

RESEARCH ARTICLE

Open Access



Virtual cleaning of works of art using deep convolutional neural networks

Morteza Maali Amiri*  and David W Messinger

Abstract

A varnish layer that is applied to a painting, generally to protect it, yellows over time, deteriorating the original look of the painting. This prompts conservators to undertake a cleaning process to remove the old varnish and apply a new one. Providing the conservators with the likely appearance of the painting before the cleaning process starts can be helpful to them, which can be done through virtual cleaning. Virtual cleaning is simply the simulation of the cleaning process. Previous works in this area required the method to have access to black and white paint regions, or physically removing the varnish first at a few spots. Through looking at the problem of virtual cleaning differently, we try to address those shortcomings. To do so, we propose using a convolutional neural network (CNN) to tackle the problem of virtual cleaning. The CNN is trained on artificially yellowed images of people, urban and rural areas, and color charts, as well as their original versions. The network is then applied to various paintings with similar scene content. The results of the method are first compared to the only physical model in the virtual cleaning field. We compare the outputs from the proposed method and the physical model by visualization as well as a quantitative measure that calculates the spectral similarity between the outputs and the reference images. These results show that the proposed method outperforms the physical model. The CNN is also applied to images of the Mona Lisa and The Virgin and Child with Saint Anne, both painted by Leonardo da Vinci. Results show both a qualitative and quantitative improvement in the color quality of the resulting image compared to their reference images. The CNN developed here is also compared to a CNN that has been developed for the purpose of image colorization in the literature to demonstrate the effectiveness of the CNN devised here, showing that the CNN architecture herein leads to a better result. The novelty of the work proposed herein lies in two premises. First, the accuracy of the method, which is demonstrated through comparison with the only physical approach derived until now. Second is the generalizability of the method which is shown through blindly applying the method to two famous works of art for which no information but an RGB image of the uncleaned artwork is known.

Keywords: Virtual cleaning, Convolutional neural network, Varnished and unvarnished artworks

Introduction

It is a well-known physical phenomenon that the appearance of varnish on the surface of a painted work of art will change over time, altering the visual qualities of the work. There are artists who did not intend for their paintings to be varnished but once the paintings were out of their hands, the artworks were varnished, usually

for protection purposes [1–3]. Although effective in protecting the artwork from dirt and pollutant, varnish application can substantially change the appearance of paintings [4], particularly after the passage of a significant amount of time. The change of the appearance of the painting due to varnishing depends on many factors, the most important of which are the type of varnish (i.e., its molecular weight) and the age of the varnish [4, 5]. Artwork cleaning is considered one of the most significant duties undertaken by conservators due to the irreversibility of the action. Cleaning is comprised of physical

*Correspondence: mm2391@rit.edu
Rochester Institute of Technology, Chester F. Carlson Center for Imaging Science, 54 Lomb Memorial Drive, Rochester, NY 14623, USA

removal of undesired deposits as well as the aged varnish from the surface, which helps reestablish the original appearance of the painted surface [6–8]. There are two main approaches that have been taken to “clean” works of art, physical and virtual. The physical approach, in which mild solvents and gel systems are usually used, apart from being time consuming, can also damage the work [9–11]. The simulation of the outcome of the aged varnish removal from a painting is referred to as virtual cleaning. Virtual cleaning supplies the conservators with a representation of the appearance change that would likely be achieved should the cleaning process be undertaken. Additionally, in some cases where the painting is not likely to go through removal of varnish sometime soon, the simulation becomes even more important as a tool to visualize the original color representation of the work [12]. It is worthwhile mentioning that virtual cleaning can be complementary to the actual physical restoration of artwork, but it cannot replace it in any way. As it was mentioned above, it could give the restorers an idea about how a painting might look should it go through the process of restoration.

There have been many studies in the area of virtual cleaning of works of art, some of the most prominent ones are mentioned here. Barni et al. [13] developed an image processing technique to virtually clean artwork. In their work, first they physically cleaned a small part of the painting. Subsequently, they found the matrix transform in the RGB domain between the cleaned part and the corresponding uncleaned part from the artwork. They then applied the same matrix transformation to the uncleaned parts of artwork and were able to virtually clean the entire piece. Papas and Pitas (2000) developed a function to recover the cleaned version of a work of art from the uncleaned one using a few different approaches [14]. They noted that using the RGB color space of the camera is not suitable as it does not closely correlate with human perception of color. Therefore, they used CIELAB color space in their work claiming it performs better due to its higher visual uniformity [14]. They also had to first physically clean the artwork in a few regions. They subsequently used the mean values of those regions in both cleaned and corresponding uncleaned image. They developed several different transfer functions, of which they showed that their linear approximation (which is based on a linear transformation from a varnished to a cleaned image) and white point

(which is based on the chromatic adaptation taken from color science) have outperformed others [14]. Elias and Cotte (2008) were able to virtually clean the famous Mona Lisa through having access to the pigments and varnish used by Leonardo da Vinci at the time of painting the work of art [15]. To do this, a color chart made out of classical paints utilized in the 16th century in Italy were made in varnished and unvarnished forms. Using these charts, they were able to deduce a mean multiplicative factor for each wavelength. The factor was then applied to the Mona Lisa's spectra leading to the virtual removal of the varnish [15]. Palomero and Soriano (2011) developed a neural network to approach the issue of virtual cleaning for the first time [16]. They trained a shallow neural network with 2 hidden layers and 30 neurons to go from a varnished painting to an unvarnished one. They again had to physically clean a part of the painting. RGB data of the cleaned and the corresponding uncleaned region of the painting were used to train the network. Using estimation methods, they were able to also estimate the spectral reflectance of the varnish layer, supposing that the varnish acts as a filter over the painting.

Trumpy et al. were the first to attempt to approach the virtual cleaning of artwork using a completely physical model for the varnish/painting system [12]. They developed the first physical approach attempting to model the effect of the varnish and obtain the spectra of the cleaned artworks. They started with the Kubelka–Munk theory [17] and developed a method that predicts the spectra of the cleaned painting. To obtain the cleaned spectra, they needed to first estimate the varnish transmittance. In order to do that, they made a few simplifying assumptions. They first assumed that a dark site of the painting consists of a perfect black that absorbs all the incident radiation. (In this paper, by “perfect” and “pure” we mean unmixed with other colors, and not grayish. Therefore any reference to perfect or pure black and white only means black and white regions that are not mixed with other colors, and are not grayish, either.) They also assumed that the varnish spectral reflectance is independent of wavelength. Through physically cleaning the dark and light sites (calling them black and white sites as well) of the painting, they were able to estimate the spectral transmittance of the varnish [12]. According to studies that we have done on the

physical model (the results of that study are not presented here), we concluded that the accuracy of the physical model depends on the purity of both black and white paints of the painting. In other words, the results showed that the physical model relies on the pure black and white to reach the highest accuracy. This model is of great importance to us, as it is the first work trying to lay out a basis model for the varnish/painting system and it is used as a reference to compare our results with.

More recently, Kirchner et al. developed a method based on the Kubelka-Munk theory attempting to virtually clean artworks [18]. In order to do that, they had to physically remove the varnish from parts of the painting and measure the spectral reflectance before and after varnish removal. One of the key measurements was to measure the pure white on the painting through which they were able to compute the transmittance and reflectance of the varnish layer using two constant Kubelka–Munk theory. After characterization of the varnish layer, they were able to digitally clean the full painting [18]. Zeng et al. developed a method based on multi-resolution image analysis and deep CNNs to reconstruct Van Gogh's drawings that have been degraded due to aging [19]. They asserted that due to lack of access to the ground truth data, they chose a digital representation of a present faded drawing and one of its less faded version reproduced in the past as input and target pair, respectively. They used a multi-scale CNN in which the images of the same scene were downsampled at different scales. The differently scaled images were then fed into different networks fit for the appropriate scales. The outputs were then averaged to obtain the final output of the model. They compared their method with another CNN based method and showed that their method has outperformed the former one. Wan, *et al.* (2020) developed a method based on a variational autoencoder to restore old images [20]. In order to do that, they formulated the image restoration as an image translation problem, where they considered clean images and the old ones as images from different domains trying to learn the mapping between them. However, they translated images across three domains: the real image domain, the synthetic image domain (images that suffer artificial degradation) and the ground truth containing images without degradation. They reasoned that directly learning the

mapping function from real images to the cleaned ones is challenging as they are not paired, prompting them to propose decomposing the translation into two stages. Here, these stages included one in which the images are mapped into the latent spaces first with synthetic images, and one in which the old ones share the same latent space. Through learning the image restoration in the latent space, they were able to learn the translation from the latent space of the corrupted images (synthetic and old images) to the latent space of the ground truth through mapping. Comparing their methods to other approaches of restoration, their methods outperformed the prior research. Linhares et al. used hyperspectral imaging to first measure the spectra of two paintings before and after varnish removal. Using this information, they were able to characterize the varnish layer which subsequently allowed for virtual removal of the varnish layer [21]. The need to specify the pure black and white, physical removal of the varnish from the painting, use of spectral reflectance and the inability to generalize the method and results to other works, are only a few shortcomings of the works reported here.

The purpose of this work is to address the shortcomings of the prior methods and find a better approach to virtually clean artwork fulfilling both public and conservator's interests in easily seeing how an old, varnished painting would visually appear without the varnish layer. In order to do that, the virtual cleaning process was approached using deep convolutional neural networks (CNNs), a technique also termed deep learning. CNNs have been applied in different areas of machine learning to solve different problems in image processing [22–25]. Of special interest to us are works in image colorization using deep learning [26–29]. Image colorization refers to the process that changes black and white image into an estimated colored image, trying to recover the original colored scene, especially from old black and white pictures. This process inspired us to use CNNs to go from a varnished work of art (a yellowed image) to an unvarnished one (a color image). To do that, images of rural and urban areas, people, and color charts were artificially yellowed. The CNN is trained to go from the yellowed images to the original colored images. The outputs of the CNN and the physical model reported by [12] are also compared to each other. Comparison of our results to the physical

model shows that the method proposed herein has outperformed this model. Two famous artworks, namely the Mona Lisa and The Virgin and Child with Saint Anne, both painted by Leonardo da Vinci, were also fed into the trained network, resulting in cleaned versions of both comparable to their cleaned versions available in the literature. Using the method proposed here, there is no need to physically remove the varnish and measure the spectral information of the painting. The need to know where the pure black and white paints are located is also erased. This approach also has the potential to be generalized to many types of artwork, given an appropriate training data set. The CNN developed in this work is also compared to the CNN developed in [27] to show the effectiveness of the CNN architecture developed here. It should be noted that training the network with images enables it to learn both spatial and spectral information, while training with single spectra or RGB triplets would only make the system learn the spectral information. In a nutshell, the power of using the CNN is that it learns the spatial patterns and their associated likely color representations (i.e., spectral information) from the training images, and then modifies the testing image accordingly.

This paper is laid out in the following manner. "Methodology" section describes the data sets used in the work, the convolutional neural network proposed, and the experiments performed. "Results and discussion" section presents the results and discussions in which the results of our method are going to be compared with the physical approach proposed by [12] along with the results obtained from applying the method to the Mona Lisa and The Virgin and Child with Saint Anne. We end our paper with conclusions summarizing the paper's contributions and outcomes and the future path for our research.

Methodology

This section first describes the data used to train and test the algorithm. The two major parts, namely, spectral and color simulations are expanded upon and described. The criteria used to evaluate the success of the method proposed here are also presented and explained.

Data

In this work, the problem of virtual cleaning of works of art is looked upon as a machine learning problem, in which a system (a CNN herein) learns to go from an

uncleaned artwork to a virtually cleaned one. The use of the CNN approach for this problem requires a set of both training and testing data for the learning process. Here, the hypothesis is that the training data does not need to be a combination of cleaned and uncleaned works of art, but instead can be images of similar content (i.e., people, landscapes, buildings, etc.) in both "cleaned" and "yellowed" states. As described below, the "yellowed" data is simulated using yellow filters. To train the networks, urban and rural images taken from the Kaggle website [30], along with images of people and color charts (described below) were used to train the CNN. Some of the images are shown in Fig. 1. We add the color charts as they represent a large range of colors which might not have been represented in the other images, despite lacking the spatial information of the images. To ensure our training data covers a wide range of colors, the color charts are simulated and added to the dataset as well.

It should be noted that the color charts are simulated using 1269 spectral reflectances of Munsell chips from the Munsell Book of Color Matte Finish Collection [31], 264 spectral reflectances of ANSI IT8.7/2 Standard Chart from Kodak, 1950 samples from the Natural Color System, 130 spectral reflectances of Artist's paint, and 24 spectral reflectances from the Macbeth ColorChecker. These datasets are the spectral data imitating natural scenes and human skin tones, and are made of matte paint, chips and the like. In terms of colors, these charts are designed to represent a broad range of common colors. Munsell and NCS themselves are built based on two famous color systems covering a very broad range of colors [32]. Other color charts are mostly used for digital camera characterization, but are also designed to represent a large range of colors. Therefore, there would not be any limitation that could pose a significant problem for this work. All of the samples were simulated into a 4 by 6 color chart, the same as the Macbeth ColorChecker with 24 chips. To make the simulated color charts, the spectral reflectances were first transformed to CIEXYZ using standard formulae and then the CIEXYZ was transformed to sRGB [32]. Reflectance spectra are converted to CIEXYZ tristimulus values using

$$\begin{aligned} X &= k \sum_{\lambda} S_{\lambda} R_{\lambda} \bar{x}_{\lambda} \Delta\lambda & Y &= k \sum_{\lambda} S_{\lambda} R_{\lambda} \bar{y}_{\lambda} \Delta\lambda \\ Z &= k \sum_{\lambda} S_{\lambda} R_{\lambda} \bar{z}_{\lambda} \Delta\lambda \end{aligned} \quad (1)$$

where, X , Y and Z show the tristimulus values, k denotes the normalizing factor, S shows spectral radiance of the light source (D65 standard illuminant was used herein), \bar{x}_λ , \bar{y}_λ , and \bar{z}_λ represent the color matching functions of the standard observer and λ denotes the wavelength. We convert the CIEXYZ tristimulus values to sRGB using

$$\begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix} = \begin{bmatrix} 3.2404542 & -1.5371385 & -0.4985314 \\ -0.9692660 & 1.8760108 & 0.0425560 \\ 0.0556434 & -0.2040259 & 1.0572252 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (2)$$

These linear RGB values are then gamma corrected as

$$f(x) = \begin{cases} 12.92u, & \text{if } u < 0.0031308 \\ 1.055u^{1/2.4} - 0.055, & \text{otherwise} \end{cases} \quad (3)$$

where, u represents RGB_{linear} . If any of the sRGB values were greater than 1, they are clipped to be equal to 1, and for values less than zero, they are clipped to be equal to zero. The problem of out-of-gamut colors are of little importance to us as all the images used are sRGB which does not impact the main purpose of this work. The final images were saved in jpg format. There are 154 simulated color charts overall and 488 images of people, urban, and rural areas.

As it is described later in section , there are two major parts to this work. The first is called spectral simulation, in which only the color charts (described above) are used. The second part is called color simulation. Many color charts are simulated and the color charts are chosen for use in the second part based on the different number of colors they contain. The more different colors they have, the higher their likelihood to be picked out to be one of the 22 images used in the second part. As mentioned above, we use the color charts to represent a wider range of colors that might not be found in the other images in our dataset. Therefore, we did our best to pick the color charts that represent diverse colors. For the second part of the work 488 images of urban and rural areas are combined with 22 images of the color charts as training samples to virtually clean specific works of art.

Procedure

This paper aims to virtually clean artwork which allows everyone, including both art enthusiasts and conservators, to readily see how an old, varnished painting would appear without varnish, referred to as virtual cleaning of works of art. To do this, a

convolutional neural network is proposed. The proposed method is trained to go from yellowed “varnished” images to color “unvarnished” images of artworks. The varnished and unvarnished samples are simulated either in the spectral domain, under the section called spectral simulation, or in the RGB color domain, under the color simulation, both explained below. To be able to compare our method to the physical model, in which the works of art are virtually cleaned in the spectral domain, we also simulate our samples in the spectral domain. The other approach to simulate the cleaned and uncleaned samples is applied in the RGB color domain hence color simulation. In that section, the only data available to virtually clean the artwork is the RGB image of the uncleaned artwork. There is no access to any other type of data (such as the spectral data of the black and white spots on the painting etc.), making this method a viable option to be applied to famous artwork for which having access to the cleaned version of the artwork, even at few spots, is simply impossible. Therefore, the color simulation section shows the generalizability of the method proposed in this work and how applicable it is to other artworks. Comparing the physical model to the results obtained in the color simulation section is simply not possible due to the limitation of the physical model, or any other method mentioned in the introduction section, as they are reliant on the spectral data on at least black and white spots of the painting. The performance of the CNN developed in this work is also compared to the network developed by [27] to show the effectiveness of the network developed in this work. This comparison is done in the color simulation section. The dataset used by both networks are the same. It should be noted that the network proposed by [27] tackles the image colorization problem with the network learning to go from a black and white image to a colorful image. However, in this work, this network is used to go from yellowed images to colorful ones, in other words, the input and output to the model is simply changed to yellowed images and colored images instead of black and white images to colored images. Everything else in the network proposed by [27] is the same as what was reported by the authors.

Spectral simulation

As mentioned above, spectral simulation here refers to simulating cleaned and uncleaned samples in

the spectral domain. For us to be able to draw a fair comparison between our method and the physical model proposed by [12], we simulate the samples in the spectral domain with the light matter interaction model in mind. To do that, we consider the Macbeth ColorChecker as an experiment to assess the feasibility of the CNN method and compare it to the physical model proposed by [12]. The reason for choosing Macbeth is it has pure black and white spectra along with other colors. The physical model relies on these areas, and also does not utilize any spatial information, allowing us to use the Macbeth ColorChecker as a test object.

To simulate the interaction of light with varnish/painting system, Fig. 2 is used as a guide.

Here, R_t , R_V , and R_P denote the spectral reflectances of the uncleaned artwork (or the total effective reflectance of the varnished work of art), varnish and the paint, respectively. T also denotes the transmittance spectra of the varnish. It should be noted that R_P here is the Macbeth ColorChecker reflectance spectra. It goes without saying that the cleaned spectra are contained in R_P . The measurement geometry is 45/0 (as it was the case in [12] as well), referring to the method of measurement (illumination/viewing) in which the illumination is at 45 degrees off axis and the observer is at 0 degrees. Using the same approach as [12], the transmittance and the reflectance spectra of the varnish are presumed to be as shown in Fig. 3. The spectral reflectance shown is simply the yellow spectrum in the Macbeth ColorChecker multiplied by a factor to make it match the small value of reflectance of the body reflectance of the varnish, as reported by [12]. The spectral transmittance is simply a logarithmic function, mimicking what was reported in [12] regarding the spectral transmittance of varnish.

In general, the varnish is yellow and the transmittance and reflectance spectra of the varnish should represent that [12, 18]. The reflectance of the varnish is very low, as [12] assumed that the body reflection of the varnish is equal to the black reflectance spectrum covered with varnish (or R_t measured over a dark area). Here, we use the Macbeth ColorChecker (and other simulated color charts) as an experiment to assess the feasibility of the CNN method and compare it to the physical model proposed by [12] to understand how well they predict the cleaned spectra of the Macbeth color chart.

It should be noted that the equation shown in Fig. 2 is used to obtain the spectral reflectance of the uncleaned samples and then those spectra, along with the cleaned spectra, are changed to sRGB as explained in section . Using 1269 spectral reflectances of Munsell chips, 264 spectral reflectances of ANSI IT8.7/2 Standard Chart from Kodak, 1950 samples of the Natural Color System, and 130 spectral reflectances of Artist's paint [33], 154 color charts are simulated in a way that each color chart has 24 chips in the same arrangement as in the Macbeth ColorChecker. Using all these samples we are able to make a fairly large set of training data, using which the CNN learns the transformation from yellowed images to images in their original color. After training the CNN, the Macbeth color chart is used to test the CNN and see how well the network can recover the original color chart from the yellowed one. At this point, we also compare the results to the results obtained from the physical model [12] using a Macbeth ColorChecker yellowed at the same level as that of the CNN input.

Color simulation

The second part of the paper focuses on simplifying the work, so that it could be applied to any RGB image of a piece of art with the goal of virtually “cleaning” it. In other words, the simulation of the cleaned and uncleaned training samples is done in the RGB color domain, hence the use of color simulation for this section. In order to do that, images of urban and rural areas, people, and color charts are combined to form a dataset of 500 images which are used as training data. Now, because we do not have access to the spectral information of all of these images, we cannot yellow them the same way as in Section . Therefore, we developed yellow filters, as shown in Fig. 4, which are multiplied point by point with all the 500 images, resulting in the yellowed (or simulated “varnished”) images.

As shown in Fig. 4, three levels of these filters are used, slightly, moderately, and highly yellow, notionally representing various degrees of aging. The RGB counts for the least yellow to the yellowest filter are $RGB = (255, 255, 179)$, $(255, 255, 128)$ and $(255, 255, 77)$, respectively. Our aim here is to train a CNN so that it can estimate the original colored images from the yellowed samples. After training, we apply this network to two famous works of art with images available from the internet, namely Mona Lisa and The

Virgin and Child with Saint Anne, taken from [15, 34] both painted by Leonardo da Vinci. To capture the image of Mona Lisa, they used a multispectral camera of high resolution using 13 filters with 40 nm band-pass [15]. 10 filters were located in the visible part of the spectrum and three in the infrared. The spectral reflectance of the Mona Lisa was captured using this multispectral camera with lighting that was made of halogen lamps with the $57^\circ/0^\circ$ illumination/receiver geometry. The spectral reflectances measured were then changed to a format of RGB image to present them in their paper. The same RGB picture is also used in this work. It should be noted that the Mona Lisa was only virtually cleaned using the technique described in [15]. The image of The Virgin and Child with Saint Anne was obtained from the database reported in [34], which is presented by the department of paintings of the Louvre Museum in Paris. An interesting point about these paintings is that their cleaned version is also available, so we can feed the varnished (uncleaned) versions through the network and compare the results to the cleaned versions as “ground truth” for evaluation. To do the training, first a judgment should be made about how yellow the filter applied to the training data should be. We chose the moderately and highly yellowed filter for The Virgin and Child with Saint Anne and Mona Lisa, respectively, after visually judging the cleaned and uncleaned artworks. The CNN is then trained on the yellowed and cleaned images in the training set, and after that, the artworks are fed into the network as a test to see how well the approach works. Finally, the results are compared to the physically cleaned versions of the artworks using pixel-level spectral similarity measurements. As noted above, the cleaned version of Mona Lisa is actually virtually cleaned [15]. To do so, they used the same set of pigments used by Leonardo da Vinci and then varnished them with the same varnish as he used. After that, they estimated the matrix transform between the varnished pigments and those of the unvarnished ones and applied the same matrix to the uncleaned version of Mona Lisa. Their approach of virtually cleaning works of art possesses an acceptable level of accuracy. Therefore, the work presented by [15] is not used to draw a comparison between our method and theirs, but because their virtual cleaning output is very accurate due to having access to pigments and varnish used by the

original artist, we use their output Mona Lisa as our reference cleaned version of Mona Lisa and try to replicate that. It is obvious that comparing our method to their method does not make sense, as they have access to information about the painting that we did not have access to. Also, we do not always have access to the type of information they had access to, therefore, choosing reference [12] to compare our method to makes more sense (the comparison is made in the spectral simulation section as explained before), as they have developed a physics-based approach which could be applied to many other artworks (although it does need have access to the painting’s black and white spectral reflectance, it is still more realistic than having access to the original pigment and varnish used by the artist). Our method does not need to have access to any type of information about the artwork either, making it applicable to many other artworks too. We also note that the two artworks themselves are not used to train the network. They are only used to test the performance of the trained network.

Convolutional neural network

Architecture

Keras, a library written in python for the purpose of deep learning, was used in this work [35]. A deep convolutional neural network (CNN) was built here with 11 layers. The architecture of the network is shown in Fig. 5.

Kernels, also referred to as filters, move over the input image and extract a specific set of features. Here, the size of the filters was arbitrarily chosen to be 3×3 , and the method was experimentally shown to be invariant to filter size. The amount by which the filter shifts over the input image is referred to as stride. If, for examples, the filter shifts one pixel at a time, it is said that the stride has been set to 1 and so on. A stride of 2 was used here which results in a dimension reduction to half that of the input after each filter operation. It was found out in this work, that using a stride of 2 works slightly better than the max-pooling layer which also results in a reduction in dimension. ReLU (Rectified Linear Unit) is a piecewise linear function that will output the input with no change if it is positive, else, it will output zero. Here, ReLU is used as the activation function. The stride 2, as mentioned, leads to the down-sampling of the

image, which is compensated for by using up-sampling later on in the network. In other words, the input and output have the same size. It should be noted that the size of the image input to the network should be 400×400 . 500 training images of urban and rural areas along with people and color charts are used to train this network. The training data was divided into two subgroups, 70% was used for training the network and 30% was used for validation. The learning rate and batch size were chosen to be 0.01 and 1, respectively, chosen based on trial and error. The number of feature maps along with their size are all specified in Fig. 5. As it was mentioned, the architecture design of this network was inspired by the CNNs that have been applied in the field of image colorization. In particular, the idea of first downsampling the image and then upsampling it was adapted from [29], however, we modified the network to use an RGB image as input instead of a black and white (usually L^* channel), as is the case for image colorization.

Application of the CNNs

As noted above, the CNN used here is restricted to an input image of size 400×400 pixels. Consequently, we need to adjust the input image to fit this CNN architecture. In so doing, in general, images are spatially blurred when going through the CNN, particularly if the size of the image is much bigger than the CNN input size. Also, the feature extraction blurs the output image from the CNN. Here we describe our method to prevent the image from becoming blurred, making the CNN model derived in this work applicable to images of varying size, while also maintaining the original size of the image. Although cropping could have been applied, we were more interested in developing a method that could feed the whole image in one setting as this expedites the process. Fortunately, we can leverage existing work from JPEG compression. In order to do that, before inputting the RGB image into the network, it is first changed into a CIELAB representation, assuming that the RGB image follows the sRGB formula, and its L^* channel is preserved and set aside. After that, the original RGB image is fed into the CNN. The output of the CNN, which is also an

RGB image, is then changed to CIELAB and its a^*b^* channels are extracted. These a^*b^* channels are then concatenated with the L^* of the input image that was set aside. This new CIELAB image, which has been partially transformed by the CNN, is then changed back to RGB. Using this simple method (setting aside L^* and then combining it with a^*b^* of the output), the sharpness of the image is maintained and the approach can be applied to images of a varying size. Fig. 6 shows this process schematically.

This process is adapted from JPEG image compression which is referred to as “visually lossless compression”, meaning one could compress the color channels significantly without noticeable changes as long as the luminance channel is left untouched [36].¹ Human vision is more sensitive to the L^* channel than to the chroma (a^*b^*) channels. The L^* channel carries information about the high frequency data of the image, such as edges in the image. Therefore, preserving L^* helps preserve those edges and consequently keep the image sharp.

It is important to note that this process is only applied to the test images that are the focus of the virtual cleaning. There is no need to do this for the training images. Therefore, after training the network, when the testing images are fed into the network, they go through the process shown in Fig. 6.

Preserving the L^* helps to improve the accuracy of the final results. The important aspect of this work is to visualize the virtually cleaned artwork as accurately as possible. The CNN is trained on images that are yellowed artificially and their colored versions, and it learns the transformation from the yellowed images to the colored ones. So, the CNN learns two sets of information, spatial and color information. The CNN is not being trained as a “classifier” in this case. It is being used to approximate the non-linear transfer function between the “clean” and “yellowed” example pairs. For example, the CNN learns where there is sky, or humans, or other objects, and at the same time it also learns how to transform them from the yellowed images to the colored ones.

The success rate of the network is computed using the per-pixel Mean Squared Error loss calculated

¹ It should be noted that chroma subsampling was developed first by Alda Bedford to develop color television. He demonstrated that the human vision is highly sensitive to black and white and less sensitive to colors in the middle of the visible spectrum, such as yellow and green and much less sensitive to colors at the end of the spectrum such as red and blue. Taking advantage of this knowledge, he was able to develop a system which discarded the majority of the blue signal and kept the majority of the green and some of the red, which is called chroma subsampling [37].

Table 1 Euclidean distance and SA mean and standard deviation (SD) values between the original and “cleaned” Macbeth color chart

Method	Euclidean distance		SA	
	Mean	SD	Mean	SD
Physical model	0.141	0.044	0.182	0.083
CNN	0.021	0.002	0.014	0.004

between the output and the original colored images. What was important to us is to see how well the output colors look and how realistic they are compared to the original images. To do the backpropagation, the difference (loss) between the output of the network and the original colored image is calculated (in the training dataset) using mean squared error. The optimizer used is RMSprop, which specifies how the loss gradient is used to update the network parameters. [35]. One of the problems of CNNs is their tendency to overfit which usually happens when the number of training data is small. To examine this, the training and validation curve is also presented below.

Evaluation metrics

Visualization of the results is the first method to evaluate the success of the approaches. The per-pixel spectral Euclidean Distance and Spectral Angle (SA) were also calculated between the original (cleaned) image and the “virtually cleaned” image [38] for accuracy evaluation. The color space used is RGB, which can be considered a spectral space with only three dimensions. Each pixel in the image is considered a vector in this color space, with the tip of the vector located at the appropriate point in the color space based on the RGB values. The spectral Euclidean distance is obtained through calculating the Euclidean distance between two pixels in that color space, i.e., between the tips of the two vectors. The spectral angle is calculated between two vectors in the color space and it is reported in radians in the range [0, 3.142]. This method has been used extensively in spectral remote sensing and this convention is followed here. The spectral angle is defined as

$$SA_k = \cos^{-1} \left(\frac{\mathbf{t}_k \cdot \mathbf{r}_k}{|\mathbf{t}_k| |\mathbf{r}_k|} \right) \quad (4)$$

where k represents the k^{th} pixel, \mathbf{t}_k and \mathbf{r}_k represent the two pixels belonging to the test and reference images, and SA_k represents the spectral angle between these two pixels.

Experimental environment

Python 3.6.10 Anaconda, Inc. is used as a base coding environment for the CNN algorithm. More specifically, the CNN codes were written and run in the TensorFlow environment, which was installed onto the Anaconda. TensorFlow is an open-source and free library used for machine learning which can be utilized across a wide range of tasks and has a specific focus on training and deep neural networks. In terms of hardware, the programs are run on a CPU. The CPU used belongs to an ordinary Lenovo laptop ideapad CORE i7, 7th Gen. Because the number of training samples is not particularly large, we are able to use a CPU here. The training of the CNN is performed using the artificially yellowed images and their corresponding colored ones. 800 epochs are used with a batch size of 1 and a learning rate of 0.01. The images used are of varying size and in the format of jpg. To check if there is overfitting in the training process, the training and validation curves are examined (shown below). Visual examination of the outputs in the training and validation samples is also performed to ensure the colors look realistic. MATLAB R2020b, the package of mathematical software was also used for data preparation. The yellowing filter, converting the images into a suitable format to be read by the CNN, and all other evaluation calculations (such as computing Euclidean distance and spectral angle) are performed in MATLAB.

Results and discussions

This section is divided into two subsections, namely, spectral simulation and color simulation, reflecting the two tasks described above. The first subsection is devoted to comparison of our results to the physical model proposed by [12] and only the simulated color charts are used for that matter. The second subsection presents the application of the proposed method to two real works of art, namely, Mona Lisa and The Virgin and Child with Saint Anne. These results show how the method can be generalized to a wide range of artwork with only the need for sufficient training data for the CNN.

Table 2 Euclidean distance and SA mean and standard deviation (SD) values between the reference and “cleaned” artwork using CNN

Method	The proposed CNN				CNN proposed by [27]			
	Euclidean distance		SA		Euclidean distance		SA	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Mona Lisa	0.0371	0.0024	0.1489	0.0209	0.0619	0.0099	0.1664	0.0349
The Virgin and Child with Saint Anne	0.0548	0.0055	0.0415	0.0103	0.0711	0.0095	0.1461	0.0251

Spectral simulation

To check if the network is overfitting the data, the training and validation loss are presented in Fig. 7.

As it is observed from Fig. 7, the error is descending for both cases of training and validation, and converges showing the lack of overfitting on the dataset. The accuracy of the results obtained when the method is blindly applied to two famous artworks could also be used to see if there was any overfitting. The higher the accuracy in that case, the smaller the chance of overfitting.

Fig. 8 shows a few of the color charts simulated in this work along with their yellowed versions. The Macbeth ColorChecker is shown on the far right and is used as a testing sample for the CNN model built in this work. The top row shows the colored (“unvarnished”) charts and the bottom row shows the yellowed (“varnished”) charts. The Macbeth ColorChecker is also used to test the physical model. Therefore, the same color chart with the same level of yellowness is used for both methods, making the comparison between different approaches possible.

As mentioned before, the Macbeth color chart is used as an experiment to assess the feasibility of the CNN method and compare it to the physical model proposed by [12]. Fig. 9 shows the results of the method proposed here using the CNN along with the result from [12], referred to as the physical model. It should be noted that the output of the physical model is spectral reflectance. Consequently, the reflectance was converted to CIEXYZ and then CIEXYZ was converted to sRGB.

Looking at Fig. 9, it is clear that the CNN has outperformed the physical model in a significant manner. To have a better understanding of the results and how they are compared in a quantifiable manner, the Euclidean distance and SA between the original Macbeth color chart and the output of the CNN

and physical model are also calculated and shown in Table 1. It should be noted that the values reported are the averaged values across the whole image; in other words, the Euclidean distance and the SA are calculated per-pixel for the color chart and the output of CNN and the physical method and then the mean value is calculated and reported.

Table 1 shows the difference between the CNN and the physical model that was also observed from Fig. 9, demonstrating that the proposed method has outperformed the physical model. The reliance of the physical model on the perfect black and white regions, and the fact that they have assumed that the black spectrum is independent of wavelength, are a reasons why the model does not have a satisfactory output. On the other hand, the CNN makes no such presumptions and simply learns to model the transformation from the input image (yellowed image here) to the output image (original colored image) making it a suitable approach to the problem of virtual cleaning, in which there are many unknowns. One of these is the complexity of the relationship between the paint and the varnish, which needs a far more sophisticated physical model to be understood thoroughly. The need to physically remove a portion of varnish from the painting is also one more impediment in the physical approach making it less suitable to be applied to precious artworks. Here, the CNN learns the combined effects of the transformation and is able to estimate the original image accurately without having to physically remove any part of the varnish.

Color simulation

Here we demonstrate the generalizability of the approach to works of art without requiring any information to be known about the artwork, and with only an RGB image of the uncleaned artwork. In this section, we were not able to compare our method to any

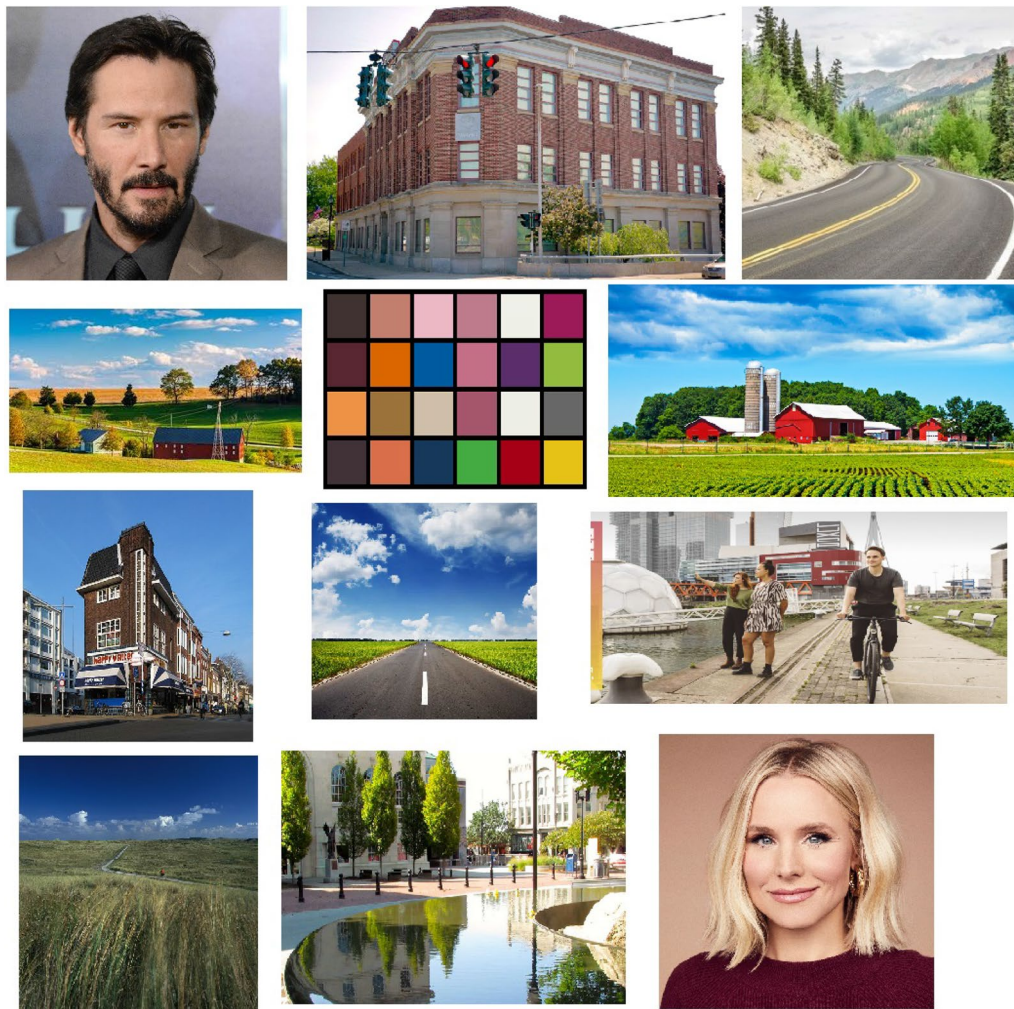


Fig. 1 Some examples of the images used for training in this work

45/0 geometry of measurement

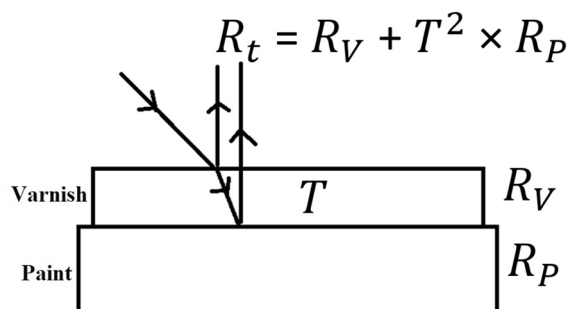


Fig. 2 Interaction of light with varnish/painting system

other model as those models require some other data to be known. In other words, other models are not as generalizable as the method proposed herein. Instead here we compare our CNN to the CNN proposed by [27] which they developed to tackle the issue of image colorization. 500 images in our training samples are all filtered with a yellow filter producing simulated “aged” images. These artificially yellowed images are used to train the network. The network is then used to virtually clean the real, degraded images of artwork, shown in Fig. 10a and e. To filter the training images in order to yellow them, there are three levels of yellowness that could be chosen, as shown in

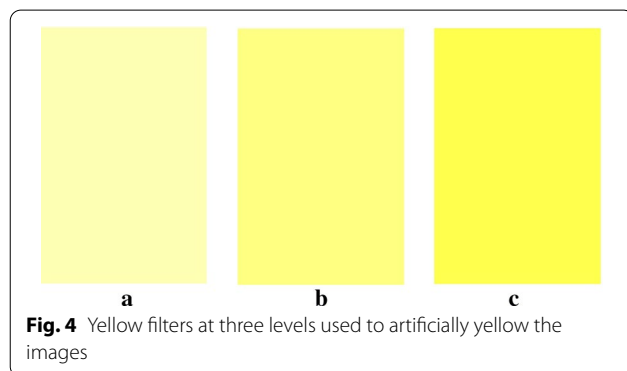
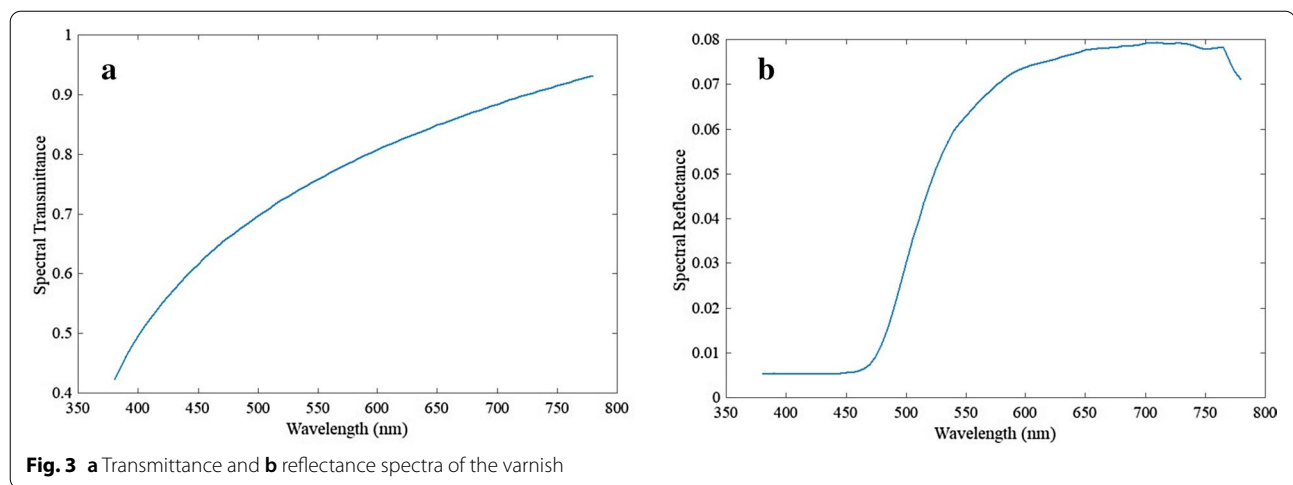
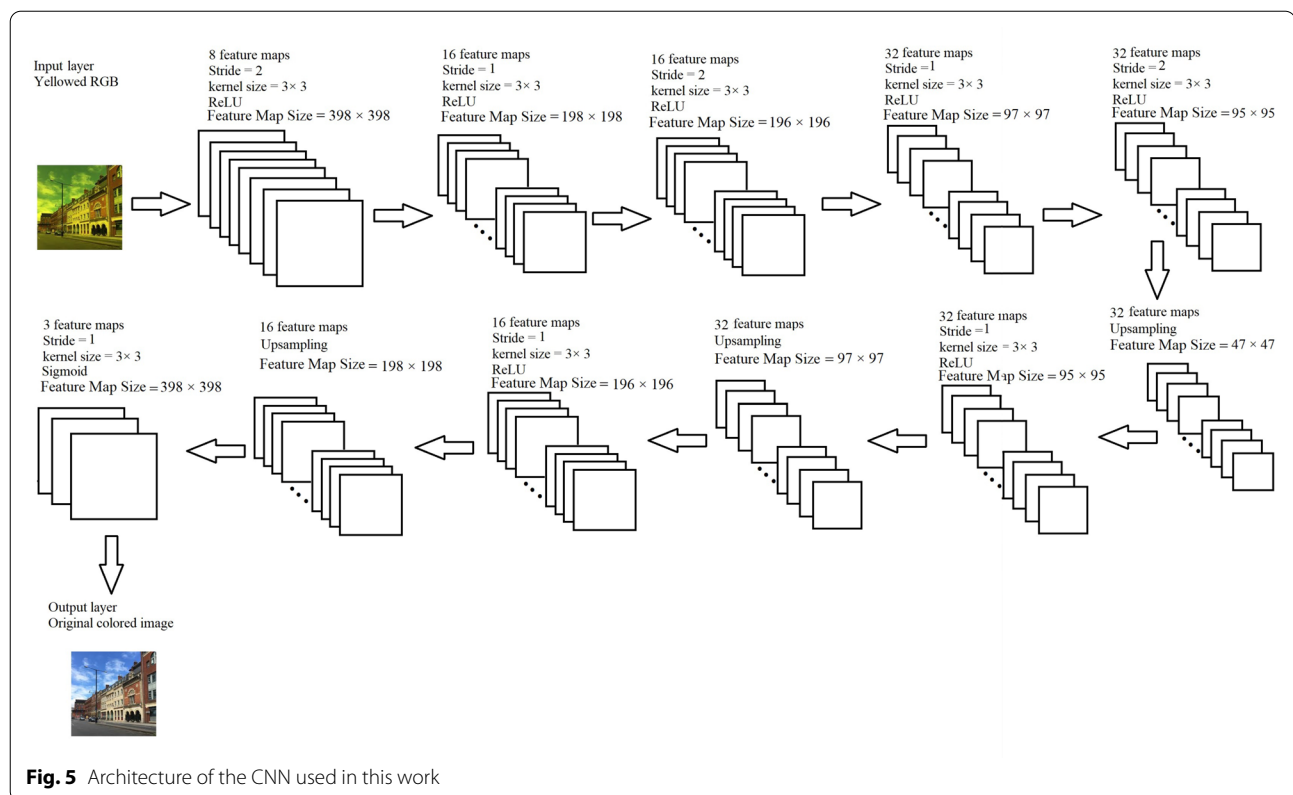


Fig. 4. To choose the level of yellowness of the filters applied to the training data, we first observe the level of yellowness that the test data (the two real artworks herein) have and infer the yellowness level to be chosen. Looking at the Mona Lisa and The Virgin and Child with Saint Anne and how yellow they are compared to their cleaned versions, we chose the moderate and high levels of yellowness (Fig. 4) to make the 500 training images yellow, and the network is subsequently tested on the Mona Lisa and The Virgin and Child with Saint Anne, respectively. Empirically, we note that the closer the images in the training set are to the desired work, in terms of both content and color, the better the results will be. As noted and described above, the cleaned version of the Mona Lisa is actually virtually cleaned by the authors themselves [15]. Fig. 10 shows the uncleaned, cleaned and virtually cleaned versions of the works used in this paper. The cleaned and uncleaned versions of these paintings

have been taken from [15, 34]. The virtually cleaned images are the results of application of the proposed CNN described here and the one proposed by [27].

As observed from Fig. 10, compared to the reference images (b and f), the proposed method does well in cleaning the artwork from a visual perspective, with The Virgin and Child with Saint Anne having been virtually cleaned at a higher (qualitative) level compared to Mona Lisa. When one zooms into the output of the CNN method proposed here, they will come across an artifact at the image boundary. This artifact might have been caused by the CNN. Considering that the CNN had no access to any other information but only an RGB image of the uncleaned artwork, this artifact is only a minor issue, as the overall result, as shown in this section, is at a very satisfactory level. Comparing our network's output to that of reference [27] shows clearly that our CNN has outperformed the CNN proposed by them.

To have a better, quantitative understanding of the results, the per-pixel Euclidean distance and SA are computed between the virtually cleaned and reference versions (b and f in Fig. 10) of these works and are shown in Fig. 11. It should be noted that the units of Euclidean distance and SA are not the same with Euclidean distance having the same unit as that of the image and SA being in radians. To make the comparison possible, the images are separately normalized through dividing them by the maximum value in the image putting the images in a similar scale ranging from 0 to 1. Here, red shows the largest differences and blue shows the smallest differences between

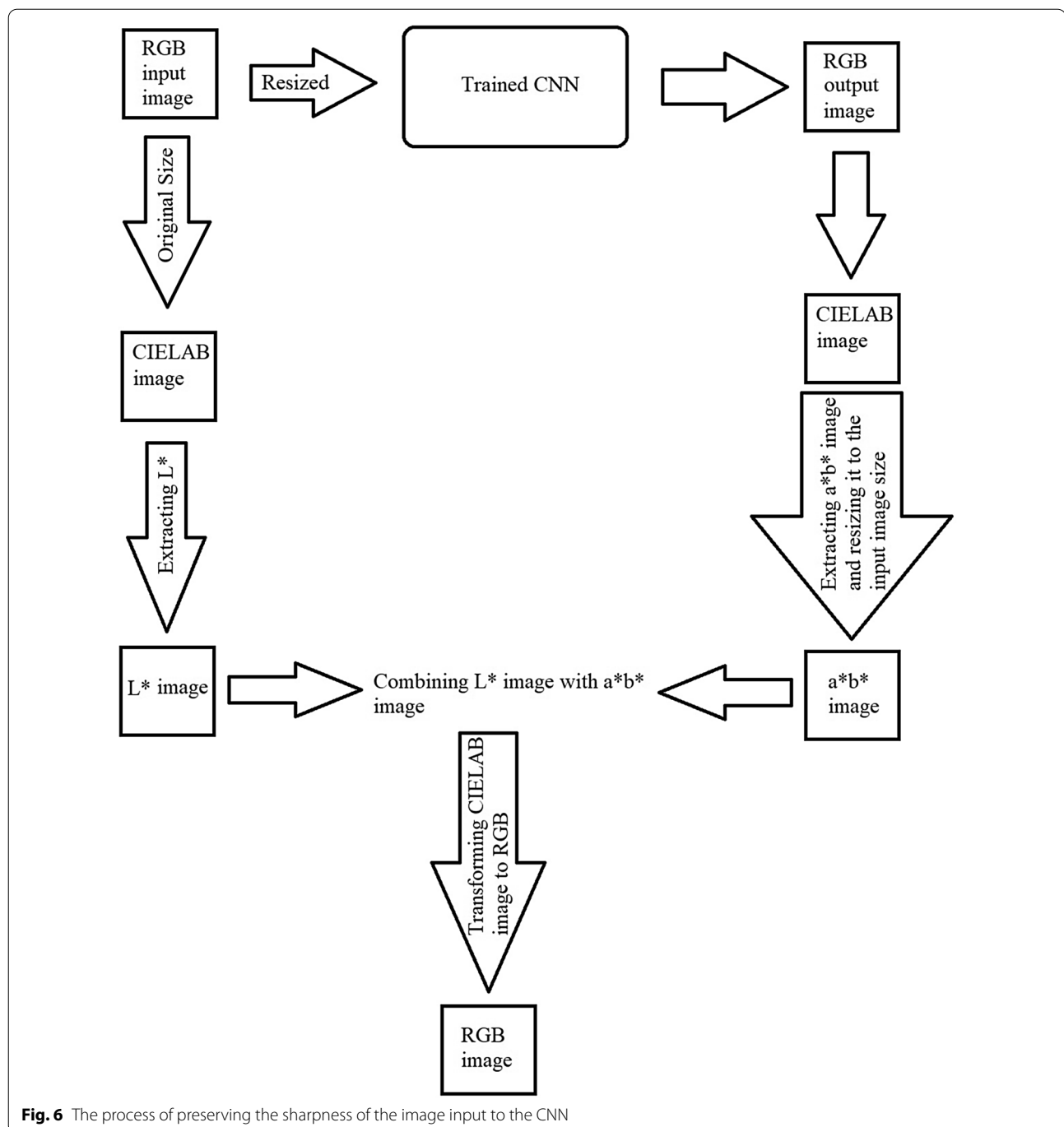


the reference and the output images of the proposed method.

From Fig. 11, we can see that the network has performed better on The Virgin and Child with Saint Anne as compared to the Mona Lisa. It is also observed that the network proposed herein has outperformed the network proposed by [27]. To make it clear which artwork has resulted in a better outcome, Table 2 is reported also, which shows the mean and standard deviation of the SA and Euclidean distance images shown in Fig. 11. Just as in Table 1, this table also reports the overall results using only one number, i.e., mean and standard deviation of SA and Euclidean distance between the output of CNN and the cleaned reference image.

As it is observed from Table 2, Mona Lisa has a lower value of Euclidean distance mean and standard deviation values compared to The Virgin and Child, but it has a much larger value of SA mean and standard deviation, showing that overall the network has

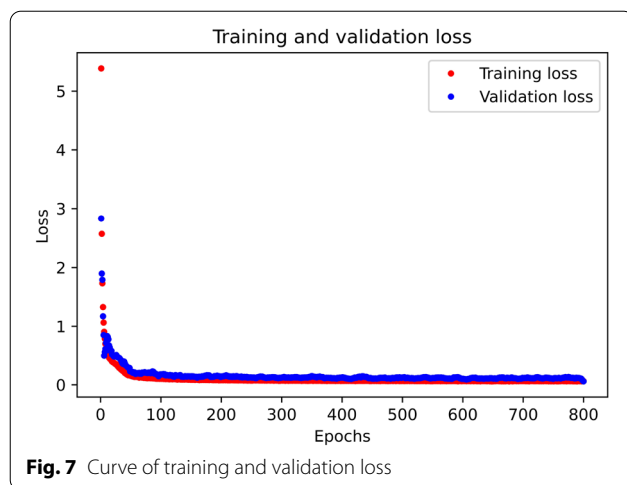
done better on The Virgin and Child with Saint Anne. This difference shows that the CNN, with a common set of training data, works better on some artworks than others, likely dependent on many factors. One reason could be that the varnish might not be the only reason for the color change, and the artwork might have experienced some other factors leading to discoloration which would not have been captured in the yellowing filters used to imitate the effect of the varnish. Also, the CNN learns specific features available in the training data. If it does not see a particular feature during training and is then tested on that feature, the CNN will fail. Therefore, the choice of the training samples and how representative they are of the testing data is of great importance. From table 2, it is also observed that our CNN has performed better than that of the CNN proposed by [27]. The reason for this could be the architecture of our network compared to that of reference [27]. Our network learns to go from yellowed image to colored image in RGB domain,



however, in reference [27] the network goes from a yellowed image in L^* domain to a colored image in a^*b^* domain. Obviously the amount of information is limited in L^* (only one channel) compared to having three channels of RGB which could lead to a less accurate result as shown herein. It is hypothesized that the CNN here has learnt two major features, one related to color and the other related to spatial features, such

as the sky, the human features, rocks, buildings, and so forth. It is interesting to see that the CNN has done better on the Virgin and Child, at least as measured in the Euclidean Distance, on the sky in each image. Again, this is likely a factor of the characteristics of the training data.

By using a convolutional neural network (CNN) with training on simple color images and their artificially



yellowed versions, we were able to virtually clean artworks with no required information about the works. The only data available to the algorithm is an RGB image of the uncleaned work of art. The results as shown in this paper were satisfactory but for a more generalized network, a great deal of attention should be paid to the datasets used to train the network and also to the level of yellowness of the artwork that are aimed to be virtually cleaned. The work proposed in this paper has two main novelties. The first is the high accuracy of the method, which was proven through

drawing a comparison between the method proposed herein with the only physical approach devised until now. The second is the generalizability of the method to different artworks with no additional information about the artwork required, in contrast to all methods devised by now. Having access to the data from at least a few parts of the cleaned and uncleaned artwork and not being able to apply the same method devised from one artwork to the other, are two main shortcomings of the prior approaches that are addressed in this work. Having to physically remove the varnish from the artwork is also another negative point from the prior work. However using the method proposed herein all these shortcomings were addressed. It is worth mentioning that the varnish properties in the case of the real artwork might spatially vary. However, in this work we assumed that the varnish properties, such as spectral transmittance and reflectance, are spatially uniform. It should also be noted that the CNN as trained here works better on realistic looking works. Looking at Fig. 1 it is obvious that the network has been trained on the images of humans, buildings and natural subjects making the network more suitable to work on natural (*i.e.*, realistic) images of art. The network as currently trained and implemented would not likely work well on abstract paintings as well as the realistic looking artworks, although that was not tested in this work.

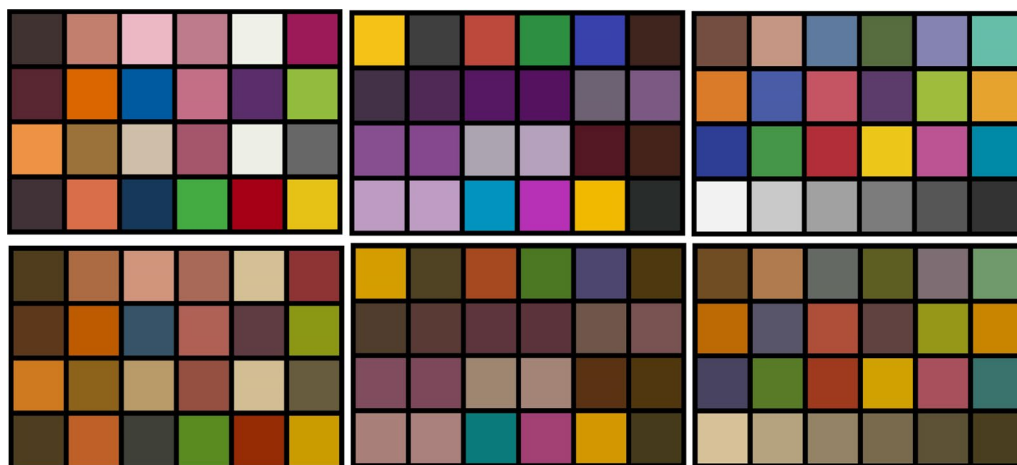


Fig. 8 Simulated color charts (top row) and their yellowed versions (bottom row)

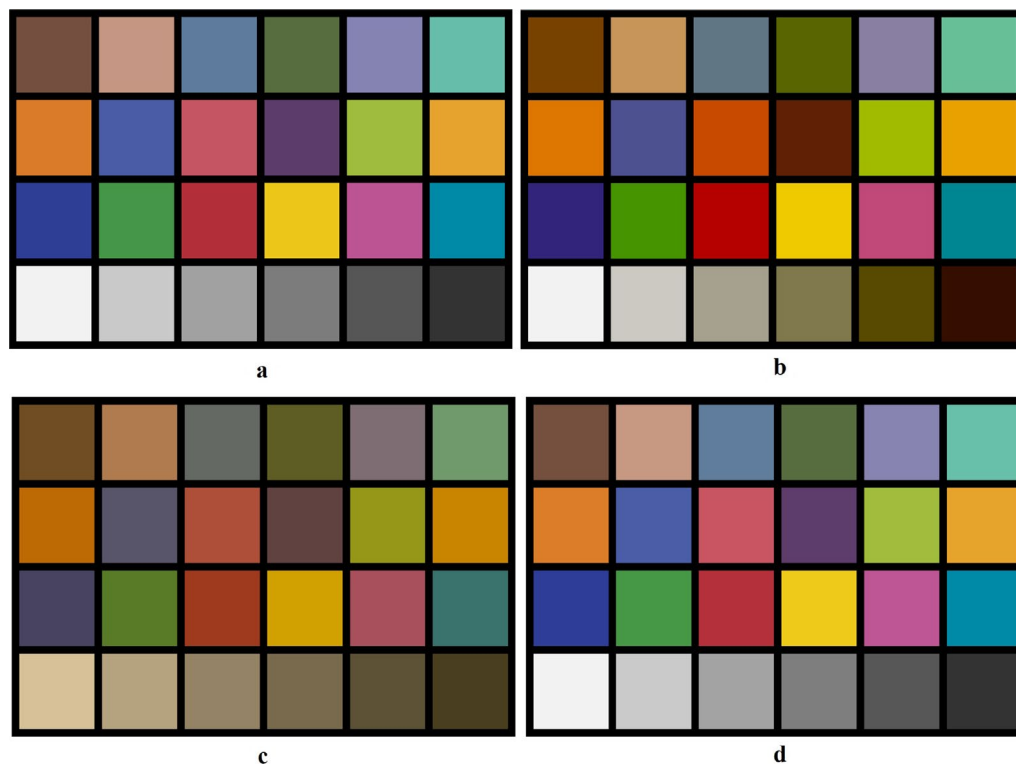


Fig. 9 a original Macbeth, b physical model output, c yellowed Macbeth, d CNN output

Conclusions

The problem of virtual cleaning has been of great interest, especially to conservators and curators in museums, which could help them visualize how a painting would be modified if they cleaned it. However, we note that virtual cleaning can not replicate or replace physical cleaning of artifacts, it is only a complementary approach to produce informative visualization of the work in its estimated original state. Many approaches have been developed and applied to solve this problem, most of which rely on the physical removal of the varnish from different parts of the painting. Reliance on finding black and white spots on the painting is another shortcoming of these approaches. In order to circumvent this problem, a CNN was applied to the issue of virtual cleaning of works of art. In order to do that, RGB pictures of people, rural and urban areas and color charts were

artificially yellowed using yellow filters and the network was trained to go from these yellowed images to the original colored ones. Comparing our work with a physical model showed that the proposed method has outperformed the physical model used to virtually clean the test targets. The method was applied to two famous works of art, namely, the Mona Lisa and The Virgin and Child with Saint Anne, leading to a satisfactory virtual cleaning of both. Comparing our network with a CNN developed to tackle image colorization issue also revealed that our network works better for the purpose of virtual cleaning. The main two novelties of the method proposed herein are the high accuracy, proven through comparing the method to the only physical based approach, and the generalizability of it, which was shown through applying the method blindly to the two famous works of art. For the latter test, the only piece of information available

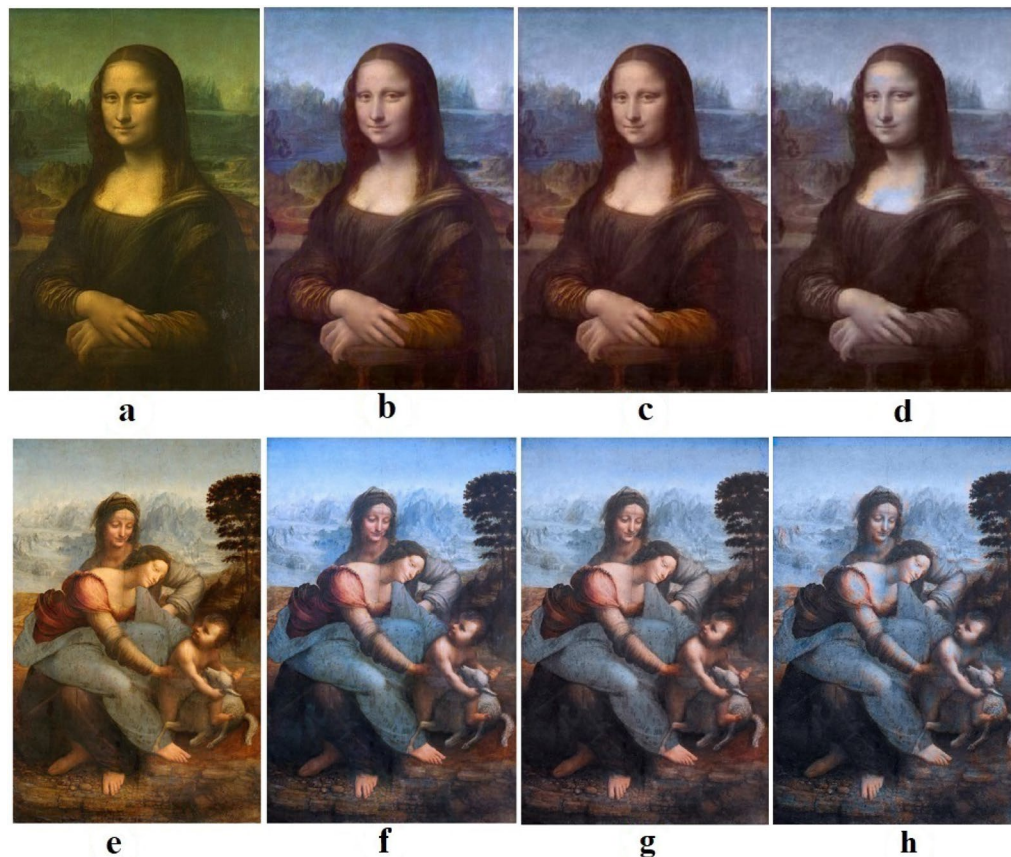


Fig. 10 **a** Uncleaned Mona Lisa, **b** virtually cleaned by [15], **c** virtually cleaned using the proposed CNN, **d** virtually cleaned using CNN proposed by [27] **e** uncleaned The Virgin and Child with Saint Anne, **f** physically cleaned, **g** virtually cleaned using the proposed CNN and **h** virtually cleaned using the CNN proposed by [27]

is the RGB image of the uncleaned work of art. One of the limitations of the network trained herein is that it might not work as well on abstract paintings, as it has been trained on the realistic looking images. The spatial and spectral information that the network has extracted from the realistic looking images might not match those of abstract ones. Also, working only on RGB images and being able to visualize the results of the virtual cleaning of artwork could be another limitation to this work. In the next research we aim to estimate the spectral reflectance data in the process of

virtual cleaning rather than only RGB visualizations. The spectral data could help the conservators with pigment mapping and identification. Another path to future research is to have access to data at different levels of cleaning. Here, we simply trained the model to transform the image to a fully cleaned final estimate of the work, but having access to data at different levels of cleaning would enable us to train a network for different levels and visualize the results. This could better aid the conservators with the choices they have to physically clean the artwork.

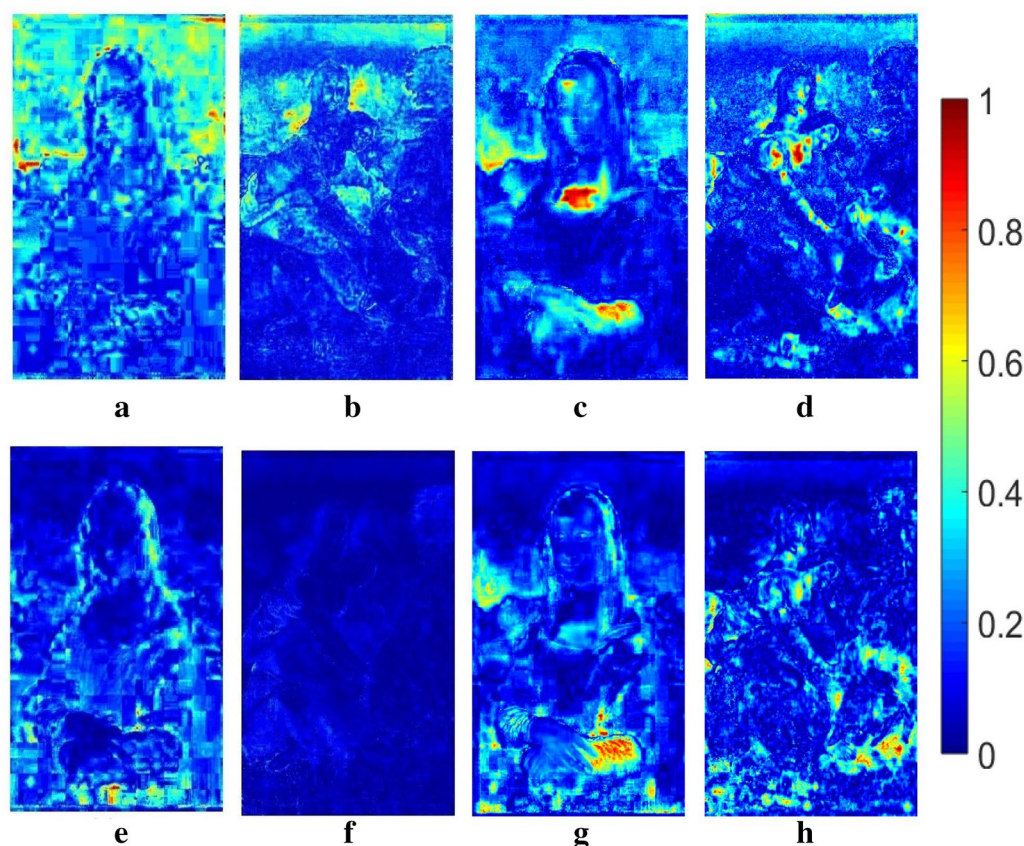


Fig. 11 **a** Euclidean distance computed between the virtually cleaned Mona Lisa using the proposed CNN and cleaned reference version, **b** Euclidean distance computed between the virtually cleaned The Virgin and Child with Saint Anne using the proposed CNN and physically cleaned version, **c** Euclidean distance computed between the virtually cleaned Mona Lisa using the CNN proposed by [27] and cleaned reference version, **d** Euclidean distance computed between the virtually cleaned The Virgin and Child with Saint Anne using the CNN proposed by [27] and physically cleaned version, **e** SA computed between the cleaned reference Mona Lisa using the proposed CNN and physically cleaned version, **f** SA computed between the virtually cleaned The Virgin and Child with Saint Anne using the proposed CNN and physically cleaned version, **g** Euclidean distance computed between the virtually cleaned Mona Lisa using the CNN proposed by [27] and cleaned reference version, **h** Euclidean distance computed between the virtually cleaned The Virgin and Child with Saint Anne using the CNN proposed by [27] and physically cleaned version. Red color shows the largest differences and blue color shows the smallest difference between the cleaned reference and the output images of the proposed method

Abbreviations

CNN:: Convolutional neural network; RGB:: Red, Green and Blue color space; CIE:: Commission Internationale de l'Éclairage; CIEXYZ:: X, Y and Z color space approved by CIE; sRGB:: standard RGB color space; JPEG:: Joint Photographic Experts Group, digital image compression method; SA:: Spectral angle.

Acknowledgements

This research was funded by Xerox Chair in Imaging Science at Rochester Institute of Technology.

Author Contributions

MMA developed the principal aspects of the algorithm, implemented it, and trained and tested the CNN and applied it to the data. DWM oversaw and advised the research. Both authors wrote the article and approved of the final version of the manuscript.

Funding

The work was done as a part of the author's Ph.D. research and was supported by Xerox chair at Rochester Institute of Technology.

Availability of the Data and Materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 10 February 2021 Accepted: 25 July 2021

Published online: 03 August 2021

References

1. Constantin S. The Barbizon painters: a guide to their suppliers. *Stud Conserv.* 2001;46:49–67.
2. Callen A. The unvarnished truth: mattness, primitivism and modernity in French painting, c. 1870–1907. *Burlingt Mag.* 1994;136:738–46.

3. Bruce-Gardner R, Hedley G, Villers C. Impressionist and post-impressionist masterpieces: the Courtauld collection. New Haven, Conn: Yale University Press; 1987.
4. Watson M, Burnstock A. An evaluation of color change in nineteenth-century grounds on canvas upon varnishing and varnish removal. In: New insights into the cleaning of paintings: proceedings from the cleaning 2010 international conference, Universidad Politecnica de Valencia and Museum Conservation Institute. Smithsonian Institution; 2013.
5. Berns RS, De la Rie ER. The effect of the refractive index of a varnish on the appearance of oil paintings. *Stud Conserv*. 2003;48:251–62.
6. Baglioni P, Dei L, Carretti E, Giorgi R. Gels for the conservation of cultural heritage. *Langmuir*. 2009;25:8373–4.
7. Baij L, Hermans J, Ormsby B, Noble P, Iedema P, Keune K. A review of solvent action on oil paint. *Herit Sci*. 2020;8:43.
8. Prati S, Volpi F, Fontana R, Galletti P, Giorgini L, Mazzeo R, et al. Sustainability in art conservation: a novel bio-based organogel for the cleaning of water sensitive works of art. *Pure Appl Chem*. 2018;90:239–51.
9. Al-Emam E, Soenen H, Caen J, Janssens K. Characterization of polyvinyl alcohol-borax-agarose (PVA-B/AG) double network hydrogel utilized for the cleaning of works of art. *Herit Sci*. 2020;8:106.
10. El-Gohary M. Experimental tests used for treatment of red weathering crusts in disintegrated granite-Egypt. *J Cult Herit*. 2009;10:471–9.
11. Gulotta D, Saviello D, Gherardi F, Toniolo L, Anzani M, Rabbolini A, et al. Setup of a sustainable indoor cleaning methodology for the sculpted stone surfaces of the Duomo of Milan. *Herit Sci*. 2014;2:6.
12. Trumpy G, Conover D, Simonot L, Thoury M, Picollo M, Delaney JK. Experimental study on merits of virtual cleaning of paintings with aged varnish. *Opt Express*. 2015;23:33836–48.
13. Barni M, Bartolini F, Cappellini V. Image processing for virtual restoration of artworks. *IEEE Multimed*. 2000;7:34–7.
14. Pappas M, Pitas I. Digital color restoration of old paintings. *IEEE Trans Image process*. 2000;9:291–4.
15. Elias M, Cotte P. Multispectral camera and radiative transfer equation used to depict Leonardo's sfumato in Mona Lisa. *Appl Opt*. 2008;47:2146–54.
16. Palomero CMT, Soriano MN. Digital cleaning and dirt layer visualization of an oil painting. *Opt Express*. 2011;19:21011–7.
17. Yang L, Kruse B. Revised Kubelka–Munk theory. I. Theory and application. *JOSA A*. 2004;21:1933–41.
18. Kirchner E, van der Lans I, Ligterink F, Hendriks E, Delaney J. Digitally reconstructing van Gogh's field with irises near Arles. Part 1: varnish. *Color Res Appl*. 2018;43:150–7.
19. Zeng Y, van der Lubbe JC, Loog M. Multi-scale convolutional neural network for pixel-wise reconstruction of Van Gogh's drawings. *Mach Vis Appl*. 2019;30:1229–41.
20. Wan Z, Zhang B, Chen D, Zhang P, Chen D, Liao J, et al. Bringing old photos back to life. In: proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2020. p. 2747–2757.
21. Linhares J, Carreira L, Bailão A, Pastilha R, Nascimento S. Chromatic changes in paintings of Adriano de Sousa Lopes after the removal of aged varnish. *Conserv Patrim*. 2020;34:50–64.
22. Dang LM, Hassan SI, Im S, Moon H. Face image manipulation detection based on a convolutional neural network. *Expert Syst Appl*. 2019;129:156–68.
23. Ratnasingam S. Deep camera: a fully convolutional neural network for image signal processing. In: proceedings of the IEEE international conference on computer vision workshops; 2019. p. 3868–3878.
24. Zhang Y, Liu Y, Sun P, Yan H, Zhao X, Zhang L. IFCNN: a general image fusion framework based on convolutional neural network. *Inf Fusion*. 2020;54:99–118.
25. Yadav SS, Jadhav SM. Deep convolutional neural network based medical image classification for disease diagnosis. *J Big Data*. 2019;6:1–18.
26. Shankar RS, Mahesh G, Murthy K, Ravibabu D. A Novel approach for Gray Scale Image Colorization using Convolutional Neural Networks. In: 2020 international conference on system, computation, automation and networking (ICSCAN). IEEE; 2020. p. 1–8.
27. Zhang R, Isola P, Efros AA. Colorful image colorization. In: European conference on computer vision. Springer; 2016. p. 649–666.
28. Dhir R, Ashok M, Gite S, et al. An overview of advances in image colorization using computer vision and deep learning techniques. *Rev Comput Eng Res*. 2020;7:86–95.
29. Anwar S, Tahir M, Li C, Mian A, Khan FS, Muzaffar AW. Image colorization: a survey and dataset. *arXiv preprint arXiv:2008.10774*. 2020;.
30. Kaggle. Datasets of urban and rural photos; 2018. <https://www.kaggle.com/dansbecker/urban-and-rural-photos>. Accessed 20 Jan 2021
31. Parkkinen JP, Hallikainen J, Jaaskelainen T. Characteristic spectra of Munsell colors. *JOSA A*. 1989;6:318–22.
32. Berns RS. Billmeyer and Saltzman's principles of color technology. New Jersey: Wiley; 2019.
33. Okumura Y. Developing a spectral and colorimetric database of artist paint materials, MS degree thesis. Center for Imaging Science, Rochester Institute of Technology, NY. 2005;.
34. Wikipedia. Category: the Virgin and child with Saint Anne (painting by Vinci)—Wikipedia, The Free Encyclopedia; 2021. [https://commons.wikimedia.org/wiki/Category:The_Virgin_and_Child_with_Saint_Anne_\(painting_by_Vinci\)](https://commons.wikimedia.org/wiki/Category:The_Virgin_and_Child_with_Saint_Anne_(painting_by_Vinci)). Accessed 07 Jun 2021
35. Gulli A, Pal S. Deep learning with Keras. Birmingham: Packt Publishing Ltd; 2017.
36. Wu H, Sun X, Yang J, Zeng W, Wu F. Lossless compression of JPEG coded photo collections. *IEEE Trans Image Process*. 2016;25:2684–96.
37. Farmer JO. See also optoelectronics; satellites, communication. *Encyclopedia of 20th-Century Technology*. 2005; p. 823.
38. Park B, Windham W, Lawrence K, Smith D. Contaminant classification of poultry hyperspectral imagery using a spectral angle mapper algorithm. *Biosyst Eng*. 2007;96:323–33.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)