperienze di conservazione a confronto*, pages 53–58.

[133] Mazzeo, R., Palazzi, C. E., Roccetti, M., and Sciutto, G. (2007). Computer-assisted pigment identification in artworks. In *Proc. of the European Conference on Internet and Multimedia Systems and Applications EuroIMSA 2007*, pages 1–6.

[134] McLachlan, G. J., Lee, S. X., and Rathnayake, S. I. (2019). Finite mixture models. *Annual review of statistics and its application*, 6:355–378.

[135] Mehri, M., Sliti, N., Héroux, P., Gomez-Krämer, P., Amara, N. E. B., and Mullot, R. (2015). Use of slic superpixels for ancient document image enhancement and segmentation. In *Document Recognition and Retrieval XXII*, volume 9402, pages 11–22. SPIE.

[136] Miliani, C., Monico, L., Melo, M. J., Fantacci, S., Angelin, E. M., Romani, A., and Janssens, K. (2018). Photochemistry of artists' dyes and pigments: Towards better understanding and prevention of colour change in works of art. *Angewandte Chemie International Edition*, 57(25):7324–7334.

[137] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., and Terzopoulos, D. (2022). Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7):3523–3542. event-title: IEEE Transactions on Pattern Analysis and Machine Intelligence.

[138] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., and Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533. number: 7540 publisher: Nature Publishing Group.

[139] Moghareh Abed, F. (2014). Pigment identification of paintings based on kubelka-munk theory and spectral images. *Ph. D. Thesis*.

[140] Monico, L., Cartechini, L., Rosi, F., Chieli, A., Grazia, C., De Meyer, S., Nuyts, G., Vanmeert, F., Janssens, K., Cotte, M., et al. (2020). Probing the chemistry of cds paints in the scream by in situ noninvasive spectroscopies and synchrotron radiation x-ray techniques. *Science advances*, 6(20):eaay3514.

[141] Monico, L., Sorace, L., Cotte, M., de Nolf, W., Janssens, K., Romani, A., and Miliani, C. (2019). Disclosing the binding medium effects and the pigment solubility in the (photo)reduction process of chrome yellows (pbcro4 pbcr1–xsxo4). *ACS Omega*, 4(4):6607–6619.

[142] Muñoz-Viñas, S. (2012). *Contemporary theory of conservation*. Routledge.

[143] Murtagh, F. and Contreras, P. (2012). Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(1):86–97.

[144] Newman, R. (1997). *Artists' pigments A Handbook of their history and characteristics*, volume 3.

[145] O'Shea, K. and Nash, R. (2015). An introduction to convolutional neural networks. arXiv:1511.08458 [cs].

[146] Osorio-Gomez, C. A., Mejia-Ospino, E., and Guerrero-Bermudez, J. E. (2009). Spectral reflectance curves for multispectral imaging, combining different techniques and a neural network. *Revista Mexicana De Fisica*, 55(2):120–124. publisher-place: Coyoacan publisher: Soc Mexicana Fisica WOS:000265620100007.

[147] Pan, N., Hou, M., Lv, S., Hu, Y., Zhao, X., Ma, Q., Li, S., and Shaker, A. (2017). Extracting faded mural patterns based on the combination of spatial-spectral feature of hyperspectral image. *Journal of Cultural Heritage*, 27:80–87.

[148] Paoletti, M. E., Haut, J. M., Plaza, J., and Plaza, A. (2019). Deep learning classifiers for hyperspectral imaging: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 158:279–317.

[149] Pappas, M. and Pitas, I. (2000). Digital color restoration of old paintings. *IEEE Transactions on Image Processing*, 9(2):291–294.

[150] Peng, J., Yu, K., Wang, J., Zhang, Q., Wang, L., and Fan, P. (2019). Mining painted cultural relic patterns based on principal component images selection and image fusion of hyperspectral images. *Journal of Cultural Heritage*, 36:32–39.

[151] Photoworks Photo Editor (2021). How to restore old photos without photoshop. https://photo-works.net/how-to-restore-old-photos.php.

[152] Picollo, M., Cucci, C., Casini, A., and Stefani, L. (2020). Hyper-spectral imaging technique in the cultural heritage field: New possible scenarios. *Sensors*, 20(10):2843.

[153] Polak, A., Kelman, T., Murray, P., Marshall, S., Stothard, D. J. M., Eastaugh, N., and Eastaugh, F. (2017). Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication. *Journal of Cultural Heritage*, 26:1–11.

[154] Pouyet, E., Devine, S., Grafakos, T., Kieckhefer, R., Salvant, J., Smieska, L., Woll, A., Katsaggelos, A., Cossairt, O., and Walton, M. (2017). Revealing the biography of a hidden medieval manuscript using synchrotron and conventional imaging techniques. *Analytica chimica acta*, 982:20–30.

[155] Pouyet, E., Miteva, T., Rohani, N., and de Viguerie, L. (2021). Artificial intelligence for pigment classification task in the short-wave infrared range. *Sensors*, 21(18):6150. number: 18 publisher: Multidisciplinary Digital Publishing Institute.

[156] Pouyet, E., Rohani, N., Katsaggelos, A. K., Cossairt, O., and Walton, M. (2018). Innovative data reduction and visualization strategy for hyperspectral imaging datasets using t-sne approach. *Pure and Applied Chemistry*, 90(3):493–506.

[157] Raimondi, V., Conti, C., Lognoli, D., and Palombi, L. (2013). Latest advancements in fluorescence hyperspectral lidar imaging of the cultural heritage. volume 9065, pages 274–284. SPIE. [Online; accessed 2022-12-19].

[158] Redman, J. (2007). Advances in Digital Imaging for Fine Art and Cultural Heritage. *NIP & Digital Fabrication Conference*, 2007(1):355–363.

[159] Ribes, A. and Schmit, F. (2002). Reconstructing spectral reflectances with mixture density networks. pages 486–491, Springfield. Soc Imaging Science Technology. WOS:000184048100102.

[160] Ribés, A. and Schmitt, F. (2003). A fully automatic method for the reconstruction of spectral reflectance curves by using mixture density networks. *Pattern Recognition Letters*, 24(11):1691–1701.

[161] Rodarmel, C. and Shan, J. (2002). Principal component analysis for hyperspectral image classification. *Surveying and Land Information Science*, 62(2):115–122.

[162] Rohani, N. (2021). *Machine Learning for Multi-Sensory Data*. PhD thesis, Evanston, Illinois.

[163] Rohani, N., Pouyet, E., Walton, M., Cossairt, O., and Katsaggelos, A. K. (2018). Nonlinear unmixing of hyperspectral datasets for the study of painted works of art. *Angewandte Chemie*, 130(34):11076–11080.

[164] Rohani, N., Pouyet, E., Walton, M., Cossairt, O., and Katsaggelos, A. K. (2019). Pigment unmixing of hyperspectral images of paintings using deep neural networks. pages 3217–3221, Brighton, United Kingdom. IEEE. [Online; accessed 2022-10-24].

[165] Rohani, N., Salvant, J., Bahaadini, S., Cossairt, O., Walton, M., and Katsagge-los, A. (2016). Automatic pigment identification on roman egyptian paintings by using sparse modeling of hyperspectral images. pages 2111–2115. ISSN: 2076-1465.

[166] Rojas, D. J. B., Fernandes, B. J. T., and Fernandes, S. M. M. (2020). A Review on Image Inpainting Techniques and Datasets. In *2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 240–247.

[167] Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. Lecture Notes in Computer Science, pages 234–241, Cham. Springer International Publishing.

[168] Rood, O. N. (1879). *Modern chromatics with applications to art and industry*. C.K.Paul, Londres.

[169] Saalfeld, S. and Tomancák, P. (2008). Automatic landmark correspondence detection for ImageJ. In *Proceedings of the ImageJ User and Developer Conference*, pages 128–133.

[170] Sabetsarvestani, Z., Sober, B., Higgitt, C., Daubechies, I., and Rodrigues, M. R. D. (2019). Artificial intelligence for art investigation: Meeting the challenge of separating x-ray images of the *Ghent Altarpiece*. *Science Advances*, 5(8):eaaw7416.

[171] Saleh, B. and Elgammal, A. (2015). Large-scale classification of fine-art paintings: Learning the right metric on the right feature. arXiv:1505.00855 [cs].

[172] Sato, T., Nakano, Y., Iga, T., Nakauchi, S., and Usui, S. (1996). Color reproduction based on low dimensional spectral reflectance using the principal component analysis. volume 1996, page 185–188. Society for Imaging Science and Technology. issue: 1.

[173] Sazli, M. H. (2006). A brief review of feed-forward neural networks. *Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering*, 50(01):0–0. number: 01.

[174] Scikit-learn (2023). Sklearn neural network mlp classifier. https://scikit-learn/stable/modules/generated/sklearn.neural_network.MLPClassifier.html. [Online; accessed 2023-02-10].

[175] Sharma, S., Sharma, S., and Athaiya, A. (2017). Activation functions in neural networks. *towards data science*, 6(12):310–316.

[176] Shen, Y., Xiao, L., Chen, J., and Pan, D. (2019). A spectral-spatial domain-specific convolutional deep extreme learning machine for supervised hyperspectral image classification. *IEEE Access*, 7:132240–132252. event-title: IEEE Access.

[177] Sherman, J. (1955). The theoretical derivation of fluorescent x-ray intensities from mixtures. *Spectrochimica acta*, 7:283–306. publisher: Elsevier.

[178] Shi, L., Babaei, V., Kim, C., Foshey, M., Hu, Y., Sitthi-Amorn, P., Rusinkiewicz, S., and Matusik, W. (2018). Deep multispectral painting reproduction via multi-layer, custom-ink printing. *ACM Transactions on Graphics*, 37(6):271:1–271:15.

[179] Shi, W., Koo, D. E. S., Kitano, M., Chiang, H. J., Trinh, L. A., Turcatel, G., Steventon, B., Arnesano, C., Warburton, D., Fraser, S. E., and Cutrale, F. (2020). Pre-processing visualization of hyperspectral fluorescent data with spectrally encoded enhanced representations. *Nature Communications*, 11(1):726. number: 1 publisher: Nature Publishing Group.

[180] Shitomi, R., Tsuji, M., Fujimura, Y., Funatomi, T., Mukaigawa, Y., Morimoto, T., Oishi, T., Takamatsu, J., and Ikeuchi, K. (2023). Unsupervised learning with a physics-based autoencoder for estimating the thickness and mixing ratio of pigments. *Journal of the Optical Society of America A*, 40(1):116.

[181] Sibi, P., Jones, S. A., and Siddarth, P. (2013). Analysis of different activation functions using back propagation neural networks. *Journal of theoretical and applied information technology*, 47(3):1264–1268.

[182] Signac, P. (1911). *D'Eugène Delacroix au Néo-Impressionnisme*. Henri Flour.

[183] Signac, P. (2021). *Journal 1894-1909*. Edition de Charlotte Hellman, Gallimard Musée d'Orsay.

[184] Singh, N., Arya, R., and Agrawal, R. (2018). Performance enhancement of salient object detection using superpixel based gaussian mixture model. *Multimedia Tools and Applications*, 77:8511–8529.

[185] SoftOrbits Photo Retoucher (2021). Old photo restoration software. https://www.softorbits.net/old-photo-restoration-software/.

[186] Soyed, E., Chaieb, R., and Kalti, K. (2020). Superpixel based segmentation of historical document images using a multiscale texture analysis. In *Proceedings of the 8th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT'18), Vol. 1*, pages 341–351. Springer.

[187] Stanco, F., Ramponi, G., and de Polo, A. (2003). Towards the automated restoration of old photographic prints: A survey. In *The IEEE Region 8 EURO-CON 2003. Computer as a Tool.*, volume 2, pages 370–374 vol.2.

[188] Striova, J., Ruberto, C., Barucci, M., Blazek, J., Kunzelman, D., Dal Fovo, A., Pampaloni, E., and Fontana, R. (2018). Spectral imaging and archival data in analysing madonna of the rabbit paintings by manet and titian. *Angewandte Chemie-International Edition*, 57(25):7408–7412. publisher-place: Weinheim publisher: Wiley-V C H Verlag Gmbh WOS:000434949200021.

[189] Sun, M., Zhang, D., Wang, Z., Ren, J., Chai, B., and Sun, J. (2015). What's wrong with the murals at the mogao grottoes: A near-infrared hyperspectral imaging method. *Scientific Reports*, 5:14371.

[190] Sun, P., Hou, M., Lyu, S., Wang, W., Li, S., Mao, J., and Li, S. (2022). Enhancement and restoration of scratched murals based on hyperspectral imaging—a case study of murals in the baoguang hall of qutan temple, qinghai, china. *Sensors*, 22(24):9780. number: 24 publisher: Multidisciplinary Digital Publishing Institute.

[191] Sun, Y., Xue, B., Zhang, M., and Yen, G. G. (2020). Completely automated cnn architecture design based on blocks. *IEEE Transactions on Neural Networks and Learning Systems*, 31(4):1242–1254. event-title: IEEE Transactions on Neural Networks and Learning Systems.

[192] Svozil, D., Kvasnicka, V., and Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks. *Chemometrics and Intelligent Laboratory Systems*, 39(1):43–62.

[193] Taufique, A. M. N. and Messinger, D. W. (2019). Hyperspectral pigment analysis of cultural heritage artifacts using the opaque form of kubelka-munk theory. volume 10986, pages 297–307. SPIE. [Online; accessed 2022-12-20].

[194] The MathWorks Inc. (2023a). Fuzzy c-means clustering. https://it.mathworks.com/help/fuzzy/fcm.html.

[195] The MathWorks Inc. (2023b). Statistics and machine learning toolbox (r2021b). https://www.mathworks.com/help/stats/index.html.

[196] The MathWorks Inc. (2023c). Subclust function (r2017a). https://it.mathworks.com/help/fuzzy/subclust.html.

[197] UNESCO Archives (1980). Recommendation for the Safeguarding and Preservation of Moving Images. http://portal.unesco.org/en/ev.php-URL_ID=13139&URL_DO=DO_TOPIC&URL_SECTION=201.html.

[198] Vaart, d. W. B. V.-v. and Lambers, K. (2019). Learning to look at lidar: The use of r-cnn in the automated detection of archaeological objects in lidar data from the netherlands. *Journal of Computer Applications in Archaeology*, 2(1):31–40. number: 1 publisher: Ubiquity Press.

[199] van Loon, A., Noble, P., de Man, D., Alfeld, M., Callewaert, T., Van der Snickt, G., Janssens, K., and Dik, J. (2020). The role of smalt in complex pigment mixtures in rembrandt's homer 1663: combining ma-xrf imaging, microanalysis, paint reconstructions and oct. *Heritage Science*, 8(1):90.

[200] Vargas, W. E. and Niklasson, G. A. (2001). Reflectance of pigmented polymer coatings: comparisons between measurements and radiative transfer calculations. *Applied Optics*, 40(1):85–94. publisher: Optica Publishing Group.

[201] Vermeulen, M., McGeachy, A., Xu, B., Chopp, H., Katsaggelos, A., Meyers, R., Alfeld, M., and Walton, M. (2022). Xrfast a new software package for processing of ma-xrf datasets using machine learning. *Journal of Analytical Atomic Spectrometry*, 37(10):2130–2143.

[202] Vermeulen, M., Smith, K., Eremin, K., Rayner, G., and Walton, M. (2021). Application of uniform manifold approximation and projection (umap) in spectral imaging of artworks. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 252:119547.

[203] Väisänen, T., Markkanen, J., Penttilä, A., and Muinonen, K. (2019). Radiative transfer with reciprocal transactions: Numerical method and its implementation. *PLOS ONE*, 14(1):e0210155. publisher: Public Library of Science.

[204] Wambugu, N., Chen, Y., Xiao, Z., Tan, K., Wei, M., Liu, X., and Li, J. (2021). Hyperspectral image classification on insufficient-sample and feature learning using deep neural networks: A review. *International Journal of Applied Earth Observation and Geoinformation*, 105:102603.

[205] Wan, Z., Zhang, B., Chen, D., Zhang, P., Chen, D., Liao, J., and Wen, F. (2020a). Bringing old photos back to life. pages 2744–2754. ISSN: 2575-7075.

[206] Wan, Z., Zhang, B., Chen, D., Zhang, P., Chen, D., Liao, J., and Wen, F. (2020b). Old Photo Restoration via Deep Latent Space Translation. *arXiv:2009.07047 [cs]*.

[207] Wang, K., Gou, C., Duan, Y., Lin, Y., Zheng, X., and Wang, F.-Y. (2017). Generative adversarial networks: introduction and outlook. *IEEE/CAA Journal of Automatica Sinica*, 4(4):588–598. event-title: IEEE/CAA Journal of Automatica Sinica.

[208] Wilhelm, H. G. and Brower, C. (1993). *The permanence and care of color photographs: traditional and digital color prints, color negatives, slides, and motion pictures*. Preservation Pub. Co, Grinnell, Iowa, U.S.A, 1st ed edition.

[209] Yang, Y. and Rushmeier, H. (2017). Special object extraction from medieval books using superpixels and bag-of-features. *Journal of Electronic Imaging*, 26(1):011008–011008.

[210] Zeng, Y. and Gong, Y. (2018). Nearest Neighbor based Digital Restoration of Damaged Ancient Chinese Paintings. In *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)*, pages 1–5.

[211] Zhang, N., Ding, S., Zhang, J., and Xue, Y. (2018). An overview on restricted boltzmann machines. *Neurocomputing*, 275:1186–1199.

[212] Zhang, Q., Cui, S., Liu, L., Wang, J., Wang, J., Zhang, E., Peng, J., Kogou, S., Liggins, F., and Liang, H. (2021). Deep learning for the extraction of sketches from spectral images of historical paintings. volume 11784, pages 11–20. SPIE. [Online; accessed 2022-10-27].

[213] Zhang, Y., Wang, G., Chung, F.-l., and Wang, S. (2022). A fuzzy system with common linear-term consequents equivalent to flnn and gmm. *International Journal of Machine Learning and Cybernetics*, 13(5):1475–1492.

[214] Zhao, H., Gallo, O., Frosio, I., and Kautz, J. (2017). Loss functions for image restoration with neural networks. *IEEE Transactions on Computational Imaging*, 3(1):47–57. event-title: IEEE Transactions on Computational Imaging.

[215] Zhao, Y. (2008). *Image segmentation and pigment mapping of cultural heritage based on spectral imaging*. Rochester Institute of Technology.

[216] Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., and He, Q. (2021). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1):43–76. publisher: Institute of Electrical and Electronics Engineers Inc.