A deep learning approach to crack detection in panel paintings

R. Sizyakin[1], B. Cornelis[2], L. Meeus[3], M. Martens[4], V. Voronin[1] and A. Pižurica[3].

[1]Department of Radio-Electronics Systems, Don State Technical University, Russian Federation .

[2]Department of Electronics and Informatics, Vrije Universiteit Brussel, Belgium.

[3]Department of Telecommunications and Information Processing, TELIN-IPI-imec, Ghent University, Belgium.

[4]Department of Art, Music and Theatre Sciences, Ghent University, Belgium.

Abstract— **The accurate detection of cracks in paintings, which generally portray rich and varying content, is a challenging task. Furthermore, traditional crack detection methods are often not well suited for recent acquisitions of paintings as they are not designed for high-resolution images and do not fully exploit the information from the different imaging modalities at hand. In this work, we propose a fast crack detection framework that alleviates the aforementioned challenges. The method consists of a morphological filtering operation followed by a classification step by means of a convolutional neural network architecture. The proposed online method is capable of continuously learning from newly acquired visual data, thus further improving classification results as more data becomes available.**

1 Introduction

Paint cracking (or *craquelure*) is the most common type of deterioration encountered in old master paintings. Generally speaking, cracks appear in paint layers when pressure develops within or on it through the influence of various factors and cause the material to break [1]. The automatic detection of crack patterns is desirable for many reason. Most importantly, crack patterns can offer insights on the structural condition and conservation history of a painting [1]. Crack detection is also used as a preprocessing step for the digital restoration of paintings [2].

Many crack detection methods have been developed over the recent years, see e.g. [2–6] and the reference therein. Still, some important challenges remain especially in terms of feature selection, which often has to be adapted for different paintings, parameter tuning, and complexity, which limit practical applicability.

In this paper, we propose a new deep learning based approach for crack detection in paintings. The method consists of two processing stages: (i) a morphological filtering stage, and (ii) a classification stage. The morphological filtering essentially ensures that the amount of pixels to be classified by the CNN in the second stage is strongly reduced as only those pixels that are similar in structure to cracks are selected. In the classification stage, we employ a convolutional neural network (CNN). CNNs demonstrated recent success in many application where they have outperformed, often by a substantial margin, traditional machine learning algorithms [7–10]. We are not aware of any reported works that apply CNNs to crack detection in paintings. Some recent works applied CNN to detect cracks in roads [11, 12]. Our problem is, however, much more challenging not only because of the huge variability of cracks in paintings but also due to complex background and the fact that some painted details can closely resemble cracks. Therefore, our approach needs to incorporate multimodal data, which together with huge spatial resolution of digitized paintings poses additional challenges for the classifier.

2 Proposed approach

The first step in the proposed method is morphological filtering of the available modalities. This operation creates a preliminary crack map in a similar way as in [4, 5]. Each filtered result is followed by a thresholding step, producing binary images. The threshold is set based on the method of Otsu [13]. The binary maps are then combined into a single one using the logical OR operation. The morphological filtering step improves greatly the classification speed. The classifier is run only on pixels marked as crack in this first stage.

The input of our classification network consists of tensors of size $m \times m \times N$. These tensors are formed by concatenating $m \times m$ sized patches extracted from the N considered modalities (the three color channels of the visual macro-photographs, the single-channel infrared macro-photographs and X-ray images, along with their grayscale morphologically filtered results add up to $N = 9$ modalities.[1]). An input sample is represented by the tensor $x(u_1, u_2, v_0) \in R^{m\times m\times N}$, where u_1, u_2 are spatial coordinates and v_0 is the index that identifies the chosen modality. For our experiments, we fix $m = 8$, resulting in tensors with dimensions of $8 \times 8 \times 9$. The convolutions over v_0 are calculated in the first layer of the CNN as follows:

$$x_1(u_1, u_2, v_1) = \rho\big(x(u_1, u_2, v_0) * w_{v_1}(u_1, u_2, v_0)\big), \qquad (1)$$

where $x_1(u_1, u_2, v_1)$ is the feature map obtained by the convolution of $x(u_1, u_2, v_0)$ with $w_{v_1}(u_1, u_2, v_0)$, indexed by v_1 (in our architecture $1 \le v_1 \le 12$ for the first, $1 \le v_2 \le 24$ for second and $1 \le v_2 \le 48$ for the third convolution layers), and where ρ is an activation function. We choose the well-known rectified linear unit (ReLU) [7], defined as $\rho = max(0, x)$. All kernels are initialised randomly in the beginning of the training procedure. The core of our CNN architecture consists of performing a cascade of convolutions at each layer $j \ge 2$ as follows:

$$x_j(u_1, u_2, v_j) = \rho\big(x_{j-1}(\cdot, v_{j-1}) * w_{v_j}(\cdot)\big), \qquad (2)$$

where, as we navigate through the subsequent layers, the resolution of $x_j(u_1, u_2, v_j)$ progressively reduces [14, 15]. The last layer of the architecture consist of a *softmax* function, which ensures that all class probabilities sum up to 1.

[1]A disk-shaped structuring element with a diameter of 5 pixels was used for all experiments.

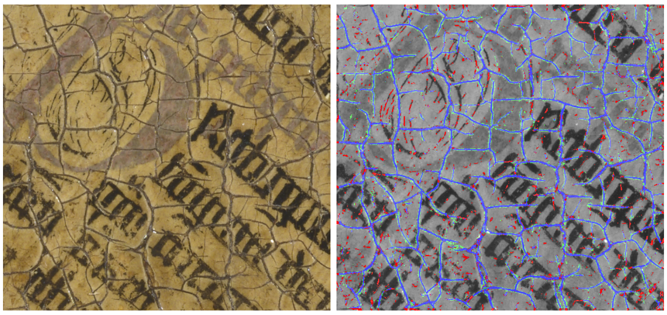

Figure 1: **Left**: A fragment of the panel *Virgin Annunciate*. **Right**: Results of the proposed method in comparison with BCTF. Blue: cracks identified by both methods; red: cracks detected by BCTF only; green: cracks detected by the proposed method only.

3 Experimental Results

All experiments are performed on a high-resolution multi-modal dataset of the *Ghent Altarpiece*, publicly available on the *Closer to Van Eyck* website[2]. In particular, we focus on three panels of the polyptych, named *Virgin Annunciate*, *Singing Angels* and *John Evangelist*. For comparsion we use the Bayesian Conditional Tensor Factorisation (BCTF) method from [3], which was also evaluated and compared to other methods in [6].

Figure 1(Left) shows crack detection results on a part of the panel *Virgin Annunciate* which is particularly challenging because some painted features resemble cracks. The results displayed in Fig. 1(Right) show that BCTF falsely labels a significant amount of pixels (such as the decorative elements around the big letter "P"), while our method successfully differentiates true cracks from those painted features. It should be noted that both methods use the same image modalities.

Similar conclusions follow from the results on the panel *Singing Angels* displayed in Fig. 2(Top). In general, the proposed method detects more cracks while reducing false detections. Furthermore, the proposed method can be trained in an online fashion, i.e. without re-training the whole network, continuously improving detection results (Fig. 2(Bottom)). An important asset is also rapid processing, especially once the network is trained. This makes our framework integrable in a fast and interactive tool that can be used by art professionals.

Figure 3 depicts results on a small area from the panel of *John the Evangelist*. For this example, only one modality was considered, namely the color photograph, for both methods. The BCTF method was trained specifically for this image, while the proposed method was pre-trained on other panels and only a small fraction of labels (approximately 3,000 patches of two types) from the present panel were used to re-train the network. This experiment indicates that the proposed method can be deployed for a variety of paintings, with relative little effort.

In general, the proposed method demonstrates potential to improve upon the current state-of-the-art methods by detecting more cracks while also reducing false detections. Furthermore, the proposed method can be trained in an online fashion, i.e. without re-training the whole network, continuously improving detection results. An important asset is also rapid processing, especially once the network is trained. This makes our framework integrable in a fast and interactive tool that can be used by art professionals.

[2]Link: http://closertovaneyck.kikirpa.be/

Figure 2: **Top**: A fragment of the panel *Singing Angels*. **Bottom**: Results of the proposed method in comparison with BCTF. Blue: cracks identified by both methods; red: cracks detected by BCTF only; green: cracks detected by the proposed method only.

4 Conclusion

In this paper, we propose a novel crack detection framework capable of handling acquisitions from different modalities. In a first step, we apply morphological filtering for a coarse initial identification of crack pixels. This step substantially reduces the amount of data to be processed later on. We then train a CNN architecture with user labelled data to further refine the results obtained in the first step. We show that our method improves upon the current state-of-the-art in this application. An additional advantage is the possibility of re-training the network using newly available data. This feature allows to improve the already obtained result without significant time costs.

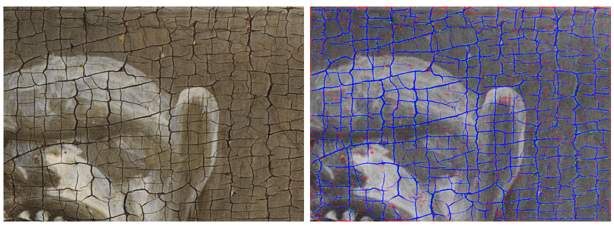

Figure 3: **Left**: A fragment of the panel *John Evangelist*. **Right**: Results of the proposed method (pre-trained on other panels and re-trained with few labels from the present panel) in comparison with BCTF (trained for this particular image). Blue: cracks identified by both methods, red: cracks detected by BCTF only, and green: cracks detected by the proposed method only.

References

[1] J. Mohen, M. Menu, and B. Mottin, "Mona lisa: Inside the painting," *Harry N. Abrams, New York, NY, USA,*, 2006.

[2] B. Cornelis, T. Ružić, E. Gezels, A. Dooms, A. Pižurica, L. Platiša, J. Cornelis, M. Martens, M. De Mey, and I. Daubechies, "Crack detection and inpainting for virtual restoration of paintings: The case of the ghent altarpiece," *Signal process*, vol. 93, no. 3, pp. 605–619, 2013.

[3] B. Cornelis, Y. Yang, Joshua T. Vogelstein, A. Dooms, I. Daubechies, and D. B. Dunson, "Bayesian crack detection in ultra high resolution multimodal images of paintings," *18th International Conference on Digital Signal Processing, IEEE*, 2013.

[4] A. Gupta, V. Khandelwal, A. Gupta, and M. C. S. Thammasat, "Image processing methods for the restoration of digitized paintings," *Thammasat Int. J. Sc. Tech*, vol. 13, no. 3, 2008.

[5] I. Giakoumis, N. Nikolaidis, and I Pitas, "Digital image processing techniques for the detection and removal of cracks in digitized paintings," *IEEE Trans. Image Processing*, vol. 15, no. 1, pp. 178–188, 2006.

[6] A. Pižurica, L. Platiša, T. Ružlić, B. Cornelis, A. Dooms, M. Martens, H. Dubois, B. Devolder, M. De Mey, and I. Daubechies, "Digital image processing of the Ghent altarpiece: supporting the painting's study and conservation treatment," *Signal Processing Magazine, IEEE* , vol. 32, 2015.

[7] A. Krizhevsky and I. Sutskever and G. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.

[8] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich, "Going deeper with convolutions," *CoRR*, vol. abs/1409.4842, 2014.

[9] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015.

[11] Z. Lei, Y. Fan, D.Z. Yimin, and J. Z. Ying, "Road crack detection using deep convolutional neural," *Image Processing (ICIP), 2016 IEEE International Conference*, 2016.

[12] Y.-Jin Cha, W. Choi, and O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Comput.-Aided Civ. Infrastruct. Eng.*, vol. 32, pp. 361–378, 2017.

[13] N. Otsu, "A threshold selection method from gray-level histogram," *IEEE Transaction on Systems, Man, and Cybernatics*, vol. SMC-9, no. 1, pp. 62–66, 1979.

[14] S. Mallat, "Understanding deep convolutional networks," *Phil. Trans. R. Soc. A*, vol. 374(2065):20150203, 2016.

[15] J. H. Jacobsen, E. Oyallon, S. Mallat, and A. W. M. Smeulders, "Multiscale hierarchical convolutional networks," *CoRR*, vol. abs/1703.04140, 2017.