Jones *et al. Heritage Science* (2022) 10:88
https://doi.org/10.1186/s40494-022-00716-3

Heritage Science

RESEARCH Open Access

Neural network-based classification of X-ray fluorescence spectra of artists' pigments: an approach leveraging a synthetic dataset created using the fundamental parameters method

Cerys Jones[1], Nathan S. Daly[2*], Catherine Higgitt[2] and Miguel R. D. Rodrigues[1]

Abstract

X-ray fluorescence (XRF) spectroscopy is an analytical technique used to identify chemical elements that has found widespread use in the cultural heritage sector to characterise artists' materials including the pigments in paintings. It generates a spectrum with characteristic emission lines relating to the elements present, which is interpreted by an expert to understand the materials therein. Convolutional neural networks (CNNs) are an effective method for automating such classification tasks—an increasingly important feature as XRF datasets continue to grow in size—but they require large libraries that capture the natural variation of each class for training. As an alternative to having to acquire such a large library of XRF spectra of artists' materials a physical model, the Fundamental Parameters (FP) method, was used to generate a synthetic dataset of XRF spectra representative of pigments typically encountered in Renaissance paintings that could then be used to train a neural network. The synthetic spectra generated—modelled as single layers of individual pigments—had characteristic element lines closely matching those found in real XRF spectra. However, as the method did not incorporate effects from the X-ray source, the synthetic spectra lacked the continuum and Rayleigh and Compton scatter peaks. Nevertheless, the network trained on the synthetic dataset achieved 100% accuracy when tested on synthetic XRF data. Whilst this initial network only attained 55% accuracy when tested on real XRF spectra obtained from reference samples, applying transfer learning using a small quantity of such real XRF spectra increased the accuracy to 96%. Due to these promising results, the network was also tested on select data acquired during macro XRF (MA-XRF) scanning of a painting to challenge the model with noisier spectra Although only tested on spectra from relatively simple paint passages, the results obtained suggest that the FP method can be used to create accurate synthetic XRF spectra of individual artists' pigments, free from X-ray tube effects, on which a classification model could be trained for application to real XRF data and that the method has potential to be extended to deal with more complex paint mixtures and stratigraphies.

Keywords: X-ray fluorescence, Convolutional neural networks, Deep learning, Transfer learning, Classification, Synthetic data, Fundamental parameters, Pigment, Painting

Introduction

X-ray fluorescence (XRF) spectroscopy is a non-invasive analytical technique for identifying chemical elements that has been used extensively to study works of art [1, 2]. It involves measuring the energy of characteristic X-rays

*Correspondence: Nathan.Daly@ng-london.org.uk

[2] National Gallery, Trafalgar Square, London WC2N 5DN, UK
Full list of author information is available at the end of the article

that are emitted when a sample is irradiated with photons from an X-ray source. The energy of the emitted X-rays is characteristic of each element, enabling identification of the elements present in a sample. With the development of scanning XRF systems capable of acquiring data on a macro scale (MA-XRF scanning), XRF is increasingly being used for in situ analysis of entire works of art, particularly relatively flat objects such as paintings and illuminated manuscripts, rather than just for individual point measurement [3–6]. While both types of systems provide information about the elemental composition in surface and subsurface layers, macroscale scanning systems can generate sub-millimetre resolution maps showing the spatial distribution of elements across an object. These maps provide a crucial tool to help understand an artist's materials and technique and to visualise subsurface features of paintings. However, depending on spectral quality and the energy overlap of characteristic X-rays, reliance on element distribution maps can be misleading as an individual map can sometimes include features that are in reality associated with other elements or background scatter, for example. Furthermore, using XRF data to identify materials such as pigments used by an artist typically relies on comparison of the signal associated with multiple elements, which can be accomplished in a more quantitative manner by working with XRF spectra. Identifying the pigments present in an object is important to inform conservation, storage and display decisions, and can also be useful to confirm an artist's palette or to aid in dating and attribution. However, to accomplish this using XRF requires background knowledge about the chemical composition of pigments available when the painting was created, as well as an understanding of the expected XRF spectra based on considerations such as energy-dependent detection sensitivity as well as secondary interactions between X-rays emitted from elements in layered pigment mixtures.

Strategies developed to reduce the burden of XRF spectral analysis for pigment classification have focused on dimensionality reduction via endmember identification and automation of aspects of the process. A number of methods exist to assign XRF spectra to chemical elements based on peak fitting or detection algorithms, requiring varying degrees of expert user input [7–13]. Extending this element assignment to pigment classification, Conover compared the elements detected in each spectrum to a table of expected major elements for each pigment [12]. However, as several pigments share major elements, distinguishing between them based on these alone was impossible and so the presence of elements in XRF spectra was instead used to improve pigment classification primarily based on the use of reflectance spectroscopy data. Alternative strategies have used the XRF spectra directly for pigment classification or clustering. Martins et al. implemented multivariate curve resolution—alternating least squares, a spectral unmixing method, with the aim of identifying pure components (in this case paints or their constituent pigment and fillers) in MA-XRF data of paintings with mixtures and layered compositions [14, 15]. However, the number of endmembers must be input, requiring some exploration of the data both visually and statistically before spectral unmixing can be performed and the identity of the endmembers assessed. Another approach by Kogou et al. reduced an MA-XRF dataset of a watercolour painting to a small number of mean cluster spectra using a self-organising map, an unsupervised machine learning method [16]. Here, however, some data reduction, namely summing of spectral bins around characteristic X-ray lines of interest, is needed to maintain a small number of output clusters that can then be manually interpreted.

Convolutional Neural Networks (CNNs) are a type of supervised machine learning technique commonly used for automating classification problems [17, 18], which learn from data without having to be fed explicit instructions or a model. They do, however, rely on large sets of labelled data for each class to train the CNN-based classification model. While such an approach could be used to develop an automatic classification model for pigments based only on XRF spectra but with no other user input, the databases of XRF spectra of pigments that are currently available provide only a single, representative XRF spectrum for each pigment (often acquired from modern, commercially available pigment samples) and so do not necessarily reflect the compositions of historical pigments, nor the compositional variations often encountered, particularly with pigments derived from mineral sources and in terms of the minor elements detected. CNNs have indeed been successfully used for pigment classification using reflectance spectral data collected on 14th century illuminated manuscripts [19]. In this research, sections of well-characterised and representative regions of the manuscripts were manually labelled, either as single pigments, or as mixtures of two pigments, to train the neural network. Whilst this approach overcame the issue of finding a large dataset for training by labelling real data, it meant that the model was limited by the composition of pigments and mixtures present in the selected training sets. Thus, while it performed well on the particular manuscripts studied, this CNN would not necessarily work for other artworks nor for pigments not represented in the manuscripts. Furthermore, as the model did not use a physics-based approach, the possibility to obtain quantitative information regarding the concentrations of the materials and pigment mixtures could not be determined.

In this research, in order to train a CNN for automatic pigment classification, a large dataset of synthetic XRF spectra was generated using the Fundamental Parameters (FP) method, a physical model based on Sherman's equations for generating XRF spectra of a sample of known composition [20]. The equations describe how the intensity of the fluorescence emitted by the elements in a sample is proportional to their concentration. The coefficient of proportionality is defined by a collection of physical coefficients called fundamental parameters and varies for each element and characteristic line. Therefore, to create synthetic spectra using the FP method, the concentration of each element in a sample (rather than simply whether an element is present or not) must be known, as well as the values for the various fundamental parameters. The FP method has previously been shown to be effective in the context of XRF analysis of paintings, for example to model absorption effects caused by paint layering and to correct for working distance variations caused by irregular paint surfaces [21, 22]. To allow for the variability of composition observed for historical pigments, multiple synthetic XRF spectra were generated for each pigment by varying the concentrations of the characteristic elements in each pigment. Once the synthetic dataset was generated, it was used to train a CNN, the performance of which was then tested using both synthetic XRF spectra and real XRF spectra collected in the National Gallery, London from reference materials of known composition. The CNN was then further fine-tuned using transfer learning techniques leveraging a small portion of the real XRF spectra collected [23–26]. This paper outlines the process of creating the synthetic dataset, training and fine-tuning the CNN, and also presents preliminary results from the classification model on a real XRF dataset of a painting.

Method

Selection of pigment classes

Due to the large variety of pigments used by artists throughout history, this study was focused on a selection of pigments typically encountered in Renaissance painting, [27] and therefore representative of the pigments present in a significant number of paintings in the National Gallery. This approach could, however, be extended to any set of pigments provided their elemental compositions are known. As some pigments, such as red lead and lead white, cannot be distinguished using XRF alone, they were grouped into a common pigment class (Table 1). This resulted in a total of 15 pigment classes: arsenic sulphides (orpiment and realgar); bone black (including ivory black); chalk; copper pigments (azurite, malachite and verdigris); gold (gold leaf); green earth; gypsum; iron earth pigments (red ochre, yellow ochre and sienna); lead–tin yellow (types I and II); lead

Table 1 Summary of pigment classes and data used to generate synthetic spectra

Pigment class	Description of data	References
Arsenic sulphides	*Average of theoretical orpiment (As_2S_3) and realgar (As_4S_4)*	[27]
Bone black	Average of values reported for bone black (1:18 Mg:Ca) and ivory black (1:8 Mg:Ca) from SEM–EDX measurements	[28]
Chalk	*Theoretical calcite ($CaCO_3$)*	[27]
Copper pigments	*Average of theoretical azurite ($2CuCO_3{\cdot}Cu(OH)_2$), malachite ($CuCO_3{\cdot}Cu(OH)_2$) and verdigris ($Cu(CH_3COO)_2Cu(OH)_2{\cdot}5H_2O$)*	[27]
Gold	Average of values for gilding on 14–15th century German and Swiss polychrome wooden sculptures from SEM–EDX measurements	[29]
Green earth	Average of values for Bohemian, Cypriot and Italian green earth pigment powders from SEM–EDX measurements	[30]
Gypsum	*Theoretical gypsum ($CaSO_4{\cdot}2H_2O$)*	[27]
Iron earth pigments	Average of values for iron earth pigment powders sourced across Europe from SEM–EDX and quantitative XRD measurements	[31, 32]
Lead–tin yellow	Average of values reported for lead–tin yellow types I and II in 14–15th century Italian paintings from SEM–EDX measurements	[33]
Lead pigments	Average of *theoretical red lead (Pb_3O_4)* and average hydrocerussite ($Pb_3(CO_3)_2(OH)_2$)/cerussite ($PbCO_3$) content of lead white in 14–16th century European paintings from synchrotron XRD measurements	[34]
Naples yellow	*Theoretical Naples yellow ($Pb_2Sb_2O_7$)*	[27]
Smalt	Average of values for smalt in sixteenth century Netherlandish and French paintings from SEM–EDX measurements	[35]
Ultramarine	Average of values for lapis lazuli in 14–18th century European paintings from SEM–EDX measurements	[36]
Umber	Average of values for umber pigment powders sourced across Europe from SEM–EDX measurements	[32]
Vermilion	*Theoretical vermilion (HgS)*	[27]

Pigment classes in *italics* use the theoretical formulas listed to generate synthetic spectra. See Table 2 for quantitative values for all pigment classes, sourced from the references listed

pigments (lead white and red lead); Naples yellow; smalt; ultramarine (lapis lazuli); umber; and vermilion.

Many of the pigments included are either naturally occurring minerals or use such materials in their synthesis, which can result in varied elemental compositions for pigments in historical paintings, as can the method of manufacture. To account for this, at least to some extent, data sourced from quantitative studies of the elemental composition of historical pigment samples where it was available were averaged and standard deviations determined, and this data was then used in the FP method calculations (Tables 1 and 2). Other commonly used pigments of the Renaissance, for example carbon-based blacks and organic colourant-based pigments such as lake pigments or indigo, were omitted due to their lack of characteristic elements detectable using typical XRF instrumentation. It should also be noted that a number of the pigments included in this study also contain characteristic elements that cannot be detected using typical XRF instrumentation, such as magnesium present in bone and ivory black,[28] but all of these characteristic elements were included in the FP method modelling. Other elements that are known to be commonly associated with certain pigments—such as strontium substitution in gypsum commonly seen in gesso on paintings [27, 37] or barium and zinc-containing accessory minerals often associated with azurite [38]—were omitted as they have not been robustly quantified in historical paint samples.

Table 2 Weight % by element (average +/- standarddeviation) for each pigment class used to generate synthetic spectra

Pigment class	H	C	O	Na	Mg	Al	Si	P	S	Cl	K	Ca	Ti
Arsenic sulphides	–	–	–	–	–	–	–	–	35 ± 6	–	–	–	–
Bone black	0 ± 0	10 ± 0	36 ± 1	–	3 ± 2	–	–	16 ± 0	–	–	–	35 ± 1	–
Chalk	–	12 ± 0	48 ± 0	–	–	–	–	–	–	–	–	40 ± 0	–
Copper pigments	2 ± 2	8 ± 3	42 ± 9	–	–	–	–	–	–	–	–	–	–
Gold	–	–	–	–	–	–	–	–	–	–	–	–	–
Green earth	–	0 ± 0	44 ± 1	0 ± 0	3 ± 1	3 ± 2	26 ± 3	0 ± 0	–	–	6 ± 2	2 ± 2	1 ± 1
Gypsum	2 ± 0	–	56 ± 0	–	–	–	–	–	19 ± 0	–	–	23 ± 0	–
Iron earth pigments	1 ± 1	1 ± 2	47 ± 5	–	1 ± 3	6 ± 6	16 ± 11	–	4 ± 6	–	1 ± 1	8 ± 10	0 ± 1
Lead–tin yellow	–	–	13 ± 3	–	–	–	3 ± 4	–	–	–	–	–	–
Lead pigments	0 ± 0	2 ± 2	13 ± 4	–	–	–	–	–	–	–	–	–	–
Naples yellow	–	–	15 ± 0	–	–	–	–	–	–	–	–	–	–
Smalt	–	–	44 ± 2	0 ± 0	–	0 ± 0	33 ± 2	–	–	–	5 ± 5	1 ± 1	–
Ultramarine	–	–	35 ± 12	11 ± 2	0 ± 0	17 ± 1	22 ± 2	–	6 ± 2	1 ± 1	3 ± 3	4 ± 2	–
Umber	1 ± 1	1 ± 2	47 ± 5	–	1 ± 0	4 ± 2	15 ± 6	0 ± 0	2 ± 3	–	1 ± 0	2 ± 1	–
Vermilion	–	–	–	–	–	–	–	–	14 ± 0	–	–	–	–

Pigment class	Mn	Fe	Co	Ni	Cu	As	Ag	Sn	Sb	Au	Hg	Pb	Bi
Arsenic sulphides	–	–	–	–	–	65 ± 6	–	–	–	–	–	–	–
Bone black	–	–	–	–	–	–	–	–	–	–	–	–	–
Chalk	–	–	–	–	–	–	–	–	–	–	–	–	–
Copper pigments	–	–	–	–	48 ± 14	–	–	–	–	–	–	–	–
Gold	–	–	–	–	0 ± 1	–	1 ± 1	–	–	99 ± 2	–	–	–
Green earth	0 ± 0	16 ± 2	–	–	–	–	–	–	–	–	–	–	–
Gypsum	–	–	–	–	–	–	–	–	–	–	–	–	–
Iron earth pigments	0 ± 0	16 ± 13	–	–	–	–	–	–	–	–	–	–	–
Lead–tin yellow	–	–	–	–	–	–	–	16 ± 6	–	–	–	68 ± 4	–
Lead pigments	–	–	–	–	–	–	–	–	–	–	–	85 ± 8	–
Naples yellow	–	–	–	–	–	–	–	–	32 ± 0	–	–	54 ± 0	–
Smalt	–	3 ± 1	4 ± 1	1 ± 0	–	8 ± 3	–	–	–	–	–	–	1 ± 1
Ultramarine	–	–	–	–	–	–	–	–	–	–	–	–	–
Umber	2 ± 3	25 ± 10	–	–	–	–	–	–	–	–	–	–	–
Vermilion	–	–	–	–	–	–	–	–	–	–	86 ± 0	–	–

Values are sourced from studies with quantitative SEM–EDX or XRD measurements, or (when in *italics*) using theoretical formulas as noted (with source references provided) in Table 1

Collection of real XRF spectra (from reference materials and paintings)

Single point spectra and area scans were collected with a macro XRF (MA-XRF) scanner (M6 Jetstream, Bruker Corp.). The M6 Jetstream is equipped with a 30 W rhodium X-ray tube with polycapillary optics, here operated at 50 kV and 600 μA, as well as a 60 cm^2 silicon drift detector set to detect 275 kilocounts/s; the source and detector are in fixed positions normal to and at 60° relative to the target surface, respectively. For all measurements the instrument was operated with a 580 μm excitation spot size, leaving an approximately 25 mm air gap between the X-ray tube nose piece and the target surface.

When acquiring XRF spectra from reference pigment samples, two distinct batches or varieties of pigment within each class were selected from among the commercially available and historic reference pigments in the collection of the National Gallery (except for the copper pigments class where one batch of each of azurite, malachite, and verdigris was used), and analysed either directly on powdered samples or on samples previously painted in various binding media on inert substrates for other studies (detailed in Additional file 1). Five spectra of each sample were collected with 15 s of real time acquisition. Although many of the pigment reference samples used are commercially available pigments that differ in composition from historical pigments or are samples that may not have been processed using traditional method, the elemental compositions of most of the pigment reference samples used have been previously characterised in other studies (see Additional file 1) [39–41]. It should be noted that the compositions of these pigment reference samples were not used to inform the elemental compositions noted in Table 2, although as discussed later in the paper, these spectra were used to refine the CNN model. Each of the resulting spectra from the reference samples were assessed and labelled using MATLAB for use as test data.

In order to determine the sensitivity of the MA-XRF instrument over the measured energy range for the creation of synthetic XRF spectra (described below), additional spectra of elemental standards (Bruker Corp.) were collected using the M6 Jetstream under the same experimental conditions as the pigment samples (see Additional file 1).

Select regions of the painting *Saint Michael Triumphs over the Devil* by Bartolomé Bermejo (NG6553, National Gallery, London) were scanned also with the M6 Jetstream under similar conditions as point measurements. With the same 580 μm excitation spot size, the scanning was completed with 580 μm spacing between pixels and a dwell time of 10 ms/pixel.

Creation of synthetic XRF spectra

The model used in this research only accounted for primary fluorescence of single layered samples, though more complex models that include secondary fluorescence and layered samples can also be incorporated in future work. Therefore, the FP method equation used in this research to calculate the intensity of primary fluorescence emitted by the element i in a sample of mass thickness d is:

$$P_i(\lambda) = \frac{G}{\sin\psi_1} C_i \epsilon_i \frac{\tau_{i,\lambda}}{\mu'_\lambda + \mu'_i} \left\{1 - exp\left\{-\left(\mu'_\lambda + \mu'_i\right)d\right\}\right\} \quad (1)$$

where G is the geometrical factor accounting for the detected irradiated sample area, C_i is the concentration of the i-th element, ϵ_i is the excitation factor of the i-th element, $\tau_{i,\lambda}$ are the photoelectric absorption coefficients for element i at irradiation wavelength λ, $\mu'_\lambda = \frac{\mu_\lambda}{\sin\psi_1}$ where μ_λ is the mass absorption coefficient of the sample at irradiation wavelength λ, $\mu'_i = \frac{\mu_i}{\sin\psi_2}$ where μ_i is the mass absorption coefficient of the sample at the fluorescence emission wavelength λ_i, ψ_1 is the angle of incidence and ψ_2 is the angle towards the detector [42]. The geometrical factor G is calculated as described by de Boer, where G is proportional to the solid angle contained by the detector surface $\Delta\Omega/4\pi$, the irradiated sample width perpendicular to the surface w, the incident beam diameter d_0, and $1/a^2$ where a is the distance from source to surface [43]. The values for the characteristic energies, relative line intensities and jump ratios were taken from Elam et al., as were the fluorescence yields [44]. However, the value for the Ca-K shell in Elam et al. was updated to 0.166 according to Wolff et al., and the absorption coefficients were calculated using the tables in Thinh and Leroux [45, 46].

To create the synthetic XRF spectra and to account for the variable elemental compositions of historical pigments, the mean and standard deviation of the concentration of the elements in the pigments of interest were calculated by aggregating quantitative data available in the literature obtained from historic paint samples and using theoretical stoichiometric compositions for pigments where appropriate (see Additional file 1). Based on this information, 200 spectra were then created for each pigment class, resulting in a dataset containing 3000 spectra in total. The concentrations used to generate the spectra were chosen randomly from a normal distribution characterised by the means and standard deviations. The intensities of the characteristic X-ray lines were calculated for each element using Eq. 1 and the resulting signals summed together to produce the full spectrum for a pigment (Fig. 1a). In order to create the characteristic Gaussian-shaped fluorescence peaks (Fig. 1b), the

characteristic X-ray lines were then convolved with a Gaussian function with variance σ^2 described as follows:

$$\sigma^2 = \left(\frac{noise}{2.3548}\right)^2 + 0.00385 * fano * E_j \quad (2)$$

where *noise* is the electronic contribution to the peak width *fano* is the fano factor, E_j is the energy of the characteristic line in keV, and 0.00385 is the energy in keV required to create an electron–hole pair in silicon [10]. The *noise* and *fano* values used were as determined by fitting acquired XRF spectra using PyMCA. Finally, it is well known that the M6 Jetstream MA-XRF instrument used for this study has limited sensitivity to detect high energy X-rays, particularly K lines higher in energy than those of the rhodium X-ray source, due to its use of polycapillary optics [4]. A sensitivity curve was therefore experimentally derived for the entire energy range of the detector (see Additional file 1) and applied to the synthetic XRF spectra to adjust the expected fluorescence peak intensities (Fig. 1c) for final comparison to acquired XRF spectra (Fig. 1d). The FP method was implemented using MATLAB and a script was written to automatically generate normalised synthetic spectra for the different pigments and produce the corresponding labels.

Convolutional neural network for classification of artists' pigments

A CNN was trained to classify XRF spectra into one of the 15 pigment classes using the Python library Keras v2.4.2 [47]. The network was comprised of three convolutional layers with max pooling layers in between. The *ReLU* activation function was used for each layer except the final layer, which used the *Softmax* activation function. Each spectrum was input into the network and the output was the label for the corresponding pigment. 70% of the synthetic dataset was used to train the network and the remaining 30% was used for testing. The network was also tested on the real XRF spectra from reference pigment and paint samples recorded using the Bruker instrument. The network took approximately 1 h 30 min to train and, once trained, took 30 s to test on each dataset.

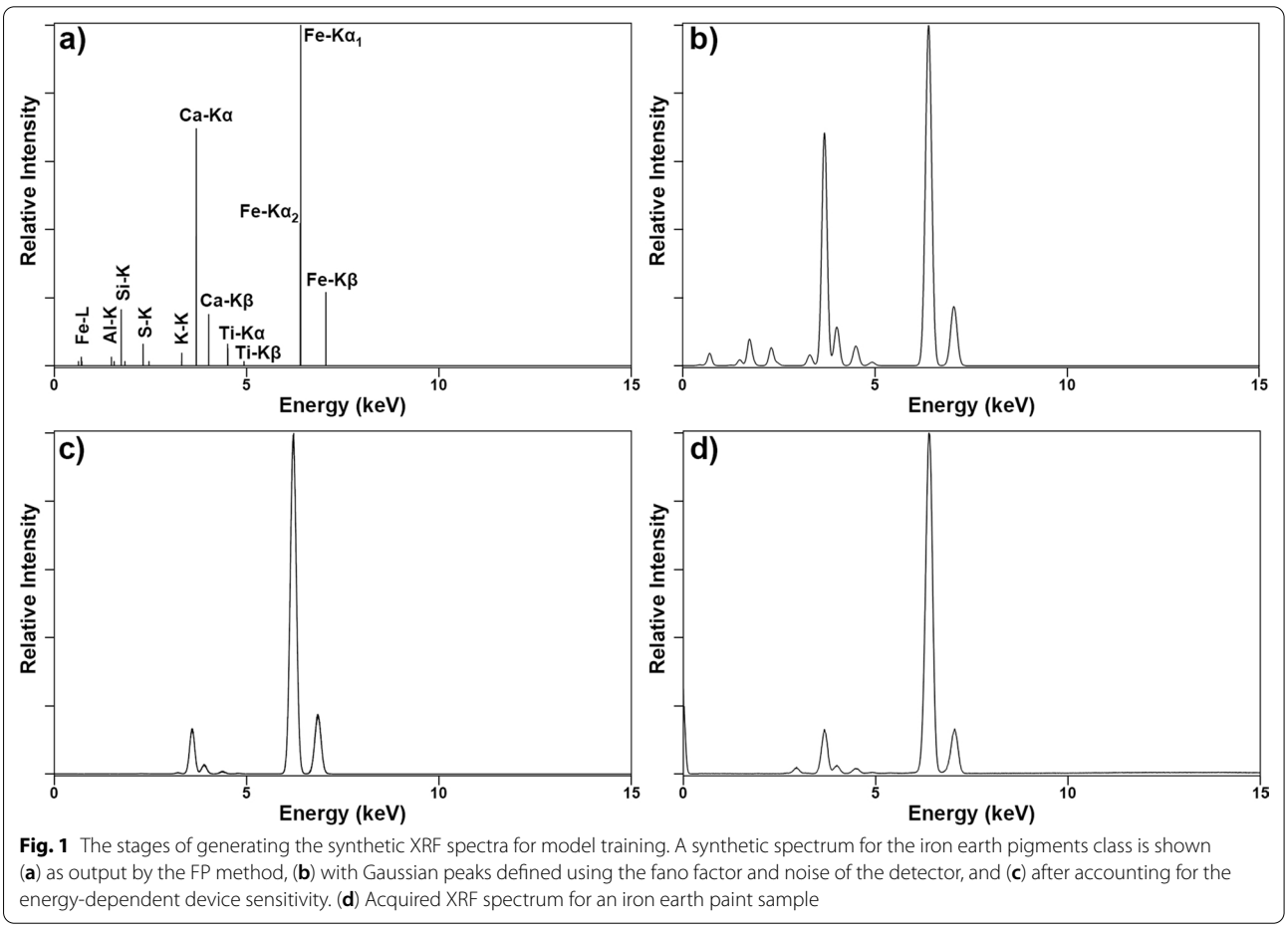

Fig. 1 The stages of generating the synthetic XRF spectra for model training. A synthetic spectrum for the iron earth pigments class is shown (**a**) as output by the FP method, (**b**) with Gaussian peaks defined using the fano factor and noise of the detector, and (**c**) after accounting for the energy-dependent device sensitivity. (**d**) Acquired XRF spectrum for an iron earth paint sample

Increasing accuracy using transfer learning

As the model was trained on the synthetic data created using the FP method (which may not exactly match real XRF spectra acquired with the M6 Jetstream), standard transfer learning—an approach whereby a machine learning model trained on one set of data can be applied to data with slightly different properties—was also used to tune the model further to improve its accuracy on real data [23–26]. This involved freezing the weights for all of the layers in the network except the final network layer. Half of the real XRF spectra from reference pigment and paint samples of known composition was then used to retrain this final layer and the remaining half was used for testing. This model with transfer learning was then used for additional testing, based on real spectra from an MA-XRF dataset from a painting as described above.

Results and discussion

Comparison of real and synthetic data

The FP method automatically generated synthetic XRF spectra that were similar to those collected using an XRF spectrometer: the characteristic X-ray lines of the elements were present at the same energies and approximately the same relative intensities, although some of the peaks at the high end of the energy range were more pronounced in the synthetic spectra while those at the low end of the energy range were diminished compared to real spectra (Fig. 2). The model developed in this research did not incorporate the argon K lines associated with atmospheric XRF measurements, or the broad continuum (bremsstrahlung radiation) and Rayleigh and Compton scatter peaks arising from the rhodium X-ray source present in the real spectra. While this meant that there were clear differences between the synthetic and real data, they do not appear to compromise the effectiveness of the model and provide the additional benefit that the spectra generated are not dependent on the specific instrument parameters and X-ray source chosen.

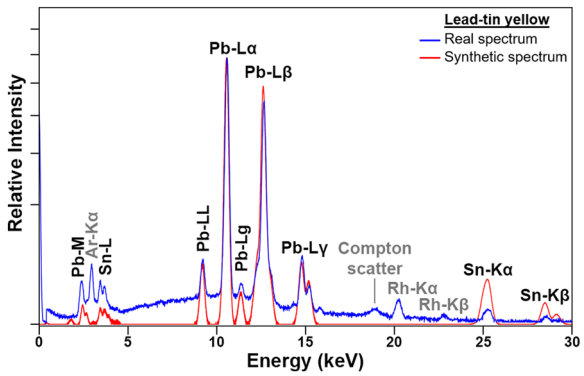

Fig. 2 A comparison of XRF spectra of lead-tin yellow. The spectrum shown in blue was recorded using the Bruker instrument in the National Gallery, while the one in red is a synthetic spectrum created for the same pigment using the FP method. The relative fluorescence intensities are plotted on a square-root scale with characteristic X-ray lines labelled for each relevant element in black and systematic features not included in the FP model labelled in grey

Classification results without transfer learning

When the initial model was trained on the synthetic spectra, the training and testing accuracy was 100%. However, when this model was applied to the real XRF spectra from reference samples the accuracy decreased to 55% (Table 3). While the reason for misclassification of each spectrum cannot be directly ascertained when using a CNN, some observations about the performance of the model can be made. While the network had the most success classifying pigments with a single major characteristic element, such as copper pigments, chalk and lead pigments, in a rare exception, the model misclassified all 10 vermilion spectra as lead pigments. The CNN also tended to misclassify pigments with the same major element, where experts would use the presence of minor peaks in the low energy range of the spectrum to distinguish them. For example, a large calcium peak is typical of bone black, chalk and gypsum, though the weak signal from phosphorous or sulphur is generally used to distinguish bone black or gypsum from chalk, respectively. The model appeared to misclassify all bone black spectra as chalk and, of the 10 gypsum spectra, misclassified one as bone black and six as chalk. Pigments which shared the same minor elements and trace impurities or where the spectra.

Included signals resulting from mixtures or layering also troubled the model. For example, two of the five gold spectra measured from a gilded panel were misclassified as chalk. The XRF spectra acquired from this reference panel contain appreciable iron and calcium signals, from the iron earth-containing bole layer and gypsum-based ($CaSO_4{\cdot}2H_2O$) gesso preparation layer respectively, which are present below the thin layer of gold leaf. A more complex model accounting for mixtures and layering should help avoid this type of misclassification.

Classification results with transfer learning

The reduced classification accuracy of the CNN when applied to real spectra from reference samples implied that discrepancies between the synthetic and real spectra could affect the performance of the model. In order to improve the accuracy of the model an adaptation process was applied using real data and transfer learning. The final layer of the CNN, prior to the *Softmax* activation function, was retrained using half of

Table 3 Pigment classification confusion matrix of real XRF spectra using the CNN trained on synthetic data

Labelled pigment class \ Predicted pigment class	Arsenic sulphides	Bone black	Chalk	Copper pigments	Gold	Green earth	Gypsum	Iron earth pigments	Lead-tin yellow	Lead pigments	Naples yellow	Smalt	Ultramarine	Umber	Vermilion
Arsenic sulphides	**9**									*1*					
Bone black			*10*												
Chalk			**10**												
Copper pigments				**10**											
Gold			*2*		**8**										
Green earth						**10**									
Gypsum		*1*	*6*				**3**								
Iron earth pigments			*5*					**5**							
Lead-tin yellow										*10*					
Lead pigments										**10**					
Naples yellow			*5*			*5*									
Smalt		*10*													
Ultramarine		*2*											**8**		
Umber														**10**	
Vermilion										*10*					

Correctly classified XRF spectra from reference samples are denoted in **bold** while misclassified spectra are denoted in *italics*

the real XRF spectra from reference samples, resulting in the accuracy of the model increasing to 96% when tested on the remaining spectra (Table 4). Although only a small number of spectra were used in training (five spectra per pigment), the accuracy of the CNN improved markedly, particularly among pigments with the same major element such as lead pigments, lead–tin yellow and Naples yellow. The dataset used to re-train the final layer was much smaller than that required to train the network from scratch: 2250 synthetic spectra were used in the initial training phase whereas only 75 real XRF spectra from reference samples were used in the transfer learning stage. The remaining misclassified spectra include two gypsum spectra and one gold spectrum (Fig. 3). Because the minor sulphur peak was not detected in the two gypsum spectra from reference samples given the experimental conditions used, the model misclassified them as chalk. The gold spectrum was misclassified as an iron earth pigment due to the significantly higher iron signal present in this spectrum than the other spectra used in transfer learning and testing, which originates from the bole layer below the gold leaf on the reference gilded panel from which the spectrum was acquired.

Table 4 Pigment classification confusion matrix of real XRF spectra using the CNN with transfer learning

Labelled pigment class \ Predicted pigment class	Arsenic sulphides	Bone black	Chalk	Copper pigments	Gold	Green earth	Gypsum	Iron earth pigments	Lead-tin yellow	Lead pigments	Naples yellow	Smalt	Ultramarine	Umber	Vermilion
Arsenic sulphides	**5**														
Bone black		**5**													
Chalk			**5**												
Copper pigments				**5**											
Gold					**4**			*1*							
Green earth						**5**									
Gypsum			*2*				**3**								
Iron earth pigments								**5**							
Lead-tin yellow									**5**						
Lead pigments										**5**					
Naples yellow											**5**				
Smalt												**5**			
Ultramarine													**5**		
Umber														**5**	
Vermilion															**5**

Correctly classified XRF spectra from reference samples are denoted in **bold** while misclassified spectra are denoted in *italics*

Classification results using spectra acquired during MA-XRF scanning of a painting—preliminary testing

In order to further validate the classification approach incorporating the transfer learning stage, preliminary testing of real spectra collected from a painting under the usual settings used for MA-XRF scanning of paintings at the National Gallery was performed. Acquisition times for MA-XRF scans of paintings, typically on the order of 10–100 ms, are much shorter than for spot measurements. While this makes the scanning of large areas and subsequent generation of element distribution maps feasible in a timely manner, individual spectra in these datasets have significantly reduced signal and greater noise. Therefore, a small number of pixels of an MA-XRF scan of an early Renaissance painting, *Saint Michael Triumphs over the Devil* by Bartolomé Bermejo, were averaged to create five spectra for preliminary testing of the classification model (Fig. 4). This painting recently underwent conservation treatment and scientific analysis in preparation for the 2019 exhibition *Bartolomé Bermejo: Master of the Spanish Renaissance* at the National Gallery, London [48]. As the pigments in the painting had been well-characterised through analysis of paint samples prepared as cross sections and the painting has a

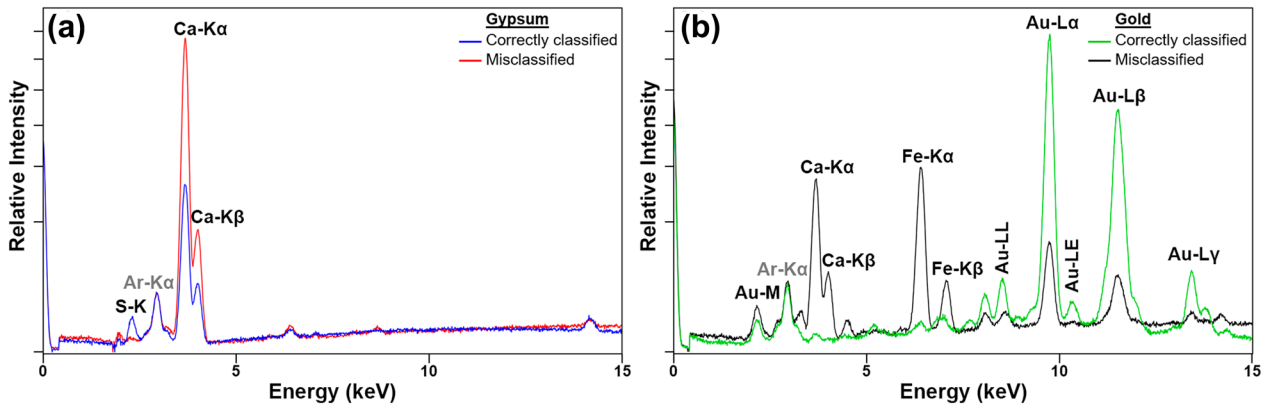

Fig. 3 Examples of XRF spectra of pigments correctly classified and misclassified using the CNN with transfer learning. **a** The spectrum shown in blue was correctly classified as gypsum, while the one in red was misclassified as chalk. **b** The spectrum shown in green was correctly classified as gold, while the one in black was misclassified as an iron earth pigment. The relative fluorescence intensities are plotted on a square-root scale with major characteristic X-ray lines labelled for each relevant element

relatively simple stratigraphy, [49] the MA-XRF scans of the painting were an ideal candidate to test the pigment classification model. The CNN correctly classified four of the five spectra, with the spectrum corresponding to a small area loss which was filled during conservation (spectrum B) assigned as gypsum instead of chalk (Table 5). Since the difference between these two pigments is what is usually a minor sulphur peak, the reduced signal-to-noise ratio of spectra from MA-XRF scanning data could make this distinction even more challenging.

Towards modelling XRF spectra of historical paintings

While the results from testing the CNN with transfer learning on both real XRF spectra obtained from reference pigment samples and extracted from select regions of a historical painting are promising, it is important to note and discuss the approximations made necessarily for this research and how the model could be improved in the future for use on historical samples and paintings.

While it could be argued that classifying spectra containing copper as copper pigments does not require machine learning, this study is a proof of concept, and the approach could readily be extended to distinguish a wider range of different pigments that share elements. This is of particular importance for the classification of the XRF spectra of nineteenth and twentieth century pigments, for example, where the rapid expansion of the synthesis of pigments resulted in a wide array of colours and materials achieved simply by altering the minor components or ratios of components present in their manufacture [50]. Further, development of the model to include both mixtures and layering of pigments rather than individual pigments will be of paramount importance to better approximate the XRF spectra obtained from historical paintings. While the network performed well on spectra obtained from XRF scanning of Bermejo's *Saint Michael Triumphs over the Devil*, many paintings involve more complex stratigraphies and intimate mixtures of pigments that would require a more sophisticated model to simulate their XRF spectra and to train a CNN to properly classify the pigments present. Such a neural network would offer an alternative to classical signal processing methods of fitting XRF spectra, potentially significantly reducing the computational time needed to model these complex systems, particularly when analysing large MA-XRF datasets. Successful modelling of pigment mixtures and layering will also rely on further refinement of the quantitative values used for the elemental composition of each pigment or pigment class as more data of this type is made publicly available. This refinement may also improve the ability of the trained CNN to distinguish between minor peaks in the XRF spectra of pigments associated with trace elements. This will be of increasing importance as this research builds to explore classification of the XRF spectra of a wider range of pigments. As discussed above, although suitable quantitative data is not currently widely available, it would also be valuable to incorporate information about additional elements known to be commonly and consistently associated with particular pigments as these minor components are often highly characteristic and informative. Indeed, it should be noted that pigment classification by neural networks is defined by the pigment classes and chemical compositions input, meaning that minor chemical components or indeed unique pigments that are not included in the model would need to be characterized using more manual analysis of XRF spectra.

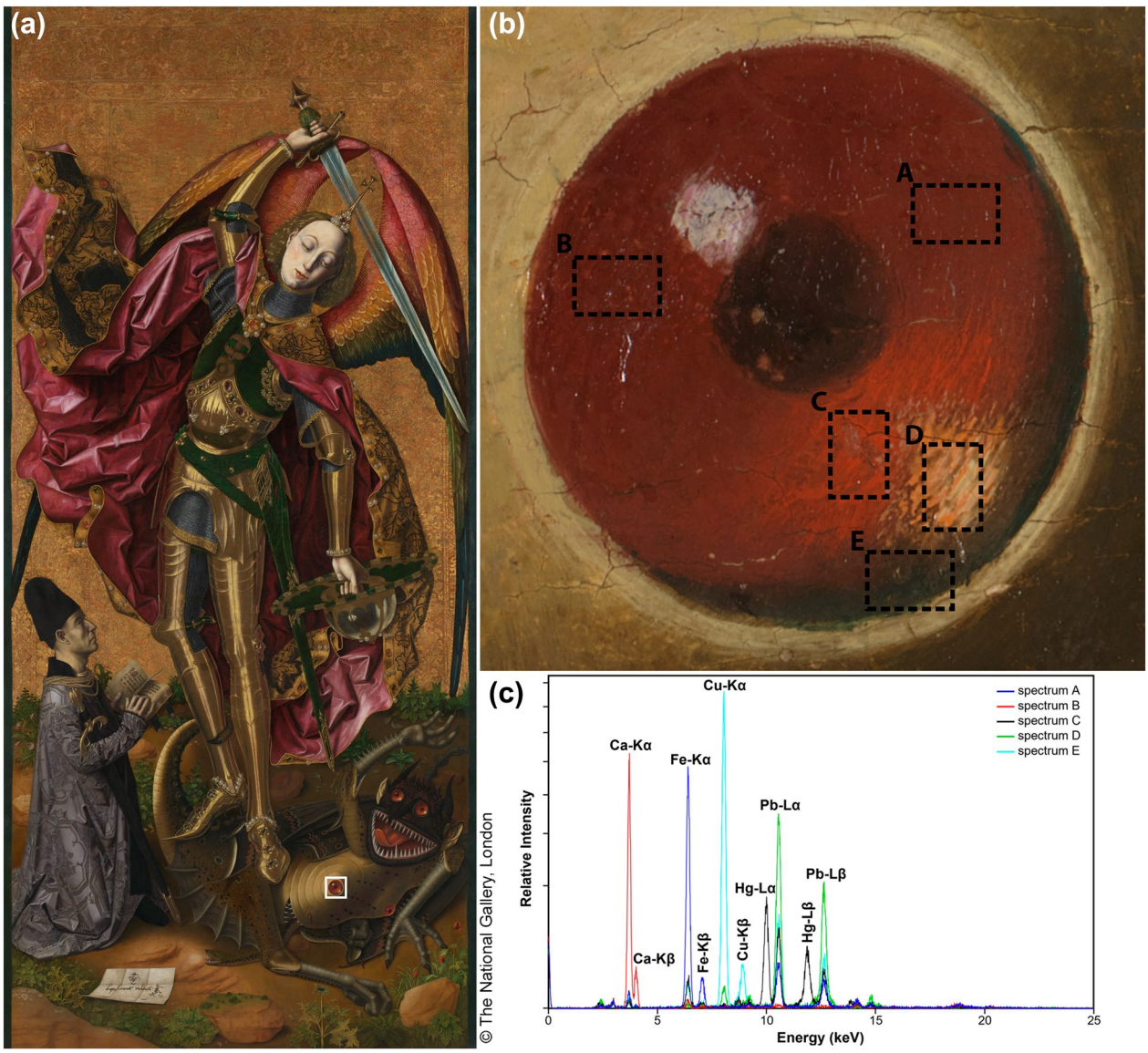

Fig. 4 Bartolomé Bermejo, *Saint Michael Triumphs over the Devil* (NG6553), 1468, oil and gold on wood, 179.7 × 81.9 cm. **a** Image of painting with detail of interest highlighted in white (image © The National Gallery, London). **b** Detail of the painting corresponding to the area of an MA-XRF scan used to extract spectra from regions A through E. **c** Spectra of regions A through E, each an average of 24 pixels of data, noted in the detail image. The relative fluorescence intensities are plotted on a square-root scale with major characteristic X-ray lines labelled for each relevant element

Although using the real XRF spectra from reference materials of known composition for transfer learning appeared to improve the performance of the CNN, it is known that the elemental composition of some of these reference pigments differs from typical historical pigment samples. It would therefore be of interest to explore using real XRF spectra acquired from well-characterised historical pigment samples for the transfer learning stage to further improve classification performance.

Beyond the classification of individual pigment spectra, what the approach described here also offers is a potential means by which MA-XRF scanning data could be processed directly to create pigment distribution maps as an alternative to the element distribution maps or element correlation plots more commonly produced by fitting and deconvoluting the XRF spectra.

Table 5 Pigment classification results of spectra from MA-XRF scanning of a historical painting (Fig. 4)

Spectrum	Pigment class based on analysis of paint samples*	Predicted pigment class
A	Iron earth pigment	Iron earth pigment
B	Chalk	Gypsum
C	Vermilion	Vermilion
D	Lead pigment	Lead pigment
E	Copper pigment	Copper pigment

*The pigment classes listed are the major pigments used in each region from which spectra were extracted and, in some cases, have been identified based on visual appearance and by extrapolation from paint samples taken from other nearby locations. More precise information beyond just the pigment class is therefore available for a number of the spectra: A—red earth; D—red lead possibly mixed with a small amount of lead—tin yellow or potentially lead white and a red lake pigment; E—verdigris

Conclusions

XRF is an analytical technique that can be used to detect the elements present at a particular location in a painting, from which an expert typically identifies which pigments are present. Motivated in part by the increasing size of XRF datasets generated during the scanning of paintings, this research aimed to automate this process by training a neural network to classify XRF spectral data into pigment classes. However, training machine learning models requires large amounts of labelled data and such datasets of XRF spectra for historic artists pigments do not currently exist and would require a large amount of time and resources to compile. Therefore, in this research, the FP method was used to create a training dataset of 3000 synthetic XRF spectra representing 15 classes of pigments in use during the Renaissance period. The synthetic spectra generated require minimal information about the geometry of the instrument used and do not consider the X-ray tube source or settings, allowing users with other instruments to easily adapt for their setup simply by acquiring XRF spectra from a series of elemental standards to account for energy-dependent device sensitivity. A CNN was trained using the synthetic spectra generated by the FP method and tested on both synthetic and real spectra of pigments with accuracies of 100% and 55%, respectively. The model failed primarily when classifying spectra with the same major elements where minor element peaks are needed for differentiation and also when signals from pigments in the layers beneath the surface were present in the spectra, though the accuracy of the model was improved considerably to 96% with the implementation of transfer learning. This involved freezing the weights in all of the CNN layers except the final one and retraining this layer using a small dataset of real XRF spectra. With this improved model, preliminary testing on spectra from MA-XRF scanning of a historical painting, where spectral signal is considerably reduced, produced encouraging results. While the small area of MA-XRF scanning studied here was relatively straightforward in composition, many areas of historical paintings are comprised of several layers and mixtures of pigments. As discussed above in more detail, future work will involve adapting the FP method so that synthetic spectra can be generated that model these phenomena, as well as expanding beyond the set of Renaissance pigments studied within this research. It is anticipated that training a new classification model on this synthetic dataset and applying transfer learning following the method in this paper will ensure the resulting classification model will be optimised for use in galleries and museums.

Abbreviations

XRF: X-ray fluorescence; CNN: Convolutional neural network; FP: Fundamental Parameters; MA-XRF: Macro X-ray fluorescence.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40494-022-00716-3.

Additional file 1: Table of commercial pigments, binding media and substrates used in the collection of real XRF spectra. Plot of XRF spectra of various element standards (Bruker) collected under typical conditions.

Acknowledgements

ND and CH would like to particularly thank Marika Spring for her help in preparing this manuscript and to extend thanks to Marika Spring and other colleagues in the Scientific Department at the National Gallery for useful discussions regarding historic pigment composition. They also gratefully acknowledge Jill Dunkerton in the Conservation Department and Isabella Kocum in the Framing Department at the National Gallery for providing access to additional reference materials for measurement. All authors would like to thank those involved in the EPSRC-funded ARTICT project for their insight and feedback on this research as well as John Delaney (National Gallery of Art, Washington D.C.) for insightful discussions.

Author contributions

MR and CH designed the research. ND and CH collected XRF datasets on pigments and paintings at the National Gallery. CJ and MR implemented the CNNs after using the FP method to generate synthetic XRF spectra of pigments. All authors assessed the results and guided improvements to the methodology. CJ and ND wrote the manuscript with contributions and editing from CH and MR. All authors read and approved the final manuscript.

Funding

This work is supported by EPSRC grant EP/R032785/1, which also directly supported the postdoctoral research fellowships of CJ and ND.

Availability of data and materials

Data are available from the authors upon reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details
[1]Department of Electronic and Electrical Engineering, University College London, Malet Place, London WC1E 7JE, UK. [2]National Gallery, Trafalgar Square, London WC2N 5DN, UK.

Received: 7 March 2022 Accepted: 15 May 2022
Published online: 13 June 2022

References

1. Shugar NA, Mass LJ. Handheld XRF for art and archaeology. Studies in archaeological sciences. Leuven: Leuven University Press; 2012.
2. Romano FP, Janssens K. Preface to the special issue on: MA-XRF "developments and applications of macro-XRF in conservation, art, and archeology" (Trieste, Italy, 24 and 25 September 2017). X-Ray Spectrom. 2019;48(4):249–50.
3. Dik J, Janssens K, Van der Snickt G, van der Loeff L, Rickers K, Cotte M. Visualization of a lost painting by Vincent van Gogh using synchrotron radiation based X-ray fluorescence elemental mapping. Anal Chem. 2008;80(16):6436–42.
4. Alfeld M, Vaz Pedroso J, van Eikema HM, Van der Snickt G, Tauber G, Blaas J, et al. A mobile instrument for in situ scanning macro-XRF investigation of historical paintings. J Anal At Spectrom. 2013;28(5):760–7.
5. Ravaud E, Pichon L, Laval E, Gonzalez V, Eveno M, Calligaro T. Development of a versatile XRF scanner for the elemental imaging of paintworks. Appl Phys A. 2015;122(1):17.
6. Romano FP, Caliri C, Nicotra P, Di Martino S, Pappalardo L, Rizzo F, et al. Real-time elemental imaging of large dimension paintings with a novel mobile macro X-ray fluorescence (MA-XRF) scanning technique. J Anal At Spectrom. 2017;32(4):773–81.
7. Van Espen P, Nullens H, Adams F. A method for the accurate description of the full-energy peaks in non-linear least-squares analysis of X-ray spectra. Nucl Instrum Methods. 1977;145(3):579–82.
8. Vekemans B, Janssens K, Vincze L, Adams F, Van Espen P. Analysis of X-ray spectra by iterative least squares (AXIL): new developments. X-Ray Spectrom. 1994;23(6):278–85.
9. Ryan CG, Etschmann BE, Vogt S, Maser J, Harland CL, van Achterbergh E, et al. Nuclear microprobe—synchrotron synergy: towards integrated quantitative real-time elemental imaging using PIXE and SXRF. Nucl Instrum Methods Phys Res Sect B. 2005;231(1):183–8.
10. Solé VA, Papillon E, Cotte M, Walter P, Susini J. A multiplatform code for the analysis of energy-dispersive X-ray fluorescence spectra. Spectrochim Acta Part B. 2007;62(1):63–8.
11. Alfeld M, Janssens K. Strategies for processing mega-pixel X-ray fluorescence hyperspectral data: a case study on a version of Caravaggio's painting Supper at Emmaus. J Anal At Spectrom. 2015;30(3):777–89.
12. Conover DM. Fusion of reflectance and x-ray fluorescence imaging spectroscopy data for the improved identification of artists' materials. PhD thesis. George Washington University. 2015.
13. Yan S, Huang JJ, Daly N, Higgitt C, Dragotti PL. When de Prony met Leonardo: an automatic algorithm for chemical element extraction from macro X-ray fluorescence data. IEEE Trans Comput Imaging. 2021;7:908–24.
14. Martins A, Albertson C, McGlinchey C, Dik J. Piet Mondrian's Broadway Boogie Woogie: non invasive analysis using macro X-ray fluorescence mapping (MA-XRF) and multivariate curve resolution-alternating least square (MCR-ALS). Herit Sci. 2016;4(22):1–16.
15. Martins A, Coddington J, Van der Snickt G, Driel B, McGlinchey C, Dahlberg D, et al. Jackson Pollock's Number 1A, 1948: a non-invasive study using macro-x-ray fluorescence mapping (MA-XRF) and multivariate curve resolution-alternating least squares (MCR-ALS) analysis. Herit Sci. 2016;4(33):1–13.
16. Kogou S, Lee L, Shahtahmassebi G, Liang H. A new approach to the interpretation of XRF spectral imaging data using neural networks. X-Ray Spectrom. 2020;50(4):310–9.
17. Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. In: Pereira F, Burges C, Bottou L, Weinberger K, editors. Proceedings of the 25th international conference on neural information processing systems. Red Hook: Curran Associates, Inc; 2012. p. 1097–105.
18. Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, editors. Computer vision-ECCV 2014. Cham: Springer; 2014. p. 818–33.
19. Kleynhans T, Schmidt Patterson CM, Dooley KA, Messinger DW, Delaney JK. An alternative approach to mapping pigments in paintings with hyperspectral reflectance image cubes using artificial intelligence. Herit Sci. 2020;8(1):1–16.
20. Sherman J. The theoretical derivation of fluorescent X-ray intensities from mixtures. Spectrochim Acta. 1955;7:283–306.
21. Alfeld M, Van der Snickt G, Vanmeert F, Janssens K, Dik J, Appel K, et al. Scanning XRF investigation of a flower still life and its underlying composition from the collection of the Kröller–Müller Museum. Appl Phys A Mater Sci Process. 2013;111(1):165–75.
22. Alfeld M, Gonzalez V, van Loon A. Data intrinsic correction for working distance variations in MA-XRF of historical paintings based on the Ar signal. X-Ray Spectrom. 2020;50(4):351–7.
23. Pan SJ, Yang Q. A survey on transfer learning. IEEE Trans Knowl Data Eng. 2010;22(10):1345–59.
24. Weiss K, Khoshgoftaar TM, Wang D. A survey of transfer learning. J Big Data. 2016;3(1):9.
25. Cheng PM, Malhi HS. Transfer learning with convolutional neural networks for classification of abdominal ultrasound images. J Digit Imaging. 2017;30(2):234–43.
26. Pires de Lima R, Marfurt K. Convolutional neural network for remote-sensing scene classification: transfer learning analysis. Remote Sens. 2020;12(1):86.
27. Eastaugh N, Walsh V, Chaplin T, Siddall R. Pigment Compendium. 1st ed. Oxford: Elsevier Butterworth-Heinemann; 2004.
28. van Loon A, Boon JJ. Characterization of the deterioration of bone black in the 17th century oranjezaal paintings using electron-microscopic and micro-spectroscopic imaging techniques. Spectrochim Acta-Part B. 2004;59(10–11):1601–9.
29. Wu Q, Döbeli M, Lombardo T, Schmidt-Ott K, Watts B, Nolting F, et al. Does substrate colour affect the visual appearance of gilded medieval sculptures? Part II: SEM–EDX observations on gold leaf samples taken from medieval wooden sculptures. Herit Sci. 2020;8(119):1-13.
30. Hradil D, Píšková A, Hradilová J, Bezdička P, Lehrberger G, Gerzer S. Mineralogy of bohemian green earth pigment and its microanalytical evidence in historical paintings. Archaeometry. 2011;53(3):563–86.
31. Elias M, Chartier C, Prévot G, Garay H, Vignaud C. The colour of ochres explained by their composition. Mater Sci Eng B Solid-State Mater Adv Technol. 2006;127(1):70–80.
32. Genestar C, Pons C. Earth pigments in painting: Characterisation and differentiation by means of FTIR spectroscopy and SEM-EDS microanalysis. Anal Bioanal Chem. 2005;382(2):269–74.
33. Martin E, Duval AR. Les deux varietes de jaune de plomb et d'etain: etude chronologique. Stud Conserv. 1990;35(3):117–36.
34. Gonzalez V, Gourier D, Calligaro T, Toussaint K, Wallez G, Menu M. Revealing the origin and history of lead-white pigments by their photoluminescence properties. Anal Chem. 2017;89(5):2909–18.
35. Spring M. New insights into the materials of fifteenth- and sixteenth-century Netherlandish paintings in the National Gallery, London. Herit Sci. 2017;5(40):1–20.
36. Klaas J. Die "Ultramarinkrankheit"-Studien zu Veränderungen in ultramarinhaltigen Farbschichten an Gemälden. Ph.D Thesis. Technischen Universität München. 2011.
37. Franceschi E, Locardi F. Strontium, a new marker of the origin of gypsum in cultural heritage? J Cult Herit. 2014;15(5):522–7.
38. Smieska LM, Mullett R, Ferri L, Woll AR. Trace elements in natural azurite pigments found in illuminated manuscript leaves investigated by synchrotron X-ray fluorescence and diffraction mapping. Appl Phys A Mater Sci Process. 2017;123:484.
39. Liang H, Lange R, Peric B, Spring M. Optimum spectral window for imaging of art with optical coherence tomography. Appl Phys B Lasers Opt. 2013;111(4):589–602.
40. Harding M. Paint Colours and Materials Brochure. https://www.michaelharding.co.uk/wp-content/uploads/2017/11/MichaelHarding_Paint_Colours_And_Materials_Brochure.pdf. Accessed 2022 Jan 27.
41. Tyler M. Aspects of the manufacture, trade and history of smalt. Ph.D Thesis. University of Glasgow. 2021.

42. De Boer DKG. Calculation of X-ray fluorescence intensities from bulk and multilayer samples. X-Ray Spectrom. 1990;19(3):145–54.
43. De Boer DKG. Angular dependence of X-ray fluorescence intensities. X-Ray Spectrom. 1989;18(3):119–29.
44. Elam WT, Ravel BD, Sieber JR. A new atomic database for X-ray spectroscopic calculations. Radiat Phys Chem. 2002;63(2):121–8.
45. Wolff T, Malzer W, Mantouvalou I, Hahn O, Kanngießer B. A new fundamental parameter based calibration procedure for micro X-ray fluorescence spectrometers. Spectrochim Acta Part B At Spectrosc. 2011;66(2):170–8.
46. Thinh TP, Leroux J. New basic empirical expression for computing tables of X-ray mass attenuation coefficients. X-Ray Spectrom. 1979;8(2):85–91.
47. Chollet F. Keras. https://www.keras.io/. Accessed 2021 May 4.
48. Ackroyd P, Billinge R, Macaro G, Peggie D, Spring M. Bermejo's saint michael triumphant: restoration, construction and painting technique. In: Treves L, editor. Bartolomé Bermejo: Master of the Spanish Renaissance. London: National Gallery Company Limited; 2019. p. 99–113.
49. Macaro G. NG6553 Bermejo-results of inorganic analysis. Internal report. Scientific Department, National Gallery, London. 2018.
50. Bomford D, Leighton J, Kirby J, Roy A. Impressionism and the modern palette in: Art in the making: impressionism. London: National Gallery Company Limited; 1990. p. 51–72.

Publisher's Note