

RESEARCH

Open Access



Virtual cleaning of works of art using a deep generative network: spectral reflectance estimation

Morteza Maali Amiri* and David W. Messinger

Abstract

Generally applied to a painting for protection purposes, a varnish layer becomes yellow over time, making the painting undergo an appearance change. Upon this change, the conservators start a process that entails removing the old layer of varnish and applying a new one. As widely discussed in the literature, helping the conservators through supplying them with the probable outcome of the varnish removal can be of great value to them, aiding in the decision making process regarding varnish removal. This help can be realized through virtual cleaning, which in simple terms, refers to simulation of the cleaning process outcome. There have been different approaches devised to tackle the problem of virtual cleaning, each of which tries to develop a method that virtually cleans the artwork in a more accurate manner. Although successful in some senses, the majority of them do not possess a high level of accuracy. Prior approaches suffer from a range of shortcomings such as a reliance on identifying locations of specific colors on the painting, the need to access a large set of training data, or their lack of applicability to a wide range of paintings. In this work, we develop a Deep Generative Network to virtually clean the artwork. Using this method, only a small area of the painting needs to be physically cleaned prior to virtual cleaning. Using the cleaned and uncleaned versions of this small area, the entire unvarnished painting can be estimated. It should be noted that this estimation is performed in the spectral reflectance domain and herein it is applied to hyperspectral imagery of the work. The model is first applied to a Macbeth ColorChecker target (as a proof of concept) and then to real data of a small impressionist panel by Georges Seurat (known as 'Haymakers at Montfermeil' or just 'Haymakers'). The Macbeth ColorChecker is simulated in both varnished and unvarnished forms, but in the case of the 'Haymakers', we have real hyperspectral imagery belonging to both states. The results of applying the Deep Generative Network show that the proposed method has done a better job virtually cleaning the artwork compared to a physics-based method in the literature. The results are presented through visualization in the sRGB color space and also by computing Euclidean distance and spectral angle (calculated in the spectral reflectance domain) between the virtually cleaned artwork and the physically cleaned one. The ultimate goal of our virtual cleaning algorithm is to enable pigment mapping and identification after virtual cleaning of the artwork in a more accurate manner, even before the process of physical cleaning.

Keywords Artwork virtual cleaning, Deep generative network, Varnished and unvarnished artworks, Spectral reflectance

Introduction

Even though there have been artists who did not intend for their works to be varnished, those works were sometimes varnished once out of the artists' hands. The varnish appearance on the surface of the artwork changes with time, causing the artwork visual properties to

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

change as well [1–3]. This change of appearance is more substantial after a significant passage of time [4]. Many factors play into the appearance alteration of the painting subsequent to varnish application. Two of the most important factors are the varnish type and its age [4, 5]. Due to the irreversibility of artwork cleaning, it is regarded as one of the tasks of highest importance undertaken by the conservators. This cleaning process constitutes physically removing the unwanted deposits and the aged varnish from the surface of the artwork aiding in reestablishing the original look of the artwork [6–8]. This process of cleaning is referred to as physical cleaning [9]. Physical cleaning can sometimes have damaging effects on the artwork along with being very time-consuming [10–12]. The simulation of the result of the varnish removal from an artwork is termed virtual cleaning. Virtual cleaning potentially provides the conservators with a likely appearance change of the painting if the cleaning process is undertaken. In some cases, the painting is not likely to undergo a thorough cleaning process anytime soon. In most cases, in fact, a small part of a painting is first cleaned. Using that cleaned portion, by estimating a relationship between the cleaned and uncleaned data of the region, the virtually cleaned version of the whole painting is estimated [13–16]. Therefore, virtual cleaning provides a method to estimate the original appearance of the painting [13]. Performing virtual cleaning in the spectral reflectance domain could also enable conservators to perform pigment mapping and identification more accurately even before a thorough physical cleaning.

Barni et al. [14] might be the original authors to report their work in virtual cleaning, albeit for *RGB* data. In their work, they first physically cleaned a part of the painting. They then found a transformation matrix between the cleaned and corresponding uncleaned portion in the *RGB* domain. Afterwards, they applied the same transformation matrix to the rest of the painting leading to a virtually cleaned artwork. Papas and Pitas stated that virtual cleaning in the CIELAB color space leads to a better result as opposed to the *RGB* color space [15]. They associated this observation with the close correlation between CIELAB and human perception. Having access to varnish and pigments that Leonardo da Vinci utilized at the time, Elias and Cotte (2008) were able to virtually clean the *Mona Lisa* [17]. They first built a chart of colors made from classical paints employed in 16th century Italy in a varnished and unvarnished state. Making use of these charts enabled them to deduce a mean multiplicative factor for each wavelength, which was then applied to the *Mona Lisa's* image spectra resulting in a virtually cleaned version of the painting [17]. Palomero and Soriano were the first to apply a neural network to the field of virtual cleaning [18]. They first physically cleaned a part of the

painting and trained a shallow neural network to go from the uncleaned painting to the cleaned one using the small physically cleaned part in the *RGB* domain. The network was then used to predict the *RGB* image of the cleaned painting. Using an estimation method, they were also able to estimate the spectral reflectance of the cleaned and uncleaned painting from *RGB* color data [18]. The estimation method was based on the Pseudo-Inverse (PI). This method is referred to as the PI because a multiplication of the pseudo-inverse matrix of *RGB* data (or any color data) and the reflectance data of the training samples is performed first in an attempt to recover the spectral reflectance of the testing samples from the *RGB* data. One point worth mentioning is that PI relies heavily on the training samples, similar to other supervised approaches. The PI works in a way that two sets of *RGB* and spectral training data are used to find the relationship between the *RGB* and spectral data. The same relationship is then applied to the testing data to estimate the spectral information [18, 19]. However, the important point is that the material types should be both the same in the training and testing data, and if the material type in the testing samples is not the same as the material in the training data, the spectral estimation of the testing data will not be accurate [19]. This is true even if the two spectra have the same *RGB* values, as visual colors might visually appear the same but the spectral reflectances could be different [20]. Let us assume, for example, the training data is a green plastic and its spectrum is also known. The PI is then used (or any supervised method for that matter) to extract the relationship between the spectral reflectance of the green plastic and its *RGB* color values. The same relationship is then applied to the testing data which is comprised of a green leaf, for which we only have the *RGB* values. The green leaf spectral reflectance estimated using this method will not be an accurate representation of its original spectral reflectance due to the spectral difference outside the visual portion of the spectrum [19]. Going back to the paper reported by [18], they estimated the spectral reflectance of the artwork using Munsell color chips, neglecting that Munsell chips are not constructed from the same material (pigments) as the ones used in the artwork, hence making the estimation of the spectra of artwork less accurate. Trumphy et al. were the first to approach the problem of virtual cleaning from a physics standpoint [13]. Using Keubelka-Munk theory as the base, they were able to develop a model that estimates the spectral reflectance of the cleaned painting. In order to do that, they made some simplifying assumptions about how the light interacted between the varnish and the painting. Some of those simplifying assumptions are that pigment particles are immersed in the binding medium, that there is varnish wetting of the

painting layer, the varnish surfaces are optically smooth while the exposed paint layer is rough, neglecting the reflection at the air/varnish interface as well as the paint/varnish interface, that a dark location completely absorbs the incident radiation, and finally, that the varnish body reflectance (which they assumed to be equal to the reflectance spectra of the uncleaned dark location of the painting) is wavelength independent [13]. This model is used below as a reference for comparison to our model. Kirchner et al. characterized the varnish layer through making a number of key measurements, particularly of the pure white on the painting using Kubelka-Munk two-constant theory [16]. Wan et al. using variational autoencoders and assuming that the image restoration can be treated as an image translation problem (in which the images are translated through three domains: the real image, a synthetic, artificially degraded image, and the ground truth with no degradation) were able to restore old images [21]. Linhares et al., using hyperspectral imaging technology, were able to characterize the varnish layer allowing them to virtually clean the artwork [22]. To do that, they measured the reflectance spectra of the painting before and after varnish removal and used that information for the subsequent characterization and virtual removal of the varnish layer [22]. The latest work in the area of virtual cleaning might belong to Maali Amiri and Messinger, where they trained a deep convolutional neural network in the *RGB* domain on an image database containing natural scenery and humans that were artificially yellowed [9]. They reported the ability to virtually clean artwork even though their network was trained on natural scenes [9]. Many of the previously reported works have shortcomings that are addressed in this work. Among them are the need to specify locations of pure black and white pigments, the need to have access to a large set of training data, low accuracy, and the inability to generalize the method and results to other works. The approach presented below overcomes these deficiencies while still providing a successful virtual cleaning of a painting.

In this paper, we develop a Deep Generative Network (DGN) to virtually clean artwork in the spectral reflectance domain. Similar to other approaches, this method requires a small part of the painting to be physically cleaned beforehand. Using that portion of the work, the DGN can virtually clean the entire painting. As a first test using simulated data, a Macbeth ColorChecker was synthetically yellowed using the method proposed in [9] in the spectral domain. It was then assumed that only a small part of it has been physically cleaned. Using the DGN, we successfully clean the rest of the color chart. The model is also applied to real hyperspectral imagery of a partially cleaned work referred to as 'Haymakers'; the

same painting used in the study by [13]¹. It is worth noting that in the case of the 'Haymakers', there is no need to simulate the varnished version of the painting as we have real data belonging to both varnished and unvarnished states of the painting. The results are shown in terms of Euclidean distance and spectral angle between the virtually cleaned and the physically cleaned artwork (computed in the spectral reflectance domain) along with visualizations of the results in the sRGB color space as well as the spectral domain. The results are compared with the physics-based model proposed by [13]. The comparison shows that the model proposed herein has outperformed the physics-based model.

The paper is laid out as follows: Sect. "Methodology" describes the data sets utilized in this work, the proposed deep generative network, and the experiments performed. Section "Results and discussion" demonstrates the results and discussions where our model results are compared with the physics-based model. We finish our paper with conclusions summarizing the paper's outcomes and contributions.

Methodology

This section describes the data used and the network algorithm along with its architecture. The criteria used to evaluate the success of the method proposed here is also presented.

Data

Hyperspectral imagery of the Macbeth ColorChecker and 'Haymakers at Montfermeil' are used in this work to test our model. The ColorChecker dataset consists of 24 color patches, and is a very suitable sample as it comprises a set of colors represented frequently works of art. Spectral reflectance data of this color chart from 400 to 700 nm at 5 nm intervals are available [23] and are used to simulate hyperspectral imagery, both "varnished" and "cleaned", of a ColorChecker target. The hyperspectral imagery of the 'Haymakers' is the same as used in [13], and was provided to us by the National Gallery of Art. There are two hyperspectral images of this work used in this study. One image is of the work after approximately 1/3 of the work was physically cleaned, while the other was collected after the full cleaning was completed. For our study, we use the small, pre-cleaned area to virtually clean the remaining 2/3 of the work, and then compare the result to the post-cleaned imagery. The hyperspectral image cubes of the 'Haymakers' contain reflectance spectra from 400 to 780 nm with a spectral sampling of 2.5

¹ We wish to thank the National Gallery of Art for providing us with the data of the 'Haymakers'.



Fig. 1 Images of the 'Haymakers at Montfermeil' **a** before removal of varnish and **b** after removal of varnish. It should be noted that this data is from a real artwork and in no way, we have done any simulation as we have with Macbeth ColorChecker

nm. A visualization (performed in the sRGB domain) of the data is shown in Fig. 1 for the varnished and unvarnished states (the pre-cleaned area of the painting is visible on the right-hand side).

Importantly for our results, we use the simulated Macbeth ColorChecker, which was artificially yellowed using the same method proposed in [9], to test the method on synthetic data, and then apply the approach to real hyperspectral imagery of a painting, both before and after physical cleaning was performed. The result of the virtual cleaning applied to the 'Haymakers' is compared with that of physics-based model. In the case of Macbeth ColorChecker, the results are not compared with the physics-based model, and it is only used as a way to test the feasibility of our model.

Deep generative network (DGN)

Neural Networks are generally able to learn non-linear transfer functions. Here, this approach learns the relationship between the spectra of those cleaned parts of the work, and the corresponding uncleaned parts. We then generalize the same relationship to other uncleaned areas. This precludes the need to use a physics-based model in which having access to samples of pure black and white on the painting is of great importance. Using the proposed method, we do not need to be concerned about the type of colors needed; literally any colors can be used for this purpose as long as they are representative of the entire work. It should also be noted that in this work we are assuming that the varnish effect is spatially uniform, an assumption that all virtual cleaning approaches make. Therefore, for only a small area of the painting we have both the cleaned and uncleaned spectral reflectance, and we use these spectra to learn the relationship between them. After discovering that relationship, we apply it to other parts of the painting that are still varnished and consequently we are able to virtually

clean the artwork. For this we use the method called deep generative network [24].

The idea of a generative network is to learn the relationship $x = f_{\theta}(z)$ mapping an image z to another image x . This approach will be applied to reconstruct the virtually cleaned hyperspectral image from the hyperspectral image of the uncleaned artwork. Our goal is to generate image $X \in \mathbb{R}^{B \times W \times H}$ (where, B is the number of spectral bands, W is the width and H is the height of the image, both in pixels) which is a virtually cleaned image of the varnished artwork. Through feeding the image cube $Z \in \mathbb{R}^{B \times W \times H}$ into the generator, an image with this characteristic will be attained. Here, Z is the hyperspectral image of the artwork before cleaning. As mentioned above, only a small area of the painting is pre-cleaned and we have the spectral images of that area for both cleaned and uncleaned conditions. Let us call the area of the painting for which we have both the cleaned and uncleaned spectra A . The spectral image of this area that is physically cleaned is called A_c and the corresponding spectral image of this area before cleaning is A_u . When Z goes through the network, the portion corresponding to A_u is taken out and the pixel-wise error between A_u and A_c is calculated to compute the loss. This is then back-propagated to the generator, through which the parameter θ of the mapping function is optimized. Fig. 2 shows the process described here.

The generator consists of different layers that are described below:

- 1) Convolution layer: this layer is comprised of a block of neurons involving the multiplication of a set of weights and biases by the input. The convolution layer will extract a particular feature of the input image. Given a convolution layer $C^{(i)}$ and biases $B^{(i)}$ and the field of view (FoV) of the feature map of the previous convolution layer $O^{(i-1)}$, $O^{(i)}$ is written as

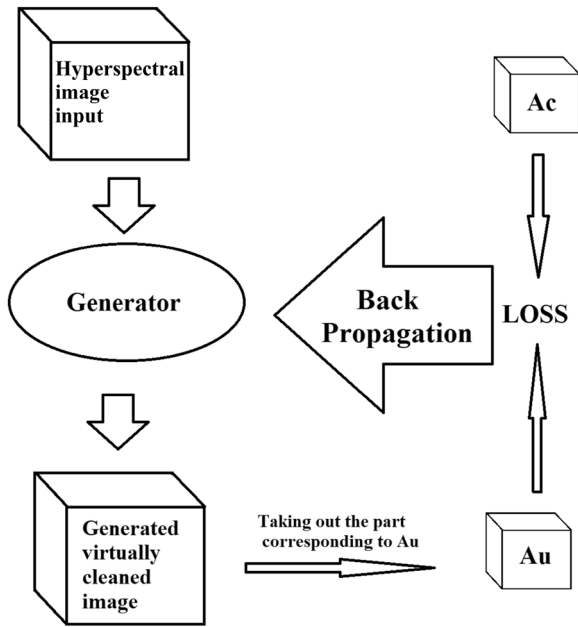


Fig. 2 Overview of the algorithm used by generative network

$$O^{(i)} = (O^{(i-1)} \cdot W^{(i)})_{f,l} + B^{(i)} \\ = \sum_{m=1}^{k^{(i)}} \sum_{n=1}^{k^{(i)}} (o_{f-m,l-n}^{(i-1)} \cdot \omega_{(m,n)}^{(i)}) + B^{(i)} \quad (1)$$

where $k^{(i)}$ is the size of the kernel, $O_{(f,l)}^{(i-1)}$ is the feature (f, l) of the feature map $O^{(i)}$ with $f = 1, 2, \dots, W$ and $l = 1, 2, \dots, H$ and $\omega_{(m,n)}^{(i)}$ is the $(m, n)^{th}$ element of the weight matrix $W^{(i)}$.

- 2) Batch normalization layer: this layer is a method for standardizing the inputs to the next layer, which has the impact of stabilizing the process of learning, and it is usually placed behind the convolution layer. The normalization is defined as

$$y = \frac{x - \text{mean}(x)}{\sqrt{\text{Var}(x) + \epsilon}} \cdot \gamma + \beta \quad (2)$$

where γ and β are learnable parameters, and ϵ is a parameter used for numerical stability. $\text{mean}(x)$ and $\text{Var}(x)$ are the mean and variance of x , respectively.

- 3) Activation layer: this layer is a nonlinear function that is attached to each neuron. It is a component of great importance as it specifies the computational efficiency of training a model and the convergence speed of the neural network. LeakyReLU is used here, defined as

$$f(x) = \begin{cases} x, & \text{if } x > 0 \\ \alpha x, & \text{if } x \leq 0 \end{cases} \quad (3)$$

where, α is a small nonzero parameter. We did not use ReLU as one of the problems of using ReLU is that its derivative is zero for negative part values which blocks the learning. However, in the case of the LeakyReLU, the derivative is a small fraction in the negative parts, which allows the gradients to flow on in the learning process.

The proposed generative network has an hourglass architecture, shown in Fig 3.

To be more specific, an image cube $Z \in R^{B \times W \times H}$, as input, goes through four main modules consisting of several blocks as follows:

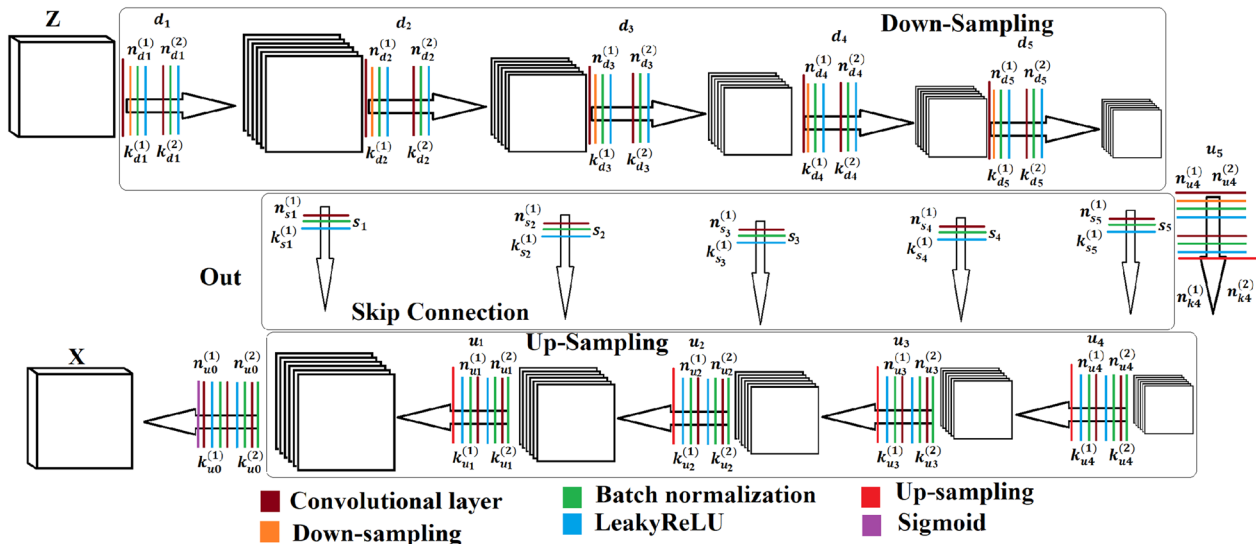


Fig. 3 The proposed generative network architecture

- 1) The down-sampling block: $d_{(i)}$ denotes the down-sampling blocks. Each $d_{(i)}$ is comprised of an initial convolution layer $C_d^{(1)}(i)$ that also performs the down-sampling step through setting the stride $S = 2$. It is then followed by the batch normalization and the LeakyReLU activation layer. The output is fed into the second convolution layer $C_d^{(2)}(i)$ with again $S = 2$. The same as the first activation layer, the second activation layer is followed by a batch normalization layer and the LeakyReLU activation function as well. $C_d^{(1)}(i)$ and $C_d^{(2)}(i)$ can be set to different kernel sizes and different numbers of filters shown as $k_d^{(1)}(i)$, $k_d^{(2)}(i)$, $n_d^{(1)}(i)$ and $n_d^{(2)}(i)$.
- 2) The up-sampling block: $u^{(i)}$ denotes the up-sampling blocks. Each $u^{(i)}$ consists of a few stacked layers. Opposite to the down-sampling blocks, the first layer here is batch normalization. It is then followed by the first convolution layer $C_u^{(1)}(i)$ with $S = 1$ and a batch normalization and LeakyReLU activation function. Its output is then fed into the second convolution layer $C_u^{(2)}(i)$. The output, after batch normalization and non-linear activation, is fed into the bilinear up-sampling layer with a factor of 2. $C_u^{(1)}(i)$ and $C_u^{(2)}(i)$, similar to the down-sampling block, can be set to different kernel sizes and different numbers of filters shown as $k_u^{(1)}(i)$, $k_u^{(2)}(i)$, $n_u^{(1)}(i)$ and $n_u^{(2)}(i)$, respectively.
- 3) Skip connection block: $s^{(i)}$ is utilized to denote the skip connection blocks. These blocks are used to connect the down-sampled data to the up-sampled data, so the residual information can be fully employed. It consists of one convolution layer, one batch normalization layer and one activation function. The number of filters and kernel size of convolution kernels in different layers can be set differently.

- 4) Output block: $o^{(0)}$ denotes the output block. It is an up-sampling block that is modified such that the up-sampling layer is replaced with one convolution layer, which is followed by one Sigmoid activation layer.

As mentioned, the input to the network is the hyper-spectral image of the uncleaned artwork $Z \in R^{B \times W \times H}$ and the generated image is $X \in R^{B \times W \times H}$. The cost function is defined as the pixel-wise difference between A_u and A_c ; defined above, A_c is the hyperspectral image of the area of the painting that is cleaned and A_u is the hyperspectral image of the same area but before cleaning. A_u belongs to X and therefore, it is changing in each iteration. Consequently, the cost function is given as

$$\min \|A_u - A_c\|^2. \quad (4)$$

To iterate to the best solution, the input to the model should be replaced with the output of the model after each iteration. As mentioned, the network has an hour-glass architecture as shown in Fig 3. Each down-sampling and up-sampling sections are comprised of 5 layers and 5 skip connections. The filter size is 3×3 in the up-sampling and down-sampling blocks but it is 1×1 in the last convolutional layer. There are 128 filters in each layer, both in the down-sampling and up-sampling blocks. There are 120 filters in the last convolutional layers equaling the spectral resolution of the hyperspectral images used. Overall there are 12 layers including input and output layers. The Adam optimization algorithm is used, chosen based on trial and error. The loss function, as mentioned before, is the Euclidean distance between the virtually cleaned area of the artwork and the physically cleaned one (A_u and A_c). The overall algorithm is shown in Algorithm 1

Algorithm 1 Deep Generative Network Algorithm

Procedure: Virtual Cleaning (A_c)
Input: Hyperspectral image of the uncleaned artwork Z
while epoch < max.epoch **do**
 $X = \text{Model}(Z)$ (Model here stands for the deep generative model.)
 $A_u = X$ (The part of the X corresponding to A_c is taken out)
 $\min \|A_u - A_c\|^2$
 $Z = X$ (replace the input with the output of the model in each iteration)
end while
Return X
End Procedure

It should be noted that this is an unsupervised approach. There is no training in the traditional sense here because it is a generative network. The error computed between A_u and A_c is back propagated to the generator which cleans the entire image using the error coming from the loss function. This cleaning process takes place step by step at each epoch, until the network reaches the maximum number of epochs.

Evaluation metrics

The virtually cleaned result is transformed into the sRGB format for visual inspection and to evaluate the success of the process. For a quantitative evaluation the per-pixel spectral Euclidean Distance and Spectral Angle (SA) are also calculated between the hyperspectral image of the physically cleaned work and the virtually cleaned hyperspectral image [25]. The spectral angle is calculated between two vectors in the spectral reflectance space and it is reported in radians in the range $[0, 3.142]$. The spectral angle is defined as

$$SA_k = \cos^{-1} \left(\frac{\mathbf{t}_k \cdot \mathbf{r}_k}{|\mathbf{t}_k| |\mathbf{r}_k|} \right) \quad (5)$$

where k represents the k th pixel, and \mathbf{t}_k and \mathbf{r}_k represent the two pixels belonging to the test and reference images. Also, the mean spectral reflectance of a few randomly chosen areas on the painting are compared between different approaches.

Experimental environment

Python 3.9.7 Anaconda, Inc. is used as a base coding environment for the DGN algorithm. More specifically, the DGN codes were written and run in the TensorFlow environment, which was installed onto the Anaconda. In terms of hardware, the programs are run on a CPU that is 11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80GHz. The training of the DGN is performed using only one image and is consequently referred to as an unsupervised learning method [24]. As mentioned before, only a small area of the image is used to compute the loss function, and the same loss is then used for the whole image to virtually clean it. 10000 epochs are used to train the model. MATLAB R2022a, the package of mathematical software was also used for evaluation computations and processing the samples used in our model.

Results and discussion

In this section, the results of applying the DGN to the problem of virtual cleaning are presented. This section is divided into two subsections in which the results for the Macbeth ColorChecker and the 'Haymakers' are presented separately.

Macbeth ColorChecker

A spectral representation of the Macbeth ColorChecker, as mentioned before, is used to test the DGN model that was developed in this work. As described above, a small part of the image of the Macbeth ColorChecker should be first physically cleaned. The rest of the painting will be virtually cleaned using hyperspectral imagery of that small part both before and after physical cleaning. A key question in this work is dependence of the performance on the choice of data that is pre-cleaned. In our simulation of the ColorChecker, we assumed that first, only white, red, green and blue patches are cleaned and the rest of it is uncleaned. In the second case, we assume that half of all the patches available on the Macbeth ColorChecker are cleaned. Through trial and error, we realized that at least four color patches are required to virtually clean the image successfully. Choosing half of all patches gives us a good idea about how the DGN works when it has data associated with all colors represented in the work. The results of these two cases are shown in Fig. 4.

As observed from Fig. 4, the DGN has been able to virtually clean the Macbeth ColorChecker and is successful in replicating the original cleaned Macbeth Chart in Fig. 4b. It is difficult to distinguish between the results in Fig. 4 as they are visually similar.

However, Table 1 presents the mean and standard deviation of the Euclidean distance and spectral angle between the virtually cleaned Macbeth ColorChecker and the original one. Note that the data are in spectral reflectance in the range $[0, 1]$. As it is observed from the table, the DGN has been able to clean the Macbeth ColorChecker at a very acceptable manner signaling that this method could be potentially applied to real artworks as well. Moreover, the DGN has resulted in a better outcome when half of all patches present on the Macbeth ColorChecker are utilized. This is not surprising. When we use half of all the patches to compute the loss function, we are using all the possible samples present in our dataset resulting in a lower error in the virtual cleaning process as shown herein.

The Euclidean distance and spectral angle measured between the virtually cleaned Macbeth ColorChecker and the original one are also presented in Fig. 5. This representation helps us see where the method is falling short. It should be noted that all of these data are normalized between 0 and 1 across all results. This helps see visually which methods, and which patches, are cleaned better than others.

It is clear that using half of all the patches leads to a better result as opposed to using only white, red, green and blue patches. Interestingly, the white patch on the color chart output by the DGN when half of all patches are used looks even better than that of DGN when white,

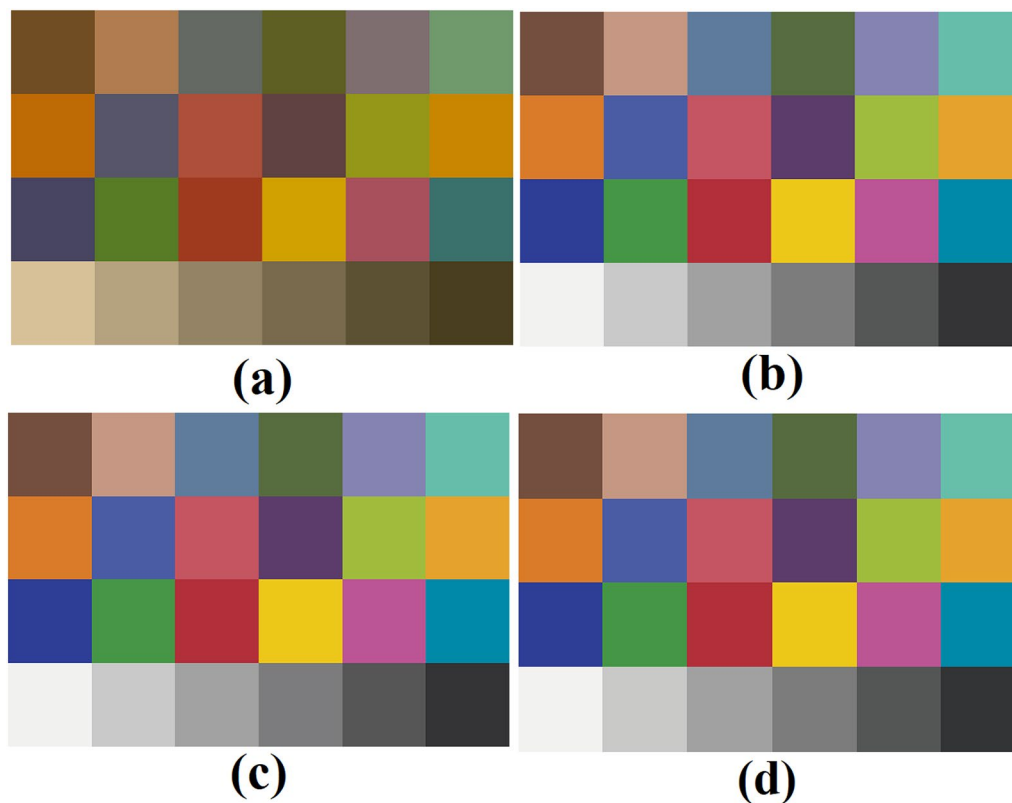


Fig. 4 **a** Unclean Macbeth ColorChecker, **b** clean Macbeth ColorChecker, **c** cleaned Macbeth ColorChecker using DGN (white, red, green and blue patches are used), **d** cleaned using DGN (half of all patches are used)

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

Method	Euclidean distance		SA	
	Mean	SD	Mean	SD
DGN (white, red, green and blue)	0.0278	0.0060	0.0134	0.0011
DGN (half of all patches)	0.0136	0.0024	0.0076	0.0013

redm green and blue patches are used. The black patch, on the other hand, has been cleaned in a more acceptable way. It should be noted that these results are normalized with respect to the maximum value present in the Euclidean distance and spectral angle computations, separately.

'Haymakers at Montfermeil'

In this section, the results of applying the DGN to the 'Haymakers' are presented and described.

As observed above from applying the DGN to the Macbeth ColorChecker, the result varies depending on which patches are chosen on the chart as data to compute the

loss function. We presented two different conditions: one with only 4 patches assumed to be physically cleaned, and the other assuming that half of all patches present are physically cleaned. For imagery of real artwork, these represent two use cases, one in which a small part of the painting is physically cleaned, and one in which an attempt is made to partially clean as many representative pigments as possible in the work. Using the data belonging to both the cleaned and uncleaned states of that small area, the DGN is able to estimate the virtually cleaned version of the whole work. To examine the same point (the impact of the area chosen to be physically cleaned), two different conditions are tested herein as well. In other words, two different experiments are performed differing only in cleaned small area the DGN uses as data for the computation of the loss function to estimate the cleaned version of the whole painting. Fig. 6 shows the two different areas chosen to perform these two experiments. The experiments are referred to as the "first" (Fig. 6a) and "second" (Fig. 6b) experiments. These two experiments differ in terms of A_c , or the small area that is physically cleaned using which the network tries to virtually clean the whole painting.

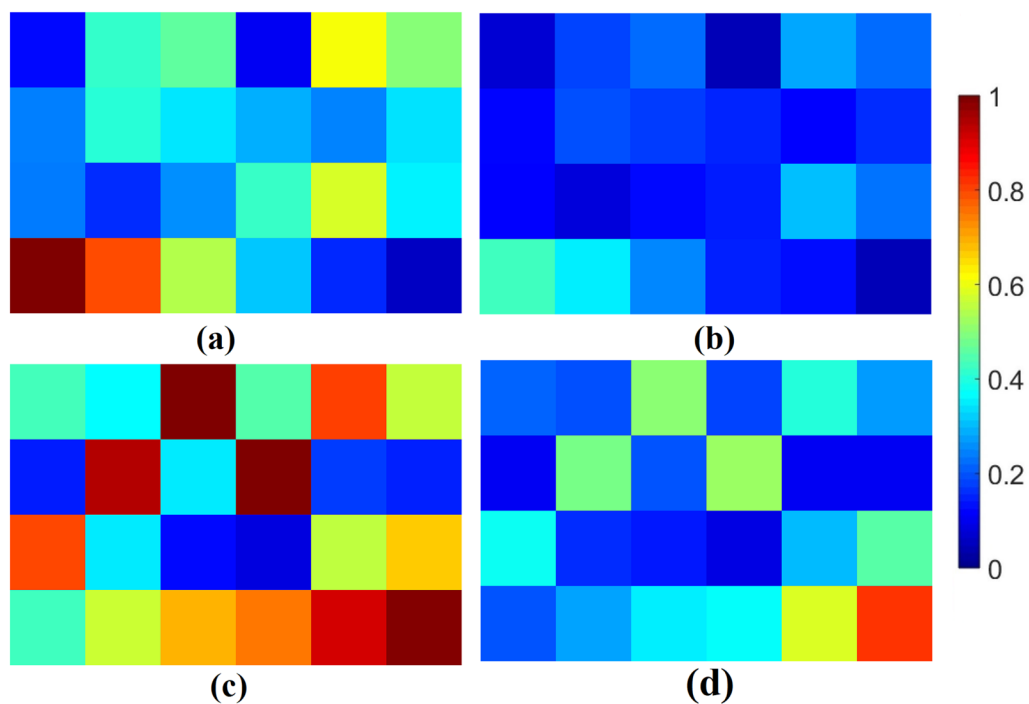


Fig. 5 Visualization of Euclidean distance computed between **a** Virtually cleaned Macbeth ColorChecker using DGN (white, red, green and blue patches) and the original one and **b** virtually cleaned Macbeth ColorChecker using DGN (half of all the patches) and the original one. Visualization of Spectral angle computed between **c** virtually cleaned Macbeth ColorChecker using DGN (white, red, green and blue patches) and the original one and **d** virtually cleaned Macbeth ColorChecker using DGN (half of all the patches) and the original one. The data is normalized between 0 and 1

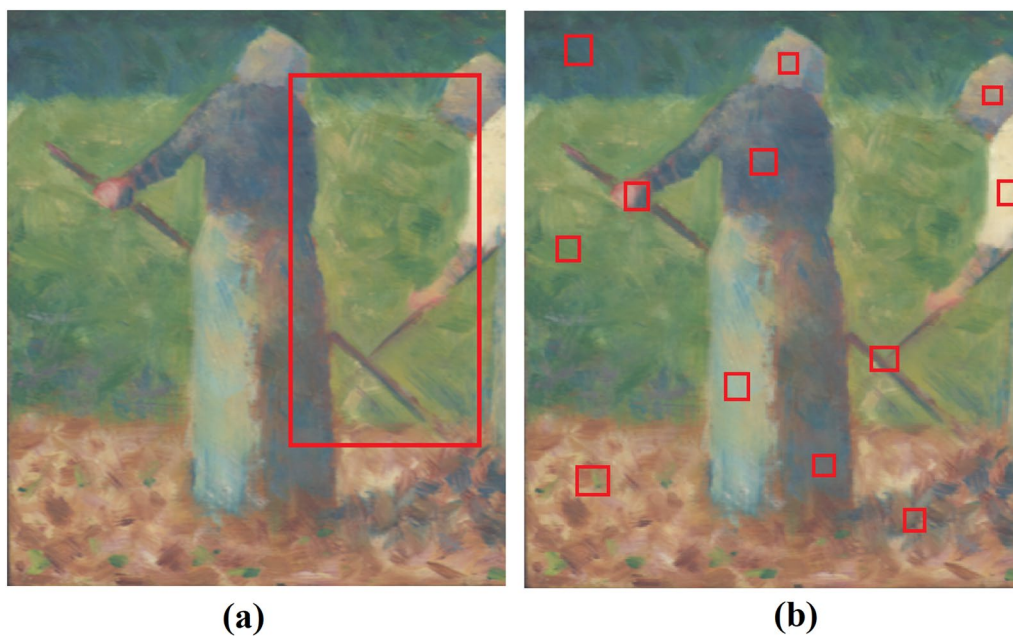


Fig. 6 Two different areas used in two experiments referred to as **a** first and **b** second experiments

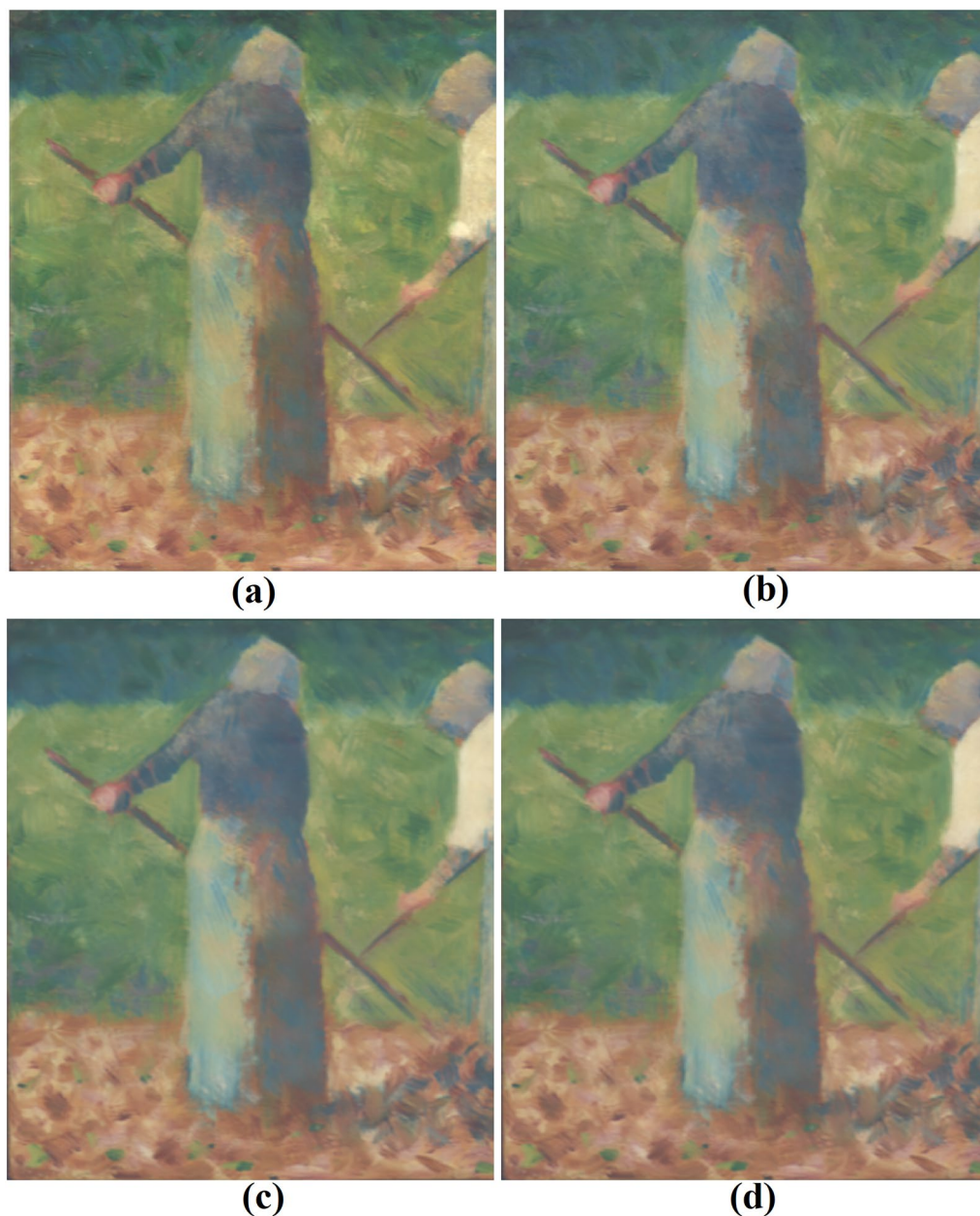


Fig. 7 **a** Virtually cleaned artwork using the physics-based model, **b** physically cleaned artwork, **c** virtually cleaned artwork in the second experiment and **d** virtually cleaned artwork in the first experiment

In the first experiment, a contiguous small area of the painting is chosen to be physically cleaned. In the second experiment, a few small areas spread over the painting are chosen. We performed the second experiment in an attempt to include as many different pigments as possible. To do that, we chose different areas spread over the painting from different pigments. We actually want to test the impact of the A_c on the final outcome of the virtual cleaning. In total there are 357,594 reflectance

spectral samples in the imagery of the painting. In the first experiment, 88,409 of those were used for training (almost 25 percent) and in the second experiment, 8860 were used for training (almost 2.5 percent).

Figure 7 shows the results for these two experiments along with the results of applying the physics-based approach. Visually, the physics-based model has not been able to clean the artwork as well as the DGN. This is subtle, but can be seen by looking closely at the grass

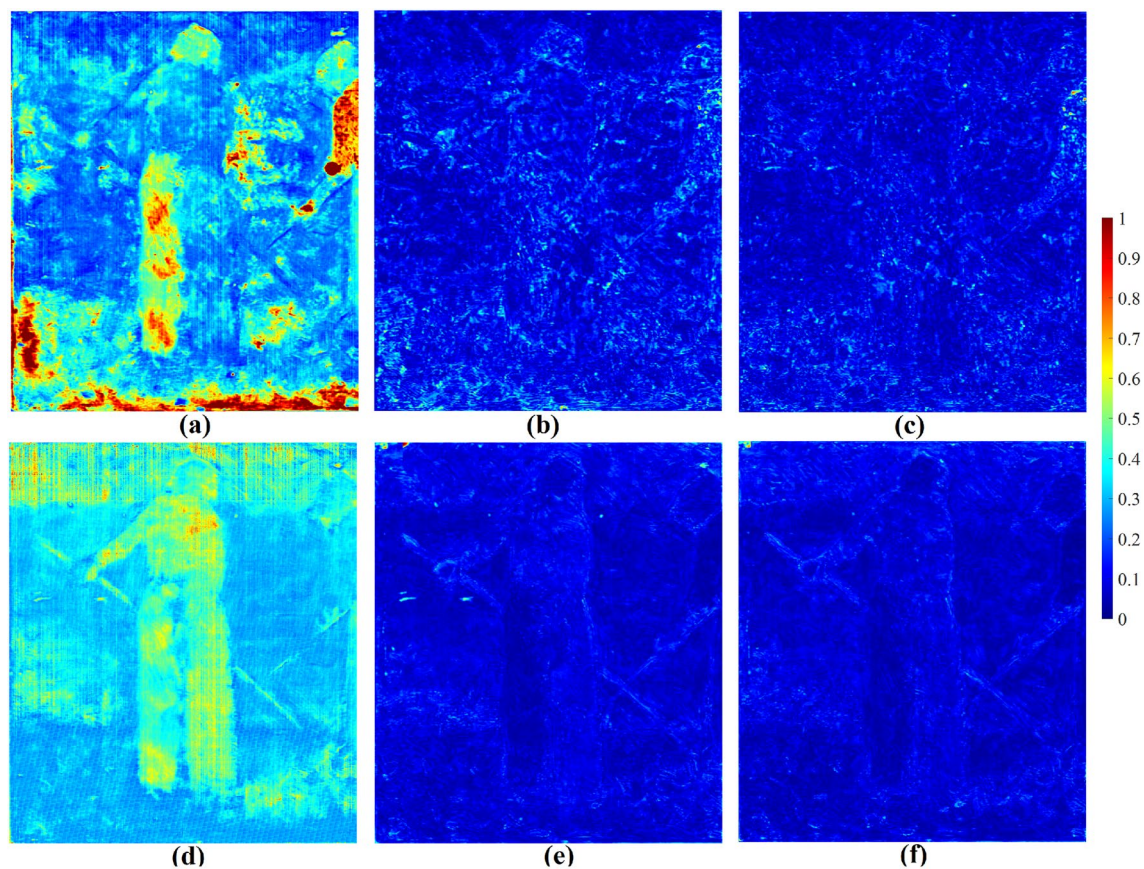


Fig. 8 ED computed between the physically cleaned artwork and the virtually cleaned one using **a** physics-based model, **b** DGN (the first experiment) and **c** DGN (the second experiment). SA computed between the physically cleaned artwork and the virtually cleaned one using **d** the physics-based model, **e** DGN (the first experiment) and **f** DGN (the second experiment). The data has been normalized between 0 and 1

in the middle of the images and noticing that there is a tint of yellow in the output of the physics-based model, not seen in the output of the DGN. One reason for this result is that there is no true black color present on the 'Haymakers' making the prediction made by the physics-based method, which relies heavily on the presence of pure black and white paints, not as accurate [9]. It is also hard to tell the difference between the outputs of the two experiments performed using the DGN.

Figure 8 shows the Euclidean distance (ED) and spectral angle (SA) measures between the virtually cleaned artwork and the physically cleaned one for the experiments performed here as well as the physics-based method. It is clear that the physics-based approach has not led to a good result. The error, in terms of ED and SA, is much lower for both the first and second experiments in the case of the DGN, than the physics-based model. Looking more closely at Fig. 8, one sees that the second experiment has led to a slightly better result than the first experiment. The reason is that in the second experiment, the area chosen to compute the loss

function by the DGN contains a better representation of the possible colors and paints in the work, helping the DGN learn the transfer function from the varnished version of the painting to the unvarnished one.

To further quantitatively examine the results shown in Fig. 8, the distributions of the spectral angle and Euclidean distance metrics are presented in Fig. 9. Overall, the distributions associated with DGN (both first and second experiment) have lower means and are narrower than those associated with the physics-based model.

To see the difference between the first and second experiment more clearly, Table 2 shows that the second experiment has led to a better result than the first experiment, in terms of mean and standard deviation of the metrics, for the reasons explained before. These experiments show that the more representative the area chosen to be physically cleaned and used by the DGN to compute the loss, the better the overall virtual cleaning outcome. It is still worth noting that the DGN has cleaned the painting at an acceptable level even in the first experiment.

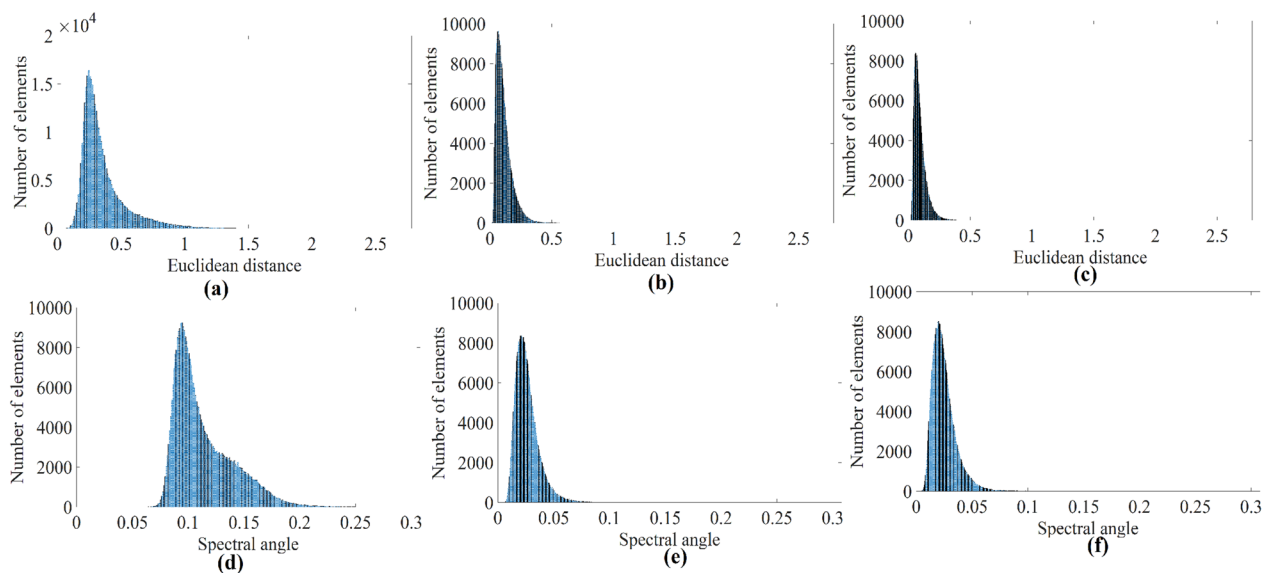


Fig. 9 **a** Distribution of Euclidean distance calculated between physically cleaned artwork and the virtually cleaned one using **(a)** physics-based model, **b** DGN (the first experiment) and **c** DGN (the second experiment). Distribution of spectral angle calculated between the physically cleaned artwork and the virtually cleaned one using **d** the physics-based model, **e** DGN (the first experiment) and **f** DGN (the second experiment)

Table 2 Euclidean distance and SA mean and standard deviation (SD) values between the physically and virtually cleaned ‘Haymakers’

Method	Euclidean distance		SA	
	Mean	SD	Mean	SD
Physics-based method [13]	0.3571	0.0547	0.1135	0.0057
DGN (first experiment)	0.1121	0.0120	0.0266	0.0092
DGN (second experiment)	0.0931	0.0082	0.0249	0.0022

We end this section by showing some of the spectral reflectance curves from the physically cleaned painting, and virtually cleaned ones using the physics-based approach, and the DGN (the first and second experiments). The curves are obtained through computing the average spectral reflectance factor of randomly chosen small areas on the painting as shown in Fig. 10.

While for some pigments all of the methods perform well, in general, we can see from Fig. 10 that the DGN (especially the second experiment) has led to a better result compared to the physics-based model. As mentioned above, in the first experiment, a rigid area has been chosen which might not be a good representative of the whole painting. However, in the second experiment, the DGN is exposed to many more pigments in the work through choosing different areas of the painting to compute the loss function. Overall, in both cases, the DGN has done an acceptable job of cleaning the painting.

Conclusions

In this work, we applied a deep generative network (DGN) to the problem of virtual cleaning of artworks. We used a simulated Macbeth ColorChecker and real hyperspectral imagery of the ‘Haymakers’ painting for this purpose. The results were compared with a well-known physics-based model both visually and in terms of Euclidean distance and spectral angle, computed in the spectral reflectance domain, between the virtually cleaned and physically cleaned artworks. Different areas of the artwork were chosen for the DGN to compute the loss function, used to estimate the cleaned version of the whole painting. The results showed that the DGN was able to outperform the physics-based model. It was also observed that the choice of the small, pre-cleaned area used by the DGN is important, and the more representative the small area is of the entire painting, the better the virtual cleaning outcome. This could be seen as one of the limitations of this method. Nonetheless, the method was able to lead to an acceptable result even when the small area was not as representative of the entire painting. Another limitation of this method might be its reliance on a pre-cleaned small area of the painting, which might not always be possible to obtain. However, if this is available, the method described here is useful to perform virtual cleaning on the entire painting.

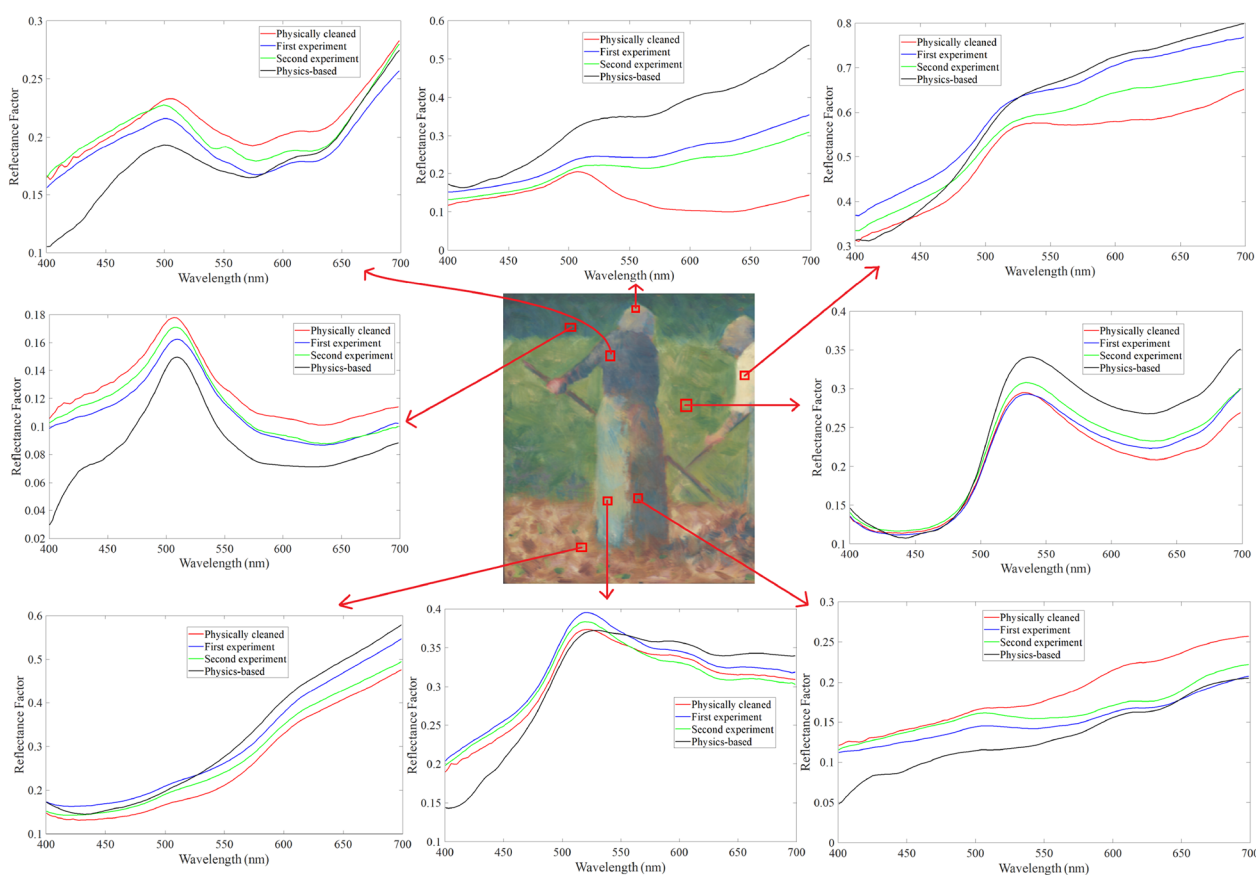


Fig. 10 Averaged spectral reflectance curves of the physically cleaned painting and virtually cleaned ones using physics-based approach and DGN computed over different areas randomly selected on the painting

Abbreviations

DGN Deep generative network
ED Euclidean distance
SA Spectral angle

Acknowledgements

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

We would also like to express our gratitude to John Delaney and Kathryn Dooley from the National Gallery of Art for providing us with access to the hyperspectral images of the 'Haymakers' collected both before and after physical cleaning.

Author contributions

MMA developed the principal aspects of the algorithm, implemented it, and trained/tested the DGN and applied it to the data. DWM oversaw and advised the research. Both authors wrote the article, read 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 the Xerox Chair in Imaging Science at the 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: 26 September 2022 Accepted: 5 January 2023

Published online: 24 January 2023

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. *Burling 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. Bajj 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. Maali Amiri M, Messinger DW. Virtual cleaning of works of art using deep convolutional neural networks. *Herit Sci*. 2021;9(1):1–19.
10. 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.
11. El-Gohary M. Experimental tests used for treatment of red weathering crusts in disintegrated granite-Egypt. *J Cult Herit*. 2009;10:471–9.
12. 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.
13. Trumpy G, Conover D, Simonot L, Thoury M, Piccolo M, Delaney JK. Experimental study on merits of virtual cleaning of paintings with aged varnish. *Opt Express*. 2015;23:33836–48.
14. Barni M, Bartolini F, Cappellini V. Image processing for virtual restoration of artworks. *IEEE Multimed*. 2000;7:34–7.
15. Pappas M, Pitas I. Digital color restoration of old paintings. *IEEE Trans Image Process*. 2000;9:291–4.
16. 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.
17. 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.
18. Palomero CMT, Soriano MN. Digital cleaning and "dirt" layer visualization of an oil painting. *Opt Express*. 2011;19:21011–7.
19. Maali Amiri M, Fairchild MD. A strategy toward spectral and colorimetric color reproduction using ordinary digital cameras. *Color Res Appl*. 2018;43(5):675–84.
20. Berns RS. Billmeyer and Saltzman's principles of color technology. New Jersey: Wiley; 2019.
21. 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.
22. 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 Patrimônio*. 2020;34:50–64.
23. Resources MCSLE. Spectral data for commonly used color products; 2018. <https://www.rit.edu/science/munsell-color-science-lab-educational-resources>. Accessed 20 Feb 2022.
24. Haut JM, Fernandez-Beltran R, Paoletti ME, Plaza J, Plaza A, Pla F. A new deep generative network for unsupervised remote sensing single-image super-resolution. *IEEE Trans Geosci Remote Sens*. 2018;56(11):6792–810.
25. 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)