Image Separation with Side Information: A Connected Auto-Encoders Based Approach

By Wei Pu, Barak Sober, Nathan Daly, Zahra Sabetsarvestani, Catherine Higgitt, Ingrid Daubechies, and Miguel R.D. Rodrigues.

I. INTRODUCTION

- **Non-invasive and Non-destructive:** macro X-ray fluorescence (MA-XRF) scanning and hyperspectral imaging …
- **Challenging tasks:** crack detection, material identification, brush stroke style analysis, canvas pattern or stretcher bar removal automated canvas weave analysis, and improved visualization of concealed features or under-drawing.

I. INTRODUCTION

- **X-ray images:** Provide insights into an artists technique and working methods, for example revealing the paintings stratigraphy and information about the painting support or even some indication of the pigments used.

- **X-ray images: However,** the X-ray image of a painting particularly those with design changes, areas of damage, hidden paintings, or paintings on both the front and reverse sides of the support will inevitably contain a mix or blend of these various features, making it difficult for experts to interpret.

I. INTRODUCTION

- **The propose:**

  1. First, self-supervised learning, connected auto-encoders that extract features from the RGB images in order to (a) reproduce both of the original RGB images, (b) reconstruct the associated separated X-ray images, and (c) regenerate the mixed X-ray image.

  2. Second, tune these auto-encoders based on the use of a composite loss function involving reconstruction losses, energy losses and dis-correlation losses.

I. INTRODUCTION

- **The propose:**

  3. Third, analysis of the effect of various hyper-parameters associated with our separation method on performance.

  4. Finally, we apply to a real dataset, showcasing state-of-the-art results over competing methods. The dataset relates to images taken from the double-sided wing panels of the Ghent Altarpiece

I. INTRODUCTION

Fig. 1. Two double-sided panels from the *Ghent Altarpiece* [51]: (left) visible RGB image of the front side, (centre) visible RGB image of the back side, (right) mixed X-ray image.

II. PROBLEM FORMULATION

- **Goal:** Separate the mixed X-ray image into two components where one component would contain features associated with the image on the front panel and the rear panel.

- **Methodology:** Dividing these images into several smaller patches that overlap with respect to the vertical and horizontal dimensions of the image.

II. PROBLEM FORMULATION

- **Linear mixing assumption:** x denotes a mixed X-ray image patch and let x1 and x2 separated X-ray image patches corresponding to the front and rear.

$$x \approx x_1 + x_2.$$

- **Further** assume that there is a mapping F that is approximately able to convert an image patch in the RGB domain into an image patch in the X-ray.

$$x \approx \mathcal{F}(r_1) + \mathcal{F}(r_2).$$

- Where mapping function F has been modelled via a 7-layer (CNN).

II. PROBLEM FORMULATION

- Minimizing the error between the sum of the two separated X-ray image patches and the original mixed X-ray image patch:

$$\min_{\mathcal{F}} \|x - \mathcal{F}(r_1) - \mathcal{F}(r_2)\|_F \, .$$

where || . || denotes the Frobenius norm.

- **Problem:** lack of constraints on the structure of x1 and x2 , the individual X-ray images obtained using the can be highly related to the corresponding RGB images.

III. PROPOSED APPROACH

A. Connected auto-encoder structure

- **Approach is based on the use of auto-encoders.**

1. **Encoder Er** is used to extract features f1 and f2 from the RGB image patches r1 and r2.

2. **Decoder Dr** is used to convert the features f1 and f2 onto an estimate of the RGB image patches $\hat{r}1$ and $\hat{r}2$.

3. **Decoder Dx** is also used to convert the features f1 , f2 and f onto an estimate of the X-ray image patches $\hat{x}1$, $\hat{x}2$, and $\hat{x}$ respectively, where f denotes a feature vector associated with the mixed X-ray image patch x;

4. **One ’addition’ process** is used in the feature and X-ray domain to get another version of the mixed X-ray $\bar{x} = \hat{x}1 + \hat{x}2$ and the corresponding feature map $f = f1 + f2$.

III. PROPOSED APPROACH

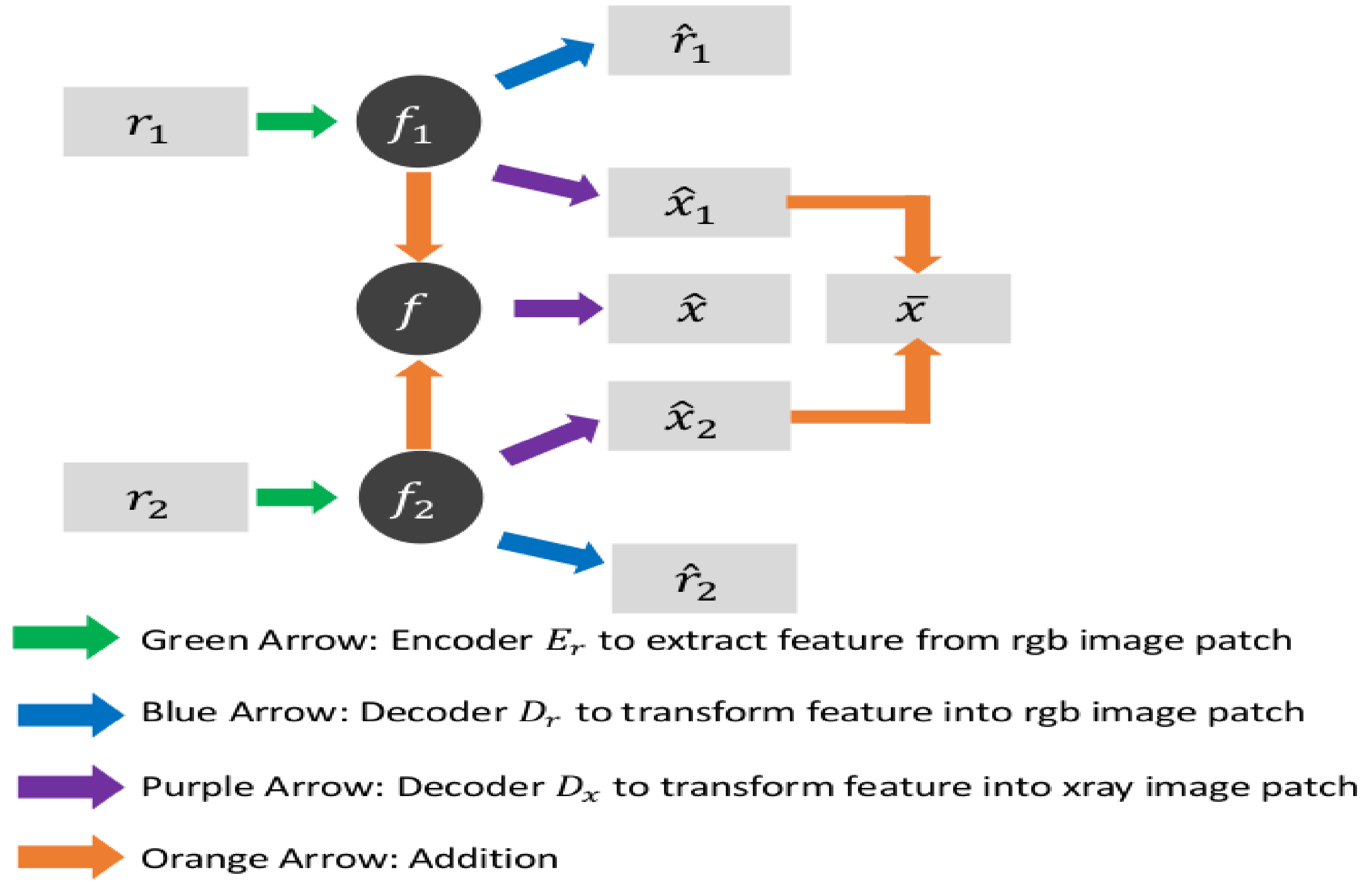


Fig. 2. Block diagram of the proposed method.

III. PROPOSED APPROACH

A. Connected auto-encoder structure

- **Shared-Features Assumption**: For a certain pair of X-ray image patch x1 and its corresponding RGB image patch r1, one postulates there is some **latent feature vector** f1 such that x1 = Dx(f1) and r1 = Er(f1) (rear panel, as well).
- **Linear-Feature Assumption:** The feature map f associated with a mixed X-ray patch x corresponds to the sum of the feature map f1 associated with X-ray image of patch x1 and feature map f2 associated with X-ray image of patch x2. The feature map can be used to reconstruct $\hat{x}$, Dx ; i.e., $\hat{x}$ = Dx(f).

III. PROPOSED APPROACH

- B.

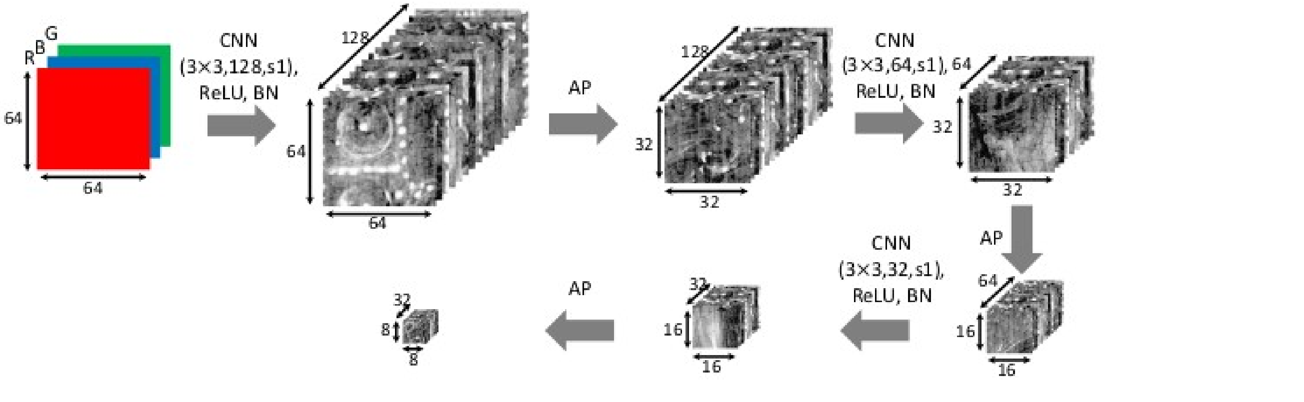


Fig. 3. Encoder E_r is modelled as 3-layer 2-dimensional CNNs , wherein each CNN layer is followed by batch normalization (BN), ReLU activation as well as average pooling (AP) layers.

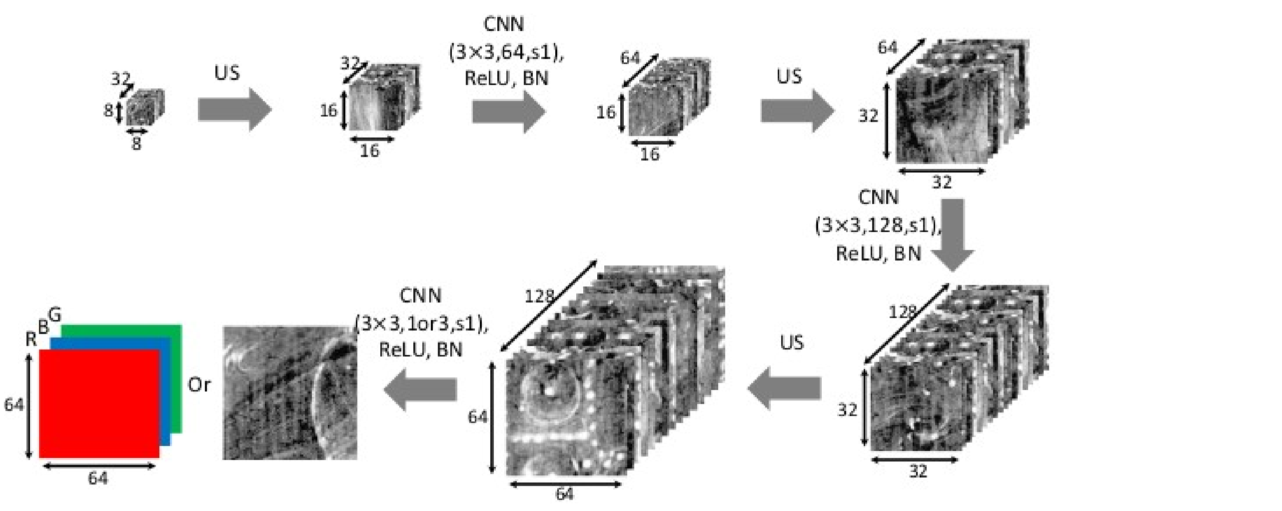


Fig. 4. Decoders D_x and D_r are modelled as 3-layer CNNs, wherein each CNN layer is followed by batch normalization (BN), ReLU activation as well as upsampling (US) layers.

III. PROPOSED APPROACH

C. Learning Algorithm

1. $$L_1 = \|r_1 - \hat{r}_1\|_F + \|r_2 - \hat{r}_2\|_F .$$

2. $$L_2 = \|x - \hat{x}\|_F ,$$

where $\hat{x} = D_x(E_r(r_1) + E_r(r_2))$.

3. $$L_3 = \|x - \bar{x}\|_F ,$$

where $\bar{x} = \hat{x}_1 + \hat{x}_2 = D_x(E_r(r_1)) + D_x(E_r(r_2))$.

III. PROPOSED APPROACH

C. Learning Algorithm

- **(Problems)**

We have also noted that these individual losses by themselves do not entirely promote reasonable results in view of the fact that:

1) It is possible to obtain degenerate results such as $\hat{x}_1 \approx x$ and $\hat{x}_2 \approx 0$ or $\hat{x}_1 \approx 0$ and $\hat{x}_2 \approx x$ by using these loss functions alone.
2) It is also possible to obtain results where a portion of the content of the X-ray image from one side appears in the X-ray image of the other side (and vice versa).

III. PROPOSED APPROACH

C. Learning Algorithm

4.

$$L_4 = \|\hat{x}_1\|_F^2 + \|\hat{x}_2\|_F^2 .$$

5.

$$L_5 = C^2(f_1, f_2), \tag{8}$$

where $C(f_1, f_2)$ denotes the Pearson correlation coefficient between f_1 and f_2 given by

$$C(f_1, f_2) = \frac{\sum(f_{v1} - \mu_1)(f_{v2} - \mu_2)}{\sqrt{\sum(f_{v1} - \mu_1)^2 \sum(f_{v2} - \mu_2)^2}}. \tag{9}$$

III. PROPOSED APPROACH

C. Learning Algorithm

$$L_{total} = L_1 + \lambda_1 \cdot L_2 + \lambda_2 \cdot L_3 + \lambda_3 \cdot L_4 + \lambda_4 \cdot L_5,$$

Where λ1 , λ2 , λ3 and λ4 are the hyper-parameters corresponding to the losses L2 , L3 , L4 and L5 , respectively.

Stochastic gradient descent (SGD) algorithm with the ADAM optimization strategy with learning rate 0.0001.

IV. EXPERIMENTAL RESULTS

A. Datasets

- Ghent Altarpiece by Hubert and Jan van Eyck. This large, complex 15th-century polyptych altarpiece comprises a series of panels including panels with a composition on both sides that we use to showcase the performance of our algorithm on real mixed X-ray data.

IV. EXPERIMENTAL RESULTS

A. Datasets

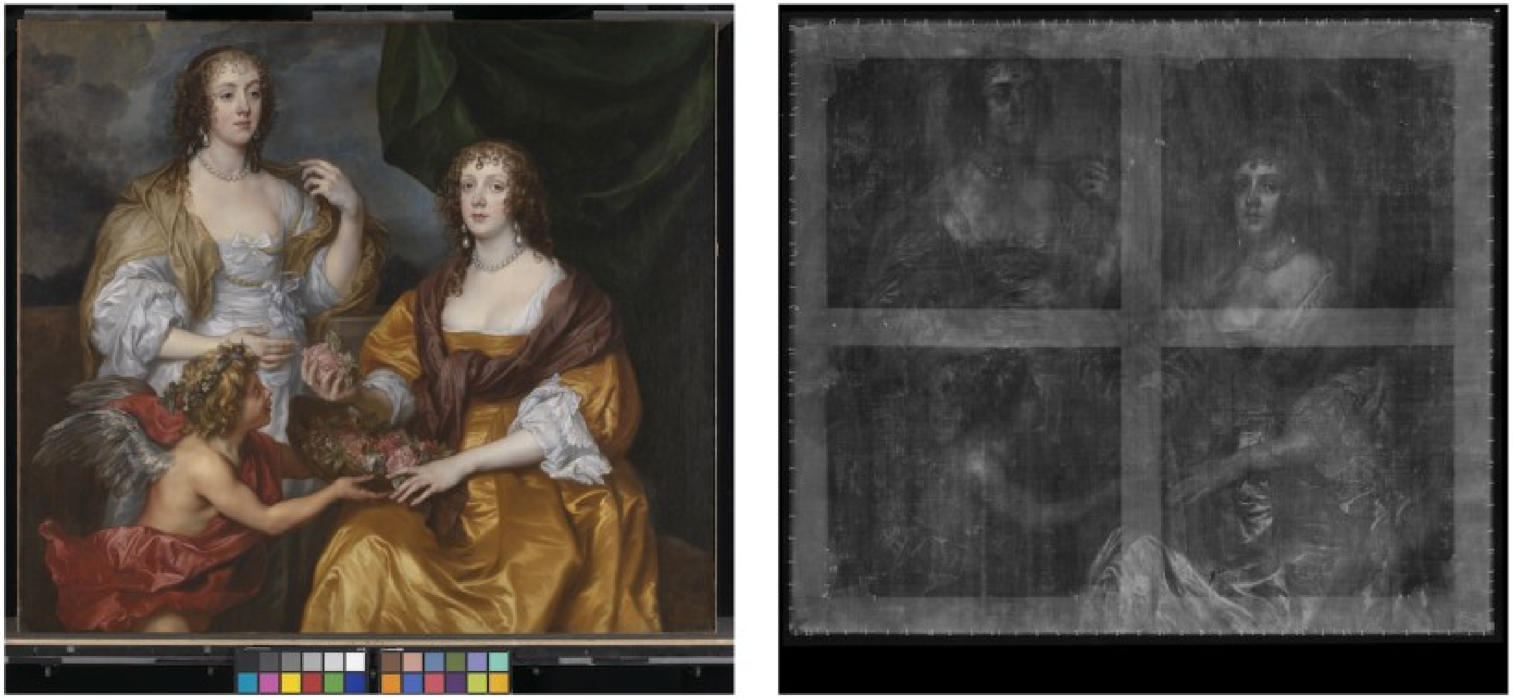
- Kitchen Scene with Christ in the House of Martha andn Mary by Diego Velzquez. This one-sided canvas painting was used to showcase the performance of algorithm on synthetically mixed X-ray data.

(a) (b)

IV. EXPERIMENTAL RESULTS

A. Datasets

- Lady Elizabeth Thimbelby and Dorothy, Viscountess Andover by Anthony Van Dyck. This canvas painting, also one-sided.



(a) (b)

IV. EXPERIMENTAL RESULTS

B. Hyper-parameter Selection Protocol

- First, we report results for the optimal values for the hyper-parameters $\lambda 1$ and $\lambda 2$ with the hyper-parameters $\lambda 3$ and $\lambda 4$ set to be equal to zero.
- Second, we report the results for the optimal values for the hyper-parameters $\lambda 3$ and $\lambda 4$ with the hyper- parameters $\lambda 1$ and $\lambda 2$ set to be equal to their optimal values from the first optimization step.
- $\lambda 1 \in [0, 10]$, $\lambda 2 \in [0, 10]$, $\lambda 3 \in [1, 10]$ and $\lambda 4 \in [0.1, 0.5]$ in steps of 0.2, 0.2, 0.2 and 0.02, respectively.

IV. EXPERIMENTAL RESULTS

C. Experiments with Synthetically Mixed X-ray Data

- **Experiment set-up (Lady Elizabeth Thimbelby and Dorothy):** Each such image is of size 1100 × 1100 pixels. These images were then further divided into patches of size 64×64 pixels with 56 pixels overlap (both in the horizontal and vertical direction), resulting in 11,236 patches.
- Random initialization.
- $$MSE = \frac{1}{2R} \sum_{r=1}^{R} \left(||X_1 - \hat{X}_1^r||_F + ||X_2 - \hat{X}_1^r||_F \right),$$

IV. EXPERIMENTAL RESULTS

C. Experiments with Synthetically Mixed X-ray Data

(a) (b) (c) (d) (e)

Fig. 7. Images used for hyper-parameter selection. (a). First RGB image. (b). Second RGB image. (c). X-ray image corresponding to first RGB image. (d). X-ray image corresponding to second RGB image. (e). Synthetically mixed X-ray image.

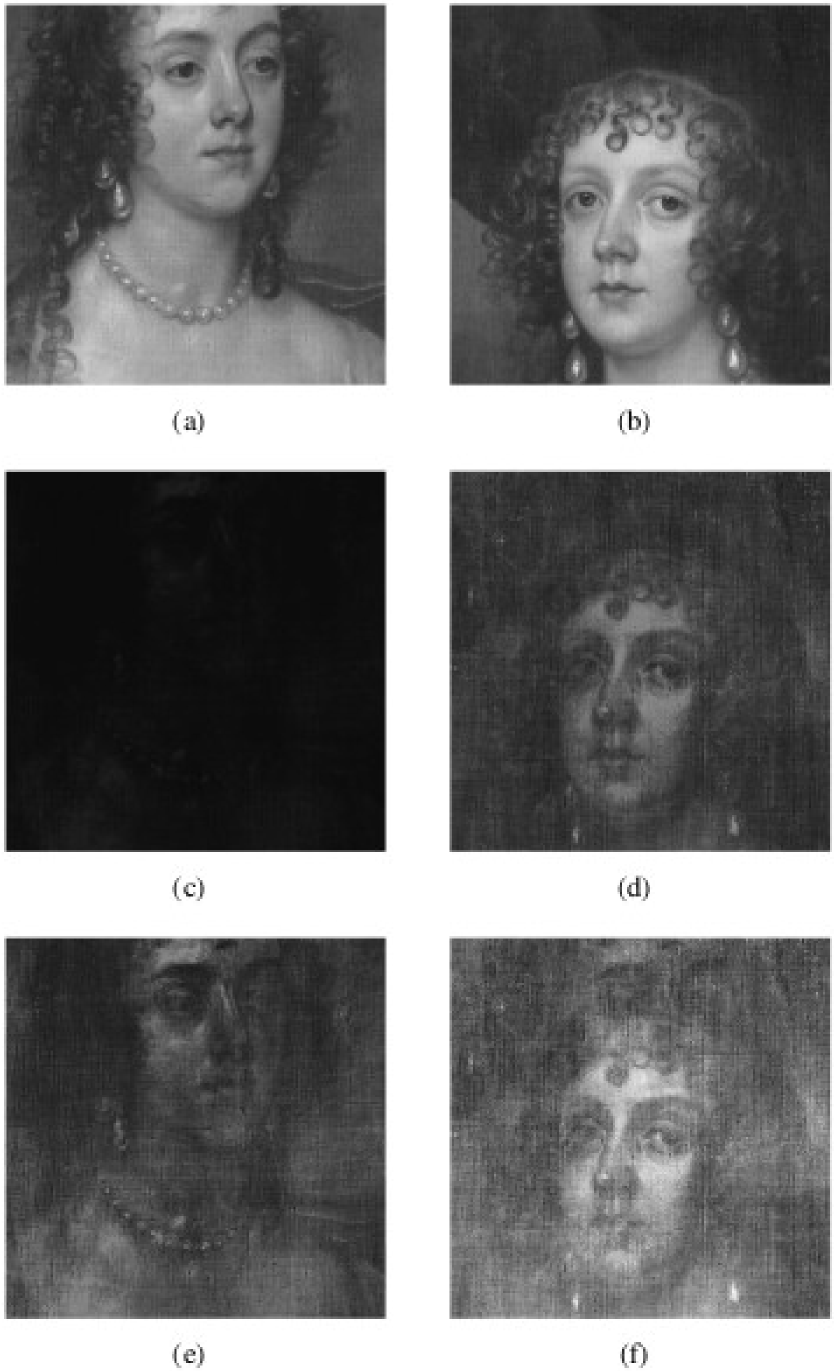


Fig. 9. X-ray separation results: (a) and (b) Separated X-ray images for $\lambda_1 = 0.1$, $\lambda_2 = 0.1$, and $\lambda_3 = \lambda_4 = 0$; (c) and (d) Separated X-ray images for $\lambda_1 = 10$, $\lambda_2 = 1$, and $\lambda_3 = \lambda_4 = 0$; (e) and (f) Separated X-ray images for $\lambda_1 = 1$, $\lambda_2 = 10$, and $\lambda_3 = \lambda_4 = 0$.

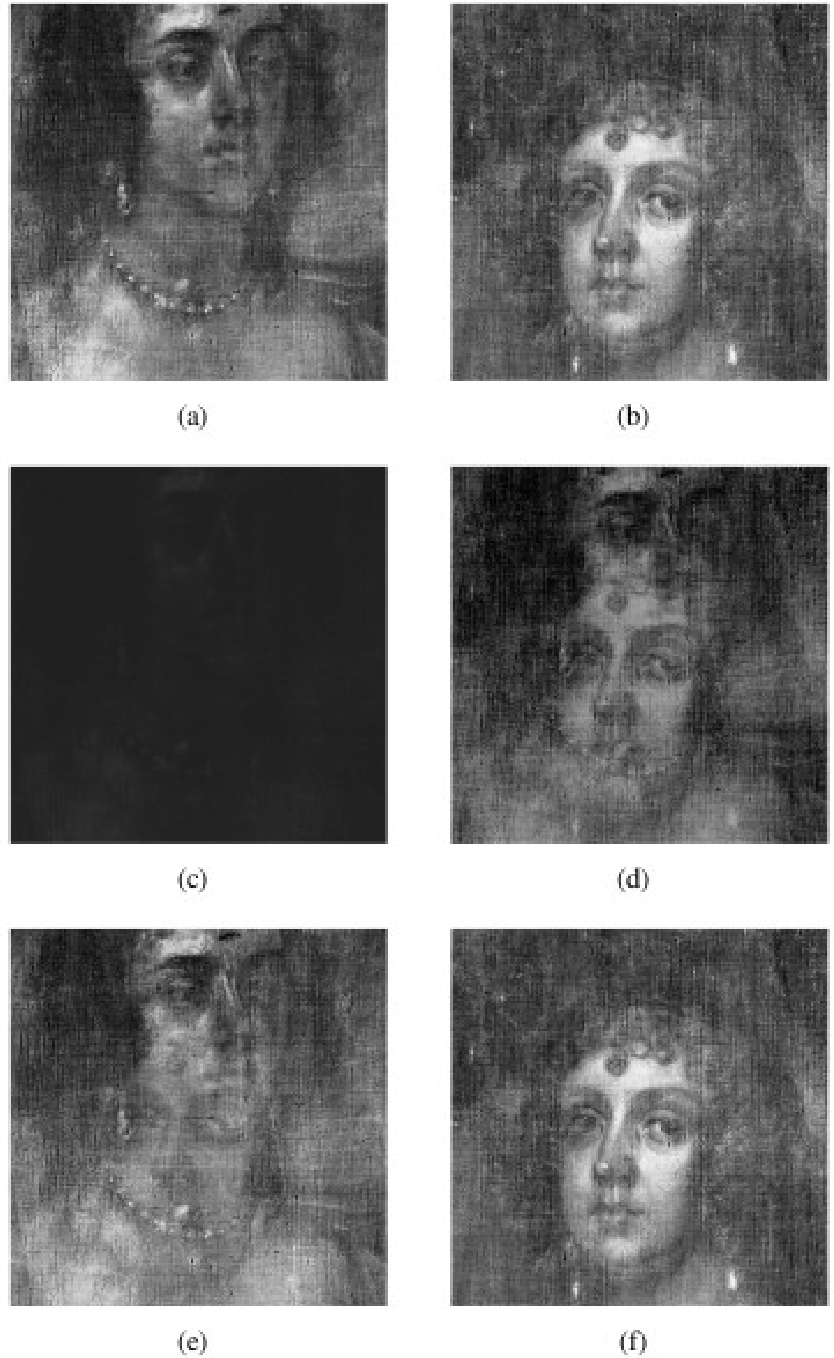


Fig. 10. X-ray separation results: (a) and (b) Case i: Separated X-ray images for $\lambda_1 = 3$, $\lambda_2 = 5$, and $\lambda_3 = \lambda_4 = 0$; (c) and (d) Case ii: Separated X-ray images for $\lambda_1 = 3$, $\lambda_2 = 5$, and $\lambda_3 = \lambda_4 = 0$; (e) and (f) Case iii: Separated X-ray images for $\lambda_1 = 3$, $\lambda_2 = 5$, and $\lambda_3 = \lambda_4 = 0$.

IV. EXPERIMENTAL RESULTS

C. Experiments with Synthetically Mixed X-ray Data

- **Experiment set-up (Kitchen Scene with Christ in the House of Martha andn Mary)** The images which are of size 1000 × 1000 pixels were divided into patches of size 64 × 64 pixels with 56 pixels overlap (both in the horizontal and vertical direction), resulting in 13,924 patches.

IV. EXPERIMENTAL RESULTS

C. Experiments with Synthetically Mixed X-ray Data

Fig. 13. Images used for synthetic data experiments. (a). First RGB image. (b). Second RGB image. (c). X-ray image corresponding to first RGB image. (d). X-ray image corresponding to second RGB image. (e). Synthetically mixed X-ray image.

IV. EXPERIMENTAL RESULTS

C. Experiments with Synthetically Mixed X-ray Data

- Best results: λ1 = 3, λ2 = 5, λ3 = 2, and λ4 = 0.3.

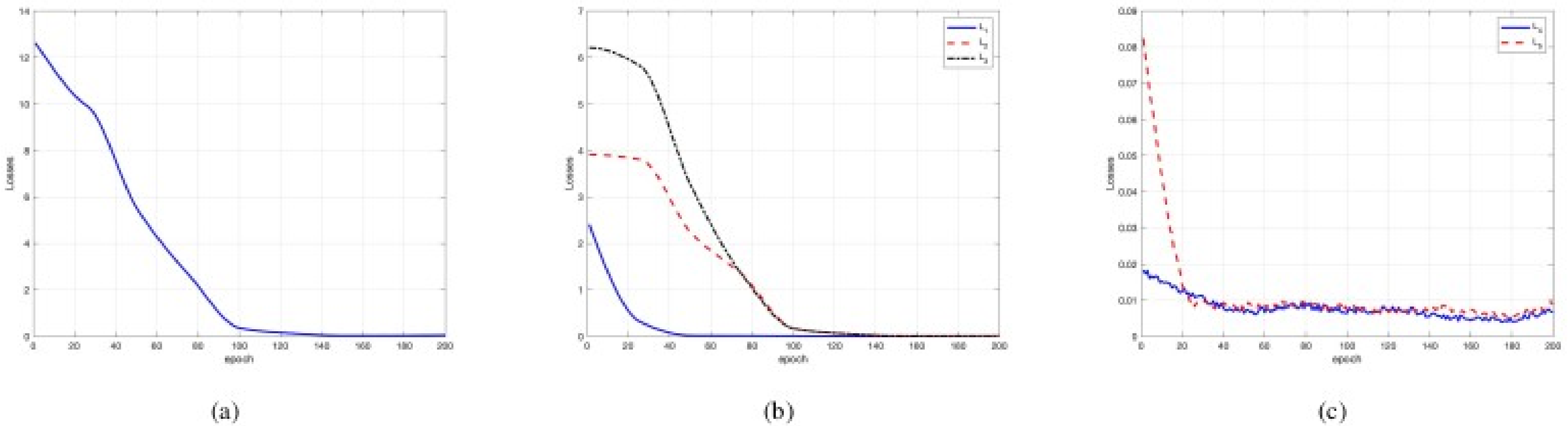


Fig. 14. Losses vs. number of epochs on synthetic data. (a). L_{total}. (b). L_1, L_2 and L_3. (c). L_4 and L_5.

IV. EXPERIMENTAL RESULTS

Fig. 15. Reconstructed images vs. number of epochs on synthetic data experiments. Columns 1 to 7 correspond to reconstructed result under 1st, 4th, 10th, 50th, 100th, 150th and 200th epoch, respectively. Rows 1 to 2 correspond to the reconstructed RGB images. Rows 3 to 4 correspond to the reconstructed X-ray images.

IV. EXPERIMENTAL RESULTS

D. Experiments with Real Mixed X-ray Data

- **Set-up:** In this experiment, we use a small area of size1000 × 1000 pixels from the Ghent Altarpiece. The previous procedure was again followed: the two RGB images and the corresponding mixed X-ray image were divided into patches of size 64×64 pixels with 56 pixels overlap (both in the horizontal and vertical direction), resulting in 13,924 patches.
- Once again, hyper-parameter values $\lambda 1 = 3$, $\lambda 2 = 5$, $\lambda 3 = 2$, and $\lambda 4 = 0.3$.

IV. EXPERIMENTAL RESULTS

D. Experiments with Real Mixed X-ray Data

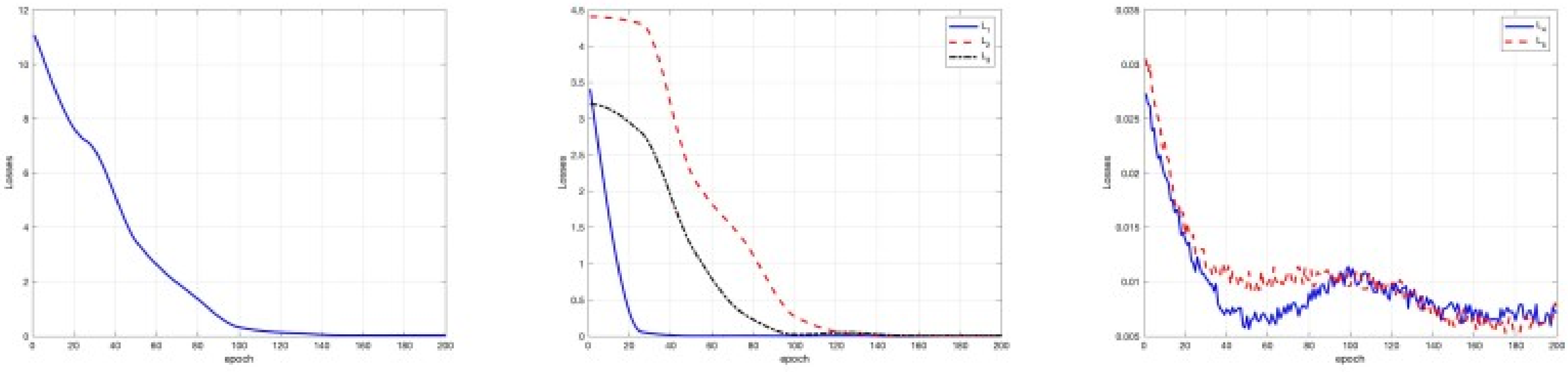


Fig. 18. Losses vs. number of epochs on real data experiments. (a). L_{total}. (b). L_1, L_2 and L_3. (a). L_4 and L_5.

IV. EXPERIMENTAL RESULTS

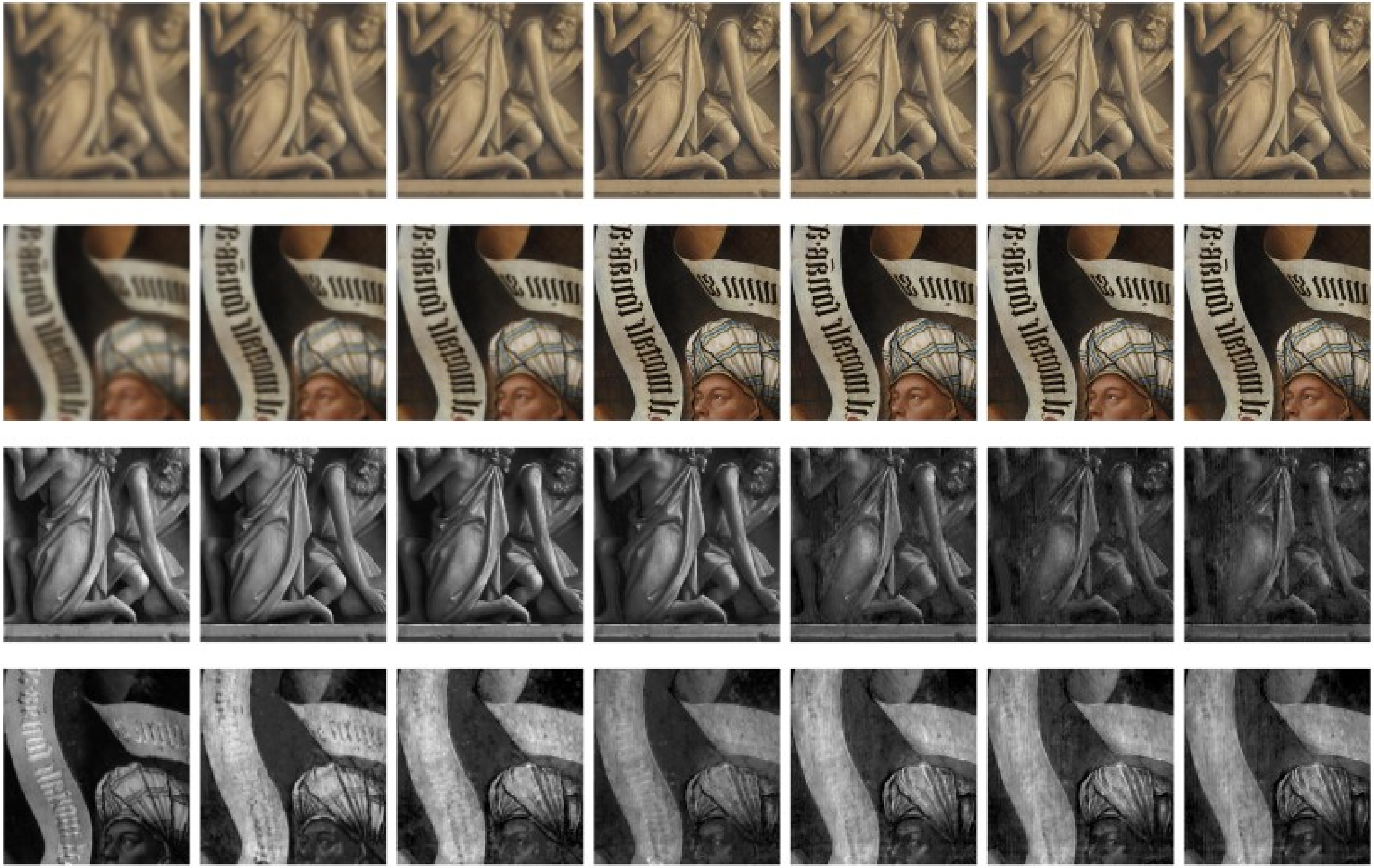


Fig. 19. Reconstructed images vs. number of epochs on real data experiments. Columns 1 to 7 correspond to reconstructed result under 1st, 4th, 10th, 50th, 100th, 150th and 200th epoch, respectively. Rows 1 to 2 correspond to the reconstructed RGB images. Rows 3 to 4 correspond to the reconstructed X-ray images.

V. CONCLUSION

- Improvement the utility of X-ray images in studying and conserving artworks.
- This approach allows image separation without the need for labelled data.
- The results from this image separation also maintain features of the support, such as wood grain and canvas weave, that are not readily apparent in the RGB images