

A Learning Based Approach to Separate Mixed X-Ray Images Associated with Artwork with Concealed Designs

Wei Pu Junjie Huang Barak Sober Nathan Daly Catherine Higgitt Pier Luigi Dragotti
University College London *Imperial College* *Duke University* *National Gallery* *National Gallery* *Imperial College*
London, UK London, UK Durham, US London, UK London, UK London, UK

Ingrid Daubechies
Duke University
Durham, US

Miguel R.D. Rodrigues
University College London
London, UK

Abstract—X-ray images are widely used in the study of paintings. When a painting has hidden sub-surface features (e.g., reuse of the canvas or revision of a composition by the artist), the resulting X-ray images can be hard to interpret as they include contributions from both the surface painting and the hidden design. In this paper we propose a self-supervised deep learning-based image separation approach that can be applied to the X-ray images from such paintings (‘mixed X-ray images’) to separate them into two hypothetical X-ray images, one containing information related to the visible painting only and the other containing the hidden features. The proposed approach involves two steps: (1) separation of the mixed X-ray image into two images, guided by the combined use of a reconstruction and an exclusion loss; (2) even allocation of the error map into the two individual, separated X-ray images, yielding separation results that have an appearance that is more familiar in relation to X-ray images. The proposed method was demonstrated on a real painting with hidden content, *Doña Isabel de Porcel* by Francisco de Goya, to show its effectiveness.

Index Terms—Art Investigation, Image Separation, Deep Neural Networks, Convolutional Neural Networks

I. INTRODUCTION

The cultural heritage sector is experiencing a digital revolution driven by the need to augment traditional invasive approaches to the technical study of works of art with non-invasive, non-destructive imaging and analytical techniques [1], [2]. With a number of new techniques being adopted for use in the cultural heritage sector, a wealth of multi-modal imaging data is being generated with a corresponding need for signal and image processing approaches to fully interrogate and visualise the resulting data. Cultural heritage institutions are also increasingly committing to making images and data relating to their collections publicly available, presenting new research opportunities.

While they have a long history of use, X-radiographs (X-ray images) still play a vital role in informing the technical study, conservation, and preservation of artworks in cultural heritage institutions due to the ability of X-rays to penetrate deep into a

painting’s stratigraphy [3], [4]. They can help to establish the condition of a painting (e.g., losses and damages not apparent at the surface), the status of different paint passages (e.g., to identify retouching, fills or other conservation interventions) or provide information about the painting support (e.g., type of canvas or the construction of a canvas or panel). X-ray images also provide insight into how the artist built up the different paint layers, thus revealing *pentimenti* – changes made by the artist during painting – which may include previous iterations of the final design or even earlier abandoned designs that were painted over when the artist revised the composition or if the painting support was reused by the artist for a completely different painting. There are many such artworks with hidden features and research from the Van Gogh Museum in Amsterdam has shown that 20 of 130 of Van Gogh’s paintings, i.e., nearly 15%, contained hidden paintings [5].

There has therefore been much interest in approaches capable of deriving clearer images of these hidden designs, in order to aid art historical scholarship and understanding of an artist and his/her work. Some researchers have proposed approaches leveraging various imaging modalities to enhance visualisation of concealed images in paintings [6], improve imaging of underdrawings [7] (e.g. preliminary sketches made on the picture support before painting), or help reveal overwritten *pentimenti* [8]. However, we are not aware of any prior work that attempts to extract images of the hidden features or paintings from the mixed X-ray images just described.

To address this challenge, in this paper, we therefore propose a new deep learning based approach to separate a mixed X-ray image associated with paintings with concealed designs into their hypothetical constituent images, corresponding to the X-rays of the visible painting and of the hidden image below. Our proposed approach – which is entirely self-supervised – leverages only the mixed X-ray image along with the RGB (visible) image acquired from the painting. Our proposed approach is also applied to Francisco de Goya’s portrait of

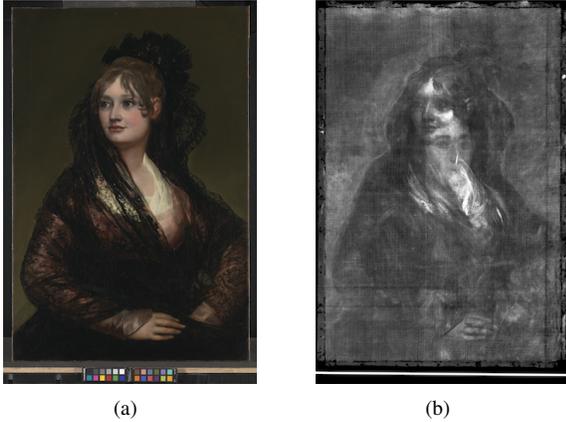


Fig. 1. Francisco de Goya, *Doña Isabel de Porcel* (NG1473), before 1805. Oil on canvas. (a). RGB image. (b). X-ray image. Images copyright of the The National Gallery, London.

Doña Isabel de Porcel, illustrated in Fig. 1, that has been painted directly on top of another portrait of a male figure [9]. An improved visualisation and a greater understanding of the underlying portrait has been possible through the use of macro X-ray fluorescence scanning, but it is also of great interest to attempt to obtain a clearer image of just the lower figure [15].

We note in passing that X-ray image separation approaches have already been proposed in a series of works such as [10-14]. However, such approaches apply to double sided panels where – in addition to the mixed X-ray image – one also has access to two RGB images associated with the front and back sides of the artwork. Our new approach applies however to a much more challenging scenario where one has access only to the mixed X-ray image (Fig. 1 b) plus a single RGB image associated with the painting visible at the surface (Fig. 1a) so that previous approaches – including ours in [10-14] – do not directly apply remove one full stop. The paper is organized as follows: in Section 2 the proposed two-step X-ray separation algorithm is presented; Section 3 shows the X-ray separation results of the proposed method on *Doña Isabel de Porcel* [9]; and Section 4 concludes the paper.

II. PROBLEM FORMULATION AND RELATED WORK

A. Problem Formulation

We wish to separate a mixed X-ray image (arising from a painting with a concealed design) containing features associated with both the visible and the concealed designs, by leveraging the availability of the visible image associated with the surface painting.

In particular, we assume that a patch of the mixed X-ray x corresponds to the linear superposition of a patch of the hypothetical X-ray associated with the visible painting x_1 and a patch of the hypothetical X-ray associated with the concealed features or painting x_2 as follows

$$x = x_1 + x_2. \quad (1)$$

We also assume there is a mapping \mathcal{F} that is approximately able to convert an image patch in the RGB domain into an image patch in the X-ray domain as follows:

$$x \approx \mathcal{F}(r_1) + x_2. \quad (2)$$

We can then cast the X-ray image separation problem as the task of learning the mapping function \mathcal{F} .

B. Mixed X-Ray Separation for Double Sided Paintings/Panels vs Mixed X-Ray Separation for Paintings with Concealed Designs.

We note that this problem involving the decomposition of mixed X-rays associated with paintings with concealed designs is much more challenging than another problem arising in art investigation involving the decomposition of mixed X-rays associated with double sided paintings/panels such as some panels of the well-known Ghent Altarpiece [16].

In particular, in challenges involving the separation of mixed X-rays associated with double sided paintings/panels – also considered by some of these authors [10-14] – one has access to patches of the mixed X-ray associated with a double-sided panel x along with patches of the visible images associated with the front panel r_1 and the rear panel r_2 . One can also posit that the mixed X-ray patch x corresponds to the sum of the hypothetical X-ray patch associated with the front panel x_1 and the hypothetical X-ray patch associated with the rear panel x_2 as follows

$$x = x_1 + x_2 \quad (3)$$

One can then posit there is a mapping function F converting a patch associated with the visible image r_i , $i = 1, 2$ to the patch of the corresponding hypothetical X-ray x_i , $i = 1, 2$ so that

$$x = F(r_1) + F(r_2) \quad (4)$$

Therefore, this alternative X-ray separation problem – which can also be tackled by learning the underlying mapping function $F(\cdot)$ as suggested in [13, 14] – is much simpler than the previous one because one has access to additional information to aid in the decomposition of the mixed X-ray onto its constituents (namely two hypothetical visible images associated with the paintings on the front and rear of the panel rather than a single visible image as in Fig. 1).

We show next how to tackle the herein proposed more challenging unmixing problem by appropriately re-purposing the learning approach original pioneered in [13, 14].

III. PROPOSED APPROACH

For the challenge of separating a mixed x-ray image arising from a painting containing a concealed design we therefore propose a two-step approach – inspired by [13, 14] – to separate each X-ray image patch x into its constituents x_1 and x_2 given the corresponding RGB image patch r_1 . The first step yields initial estimates of individual X-ray images x_1 and x_2 , and whereas the second step produces a refined version of these images.

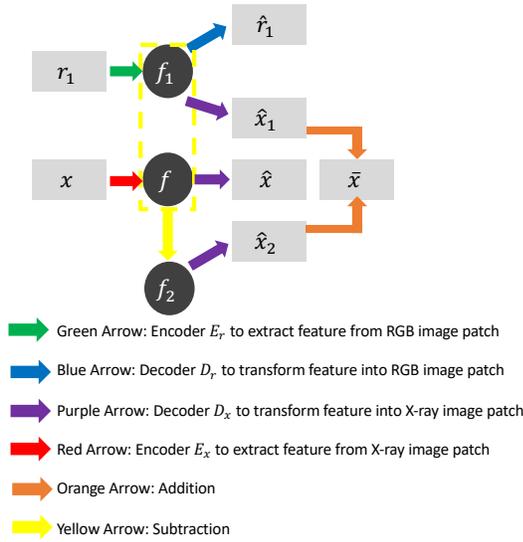


Fig. 2. The proposed connected auto-encoder network for initial separated X-ray image estimation.

A. Approach

1) *Step 1: Individual X-Ray Images Estimation:* In the first step, we propose to obtain initial estimates of the individual X-ray images of the surface and the hidden paintings using the connected auto-encoder structure shown in Fig. 2. We build the connected auto-encoder perform various operations:

- It extracts feature f_1 from the RGB image patch r_1 using an encoder E_r (represented by the green arrow).
- It extracts features f from the mixed X-ray image patch x by encoder E_x (represented by the red arrow).
- It derives latent features f_2 corresponding to the hidden painting by computing $f_2 = f - f_1$ (represented by the yellow arrows).¹
- It reconstructs X-ray image patches \hat{x}_1 , \hat{x}_2 and \hat{x} from f_1 , f_2 and f by decoder D_x (represented by the purple arrow).
- It regenerates RGB image patch \hat{r}_1 from f_1 using a decoder D_r (represented by the blue arrow).
- It also recovers the mixed X-ray image patch \bar{x} by computing $\bar{x} = \hat{x}_1 + \hat{x}_2$ (represented by the orange arrows).

Architectures: We note that this auto-encoder is a modified version of the connected auto-encoder originally proposed in [13, 14], allowing us to cater for the more challenging scenarios where one has access to a single RGB image – in lieu of two RGB images – to aid in the X-ray separation problem. It turns out that the ability of this new approach to carry out the more challenging X-ray separation process also demands the use of more complex training loss functions as described in the sequel. We model the encoders E_r , E_x and decoders D_r and D_x associated with the connected auto-

¹This operation is motivated by the success of similar operations in our previous work [13,14]

encoder in Fig. 1 using 3-layer CNN. This choice – which was also adopted in [13,14] – is due to the fact that CNN normally outperforms other deep learning structures in image processing applications. For E_r , D_r and D_x , each CNN layer is followed by batch normalization and ReLU activation layers.

Learning algorithm: We train the parameters of the different networks composing the connected auto-encoder structure using a composite loss function consisting of a reconstruction loss and an exclusion loss as follows:

$$L = L_1 + \lambda \cdot L_2 \quad (5)$$

where L_1 is the reconstruction loss, L_2 is the exclusion loss, and λ is a regularization parameter.

Our reconstruction loss measures the discrepancy between the mixed X-ray image patch reconstruction and the original one, a synthesized mixed X-ray image patch and the original one, and the visible image reconstruction and the original one, as follows:

$$L_1 = \|r_1 - \hat{r}_1\|_F + \|x - \hat{x}\|_F + \|x - \bar{x}\|_F, \quad (6)$$

where $\hat{r}_1 = D_r(E_r(r_1))$, $\hat{x} = D_x(E_x(x))$, $\bar{x} = D_x(E_r(r_1)) + D_x(E_x(x) - E_r(r_1))$ and $\|\cdot\|$ denotes the Frobenius norm.

In turn, our exclusion loss – which is adopted from [17] – measures the correlation between two edge maps at multiple spatial resolutions as follows:

$$L_2 = \sum_{n=1}^N \|\Psi(f^{\downarrow n}(\lambda_1 |\nabla \bar{x}_1|), f^{\downarrow n}(\lambda_2 |\nabla \hat{x}_2|))\|_F, \quad (7)$$

where $\Psi(x_1, x_2) = \tanh(x_1) \odot \tanh(x_2)$, \odot is element-wise multiplication, $f^{\downarrow n}()$ denotes the downsampling operation by a factor of 2^{n-1} with bilinear interpolation, $\nabla \bar{x}_1$ and $\nabla \hat{x}_2$ denote the gradients of \bar{x}_1 and \hat{x}_2 , respectively, and \bar{x}_1 is the image patch corresponding to a labeled X-ray image of the surface painting, shown in Fig. 4 (c). Here, \bar{x}_1 in Fig. 4 (c) is mostly derived from the mixed X-ray image, though a grayscale of the RGB image associated with the surface painting replaces the central portion of the mixed X-ray image because of the observation in Fig. 4 (b) that the content of the hidden layer is mainly located in this area. We set $N = 3$, $\lambda_1 = \sqrt{\frac{\|\hat{x}_2\|_F}{\|\bar{x}_1\|_F}}$, and $\lambda_2 = \sqrt{\frac{\|\bar{x}_1\|_F}{\|\hat{x}_2\|_F}}$.

We note – as alluded to above – that the inclusion of an exclusion loss in the overall loss function is critical to aid the separation process. Its inclusion is also a distinguishing feature between our new algorithm compared our previous algorithm originally proposed in [13, 14].

Then, stochastic gradient descent (SGD) with the ADAM optimization strategy is utilized to solve (5) where learning rate is chosen to be equal to 10^{-6} .

Initialization: Finally, we initialize the parameters of auto-encoders E_r , E_x , D_r and D_x using a particular warm start procedure in order to avoid the learning process converging to undesired minima. In particular, we adopt the initialization process showcased in Fig. 3 by leveraging our previous X-ray image separation approach applicable to double-sided paintings [13,14]. Here, mixed X-ray image patch x_d and RGB

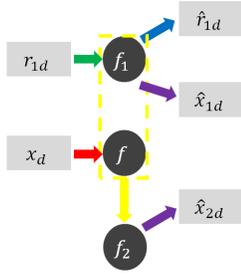


Fig. 3. Diagram of initialization training network.

image patch of one single side r_{1d} of a double-sided painting are used as inputs, while r_{1d} and separation results x_{1d} and x_{2d} obtained using [14] are used as the labels for the outputs. The auto-encoders E_r , E_x , D_r and D_x are then trained by minimizing the loss function given by

$$L_{init} = \|r_{1d} - \hat{r}_{1d}\|_F + \|x_{1d} - \hat{x}_{1d}\|_F + \|x_{2d} - \hat{x}_{2d}\|_F, \quad (8)$$

where \hat{r}_{1d} , \hat{x}_{1d} and \hat{x}_{2d} are the outputs of the connected auto-encoders using double-sided paintings, and correspond to the reconstructed RGB image patch, X-ray image patch of the front side, and X-ray image patch of the rear side, respectively.

2) *Step 2: X-Ray Image Refinement*: In the second step, we obtain final separation results x_1 and x_2 by evenly splitting the error map $e = x - \hat{x}_1 - \hat{x}_2$ into \hat{x}_1 and \hat{x}_2 , shown as follows:

$$x_1 = \hat{x}_1 + e/2, \quad x_2 = \hat{x}_2 + e/2. \quad (9)$$

The rationale for adopting this procedure has to do with the fact that the use of the exclusion loss causes the initial estimates \hat{x}_1 and \hat{x}_2 not to have any shared mutual content. While eliminating mutual information improves the reconstruction accuracy associated with connected auto-encoders, it does not accurately reflect the true situation with X-ray images. For example, an artist may have incorporated parts of an otherwise abandoned composition within a final (visible) painting or passages in the final painting may directly overlie abandoned passages. Such regions of the X-ray image are thus actually 'shared' between x_1 and x_2 , although it must be borne in mind that x_1 and x_2 are hypothetical images with x_1 representing a hypothetical X-ray image associated with the finally visible painting and x_2 containing all other features that are not visible at the surface of the painting. A further complication is that much detailed information including the canvas pattern visible in the original mixed X-ray image disappear in the separation results of the first step. While it could be argued that details such as the canvas pattern are not associated with the final visible image and should therefore appear in x_2 , when studying X-ray images of paintings researchers are familiar with being able to see features associated with the canvas of panel support of the painting. The refinement operation proposed here helps to remedy these issues and results in more

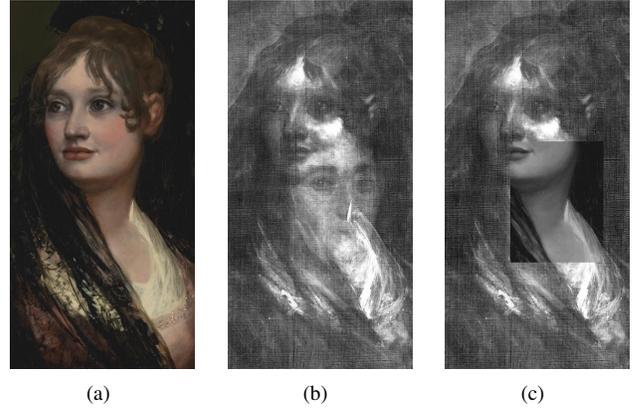


Fig. 4. Detail of Isabel's face in *Doña Isabel de Porcel*: (a) Visual RGB image of the surface painting. (b) Mixed X-ray image. (c) Labeled X-ray image of the surface painting.

realistic X-ray image separations as shown in the subsequent results.

IV. EXPERIMENTAL RESULTS

We now showcase that our proposed approach can lead to a plausible decompositions of mixed X-ray images of paintings with hidden decompositions by applying it to the painting *Doña Isabel de Porcel* by Francisco Goya shown in Fig. 1. In particular, in this experiment a small area of the whole painting, wherein both the content of surface painting and hidden painting are obvious in the X-ray image, is utilized to test the proposed method (see Fig. 4). The size of the images in Fig. 4 is 500×1000 pixels. The patch size is set to be 50×50 , and we have 45 overlapping pixels associated with both horizontal and vertical dimensions, resulting overall in 17381 patches.

Fig. 5 (a) and (b) shows the separation results after the first stage of the calculation. The re-synthesized mixed X-ray – obtained by adding the individual ones in Fig 5(a) and (b) – together with the error map – obtained by subtracting the re-synthesized mixed-X-ray image from the original mixed X-ray – are shown in Fig. 5 (c) and (d). After dividing the detailed information in Fig. 5 (d) into x_1 and x_2 , the final separation results, after the second refinement step, are shown in Fig. 5 (e) and (f). While it is clear there are still some remaining issues with the image separation, particularly around Isabel's headdress, these final images have more of the character that would be anticipated for X-ray images and are likely to feel more familiar and therefore be more appealing to end users.

The separation issues around Isabel's headdress partially reflect a potential problem with the mapping function, $F(\cdot)$. X-ray images of paintings are dominated by passages painted with heavy metal-containing pigments, particularly those containing lead. In Old Master paintings, the most common white pigment is the lead-based pigment, lead white, and thus typically areas of pale coloured paint at the surface of the painting will appear bright in the X-ray image, particularly if thickly painted. By contrast dark paint passages tend frequently to



Fig. 5. Separation results. (a) Initial separated X-ray image of the surface painting. (b) Initial separated X-ray image of the hidden painting. (c) Synthetic mixed X-ray image using (a) and (b). (d) Error map of the synthetic mixed X-ray image. (e) Refined separated X-ray image of the surface painting. (f) Refined separated X-ray image of the hidden painting.

also appear dark in X-ray images. However, in Isabel's black headdress, a small amount of a copper-containing pigment has been used in the black paint and this appears to cause an artefact in the region of the headdress. In future work, it is hoped to leverage the existence of other imaging modalities in the image separation problem which may help to address issues such as this, and the problem of the elimination of mutual information.

V. CONCLUSION

X-radiography is a useful tool in the technical study of artworks as, amongst its other benefits, it is capable of providing insights into hidden compositions and *pentimenti*. However, when hidden designs exist under the visible surface the resulting X-ray images contain features associated with both surface and sub-surface features (they are 'mixed') and as a result it is difficult for experts to visually interpret such images. To improve the utility of these X-ray images, it is desirable to separate the content into two (hypothetical) images, each pertaining to only one layer or composition. This paper proposed a novel self-supervised learning algorithm based on a two-step procedure. In the first step, a connected auto-encoder structure as well as a joint reconstruction and

exclusion loss are proposed to separate the mixed X-ray image. In the second step, the error map is evenly allocated into the separated images to produce separated X-ray images of more familiar appearance. This proposed method is demonstrated with the experiments on images from the painting *Doña Isabel de Porcel* by Francisco de Goya, leading to the first visualisation of the hypothetical X-ray image associated with the concealed male in the painting.

REFERENCES

- [1] X. Huang, E. Uffelmann, O. Cossairt, M. Walton and A. K. Katsaggelos, "Computational Imaging for Cultural Heritage: Recent developments in spectral imaging, 3-D surface measurement, image relighting, and X-ray mapping," *IEEE Signal Process. Mag.*, vol. 33, no. 5, pp. 130-138, Sept. 2016
- [2] H. Yang, J. Lu, W. Brown, I. Daubechies, and L. Ying, "Quantitative canvas weave analysis using 2-D synchrosqueezed transforms: Application of time-frequency analysis to art investigation," *IEEE Signal Process. Mag.*, vol. 32, no. 4, pp. 55-63, Jul. 2015.
- [3] J. Lang and A. Middleton, "Radiography of cultural material," *Routledge*, 2005.
- [4] Joseph Padfield, David Saunders, John Cupitt, and Robert Atkinson, "Improvements in the acquisition and processing of X-ray images of paintings," *National Gallery Technical Bulletin*, 23:62-75, 2002.
- [5] S. van Heughten, "Radiographic images of Vincent van Gogh's paintings in the collection of the Van Gogh museum," *Van Gogh Museum Journal* (1995) 63-85.
- [6] G. Van der Snickt, A. Martins, J.K. Delaney, K. Janssens, J. Zeibel, M. Duffy, et al., "Exploring a hidden painting below the surface of Rene Magritte's *Le Portrait*," *Appl. Spectrosc.* 70 (2016) 57-67.
- [7] P. Kammerer, E. Zolda, R. Sablatnig, "Computer aided analysis of underdrawings in infrared reflectograms," in *Proceedings of the Fourth International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*, 2003, pp. 19-27.
- [8] G. J. Tserevelakis, I. Vrouvaki, P. Siozos, K. Melessanaki, K. Hatzigiannakis, C. Fotakis, et al., "Photoacoustic imaging reveals hidden underdrawings in paintings," *Sci. Rep.*, 7 (2017) 747.
- [9] *Doña Isabel de Porcel* by Francisco de Goya. <https://www.nationalgallery.org.uk/paintings/francisco-de-goya-dona-isabel-de-porcel>
- [10] N. Deligiannis, J. F. C. Mota, B. Cornelis, M. R. D. Rodrigues and I. Daubechies, "Multi-Modal Dictionary Learning for Image Separation With Application in Art Investigation," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 751-764, Feb. 2017.
- [11] Z. Sabetsarvestani, B. Sober, C. Higgitt, I. Daubechies, and M. R. D. Rodrigues, "Artificial intelligence for art investigation: Meeting the challenge of separating X-ray images of the Ghent Altarpiece," *Science Advances*, vol. 5, no. 8, pp. 751-764, 2019.
- [12] Z. Sabetsarvestani, F. Renna, F. Kiraly, and M. R. D. Rodrigues, "Source Separation with Side Information Based on Gaussian Mixture Models With Application in Art Investigation," submitted to *IEEE Trans. Signal Process.*
- [13] W. Pu, B. Sober, N. Daly, C. Higgitt, I. Daubechies and M. Rodrigues, "A connected auto-encoders based approach for image separation with side information: with applications to art investigation," in *ICASSP*, Spain, 2020.
- [14] W. Pu, B. Sober, N. Daly, Z. Sabetsarvestani, C. Higgitt, I. Daubechies and M. Rodrigues, "Image Separation with Side Information: A Connected Auto-Encoders Based Approach," submitted to *IEEE Trans. Image Process.*
- [15] Spring, M., Billinge, R., Treves, L., Von Aderkas, N., Higgitt, C., van Loon, A., Dik, J., 2016. Goya's Portraits in the National Gallery: their Technique, Materials and Development. *National Gallery Technical Bulletin* 37, 78-104.
- [16] Closer to Van Eyck. <http://clostertovaneyck.kikirpa.be/ghentaltarpiece>. Accessed: 2019-01-17.
- [17] X. Zhang, R. NG and Q. Chen, "Single Image Reflection Separation with Perceptual Losses" in *CVPR*, pp. 4786-4794, 2018.