Romanian Reports in Physics **72**, 805 (2020)

PHYSICAL DEGRADATION DETECTION ON ARTWORK SURFACE POLYCHROMIES USING DEEP LEARNING MODELS

L.M. ANGHELUȚĂ[1], A. CHIROȘCA[2]

[1] National Institute for Research and Development in Optoelectronics,
409 Atomiștilor, Măgurele, România
E-mail: *laurentiu@inoe.ro*
[2] University of Bucharest, Faculty of Physics, Department of Nuclear Physics,
Magurele, Romania
E-mail: *alecsandru.chirosca@gmail.com*

Received December 6, 2019

Abstract. This paper presents the application of a Deep Learning algorithm for a classified accurate detection of different types of physical damages in artwork polychrome surfaces. The algorithm was trained for the automated detection of three classes of typical surface deteriorations: cracks, blisters and detachments (losses). The image sets used in this study were previously recorded for the purpose of detailed surface 3D reconstruction by the means of macro-photogrammetry of a wood painting. These high-resolution images were captured using 2:1 optical macro magnification with a generous overlapping, following the 3D reconstruction methodology, and provided high quality details of the surface features to be classified. Specific activation maps are used to visually emphasize the detected potential deteriorated areas. The purpose of this work was on one hand to validate a process of reusing photogrammetry image data sets, used 3D reconstruction, for machine learning feature detection training and on the other hand to provide a starting point for the development of an affordable real-time surface damage assessment system.

Key words: cultural heritage, image analysis, deep learning, photogrammetry, physical damage monitoring.

1. INTRODUCTION

All artwork polychrome surfaces are prone to physical and visual deterioration due to a multitude of natural or anthropic factors. The risks factors that these surfaces are subjected to starts with the artist technique and used materials and are influenced by age and the preservation conditions. Preventive conservation measures [1] have the purpose to thwart the ageing and environment-based deteriorations of artworks by monitoring and controlling the storage and exhibition conditions (light, humidity, temperature, manipulation or pollutants). Another important task in the preventive conservation is the periodical imaging and physico-chemical documentation and investigations. These complementary methods are improving the preventive conservation policies with precise measurements and statistics that can foresee emerging deterioration effects.

1.1. POLYCHROME SURFACE DETERIORATION FACTORS

There are many issues of the artworks' painted layers that can be traced back to the artist's used materials and his method of mixing pigments and support treatment. The artwork chosen for this study presented different types of physical deteriorations on painted layer level. To test our proposed Deep Learning algorithm we focused on three common types of damage: flaking blisters, cracks and losses.

A common type of deterioration are the blisters (or bubbling) is a form of localized loss of adhesion of a paint film. With ageing this film becomes more rigid and the lack of adhesion is turned into flaking. This effect can appear between two painted layers or between a painted layer and the support. On wood support the most common cause for this type of damage is the moisture beneath the paint film, mostly resulted from inadequate protection of new woodwork before painting.

Perhaps the most common type of damage in painting layers are the cracks. These are separations in the paint layer, ground or support perpendicular to the surface of the painting. Cracks have many forms (fissures, crocodiling, crazing, checking, craquelure etc.) and are caused by different factors: ageing, physical impact, framing stress, expansion or contraction of the substrate or the application of hard-drying coatings over soft ones.

Combinations of cracks with paint or ground layer separations usually lead to flaking and material losses. These areas have missing material and can be caused by aforementioned combination of other types of damages or by physical impacts like abrasions, tears etc.

1.2. STATE OF THE ART IN POLYCHROME SURFACE CONSERVATION MONITORING

The usual practice for monitoring the physical quality decay of artwork painted surfaces is the visual inspection. There are special cases when the restoration processes are monitored with periodic 3D digitization [2]. Although it is not standardized, this approach is more and more used, especially using photogrammetry [3] due to its low cost and high accuracy. There are other important methods that were studied regarding the effect of environment vibrations on museum items [4] and experiments were performed using laser speckle-pattern interferometry to see the limits of crack emergence on painted wood [5] or record the artwork reactions to shock and vibrations during transportation [6].

Visual inspection of artwork physical quality can be subjected to human error by a multitude of factors. Even photographic documentation that will also be subjected to human inspection can become cumbersome due to the high amount of data to be analyzed. In this regard, automatic interpretation methods of digital photographs can greatly improve the speed of surface deterioration detection, classification and documentation and are can be incorporated into a real-time alerting system.

Computer vision is an interdisciplinary field that is focused on developing machine algorithms and methods that are based on human eye complexity with the purpose of improving the quality and speed of human visual abilities. A significant part of computer vision technologies are relying on developing performant artificial intelligence (A.I.) algorithms that are mainly used for image pattern recognition. When speaking of A.I., there are two approaches, both subfields of A.I.: machine learning and deep learning. Deep learning is a new field of AI that is based on a multi-layered artificial neural network. All modern AI based systems like Google Translate, Netflix, Alexa/Siri chat bots or self-driving cars would not exist without deep learning.

On a lower scale, deep learning algorithms have found their way into similar image-based inspection and monitoring applications in both civil engineering and cultural heritage. The most common deep neural network class used in image interpretation is the convolutional neural network (CNN). These convolutional networks are inspired from biology because the neuron connection pattern resembles that of the animal visual cortex [7].

There have been reported many advantages of using CNN algorithms in civil engineering for cement and structure (*e.g.* bridges, buildings) inspection that in combination with UAV technology present several advantages. The common subjects for these kind of inspections were cracks and detachments. The advantages can be summed up as: low-cost inspection, high resolution imaging, minimal human intervention, predictive maintenance [8–11]. Another type of applications is where the CNN algorithms are used for automated building or building component detection on different imaging sources for classification [12, 13]. Similar application but in a totally different field is the anomaly detection in nuclear reactors [14].

Similar works with our approach have been reported only for cracks and craquelure (a unique network of cracks) detection using CNN or BCTF models [15, 16].

Our approach is based upon the usage of encoder pattern using a VGG-16 specialized convolutional network developed at the Visual Geometry Group from Oxford University. This network comes pre-trained for classic applications in computer vision object detection and improve the training for our specific applications [17, 18].

2. MATERIALS AND METHODS

Specialized convolutional networks are widely used in Computer Vision for image identification and segmentation [19] while for our needs we have to adjust the classification part within the last 8 layers due rendering the first 8 encoding layers to non-trainable thus preserving the original weights provided by the model. The network output consists of a dense layer providing an output of 3 (crack, clean and exfoliation). The resulting model consists of 14.716.227 parameters from which only 12.980.739 parameters are trainable.

Our model was developed using Keras [20] with Tensorflow [21] using a desktop computer with 32 GB of RAM and Intel Core i7 CPU with NVIDIA GPU. Although the usage of GPU for running large scale machine learning shows considerable improvements, our approach developed and implemented the full model using only CPU, thus ensuring compatibility with most of the environments.

The image set is part of a photogrammetric documentation of a XIXth century icon painted on wood, representing St. Constantine and Helen [22]. The 3D digitization work was realized using a 36 megapixel full frame camera with a 60 mm macro lens with extension rings for a total optical magnification of 2:1. For this reason, the resulted orthophoto mosaics of two 6 × 4 cm areas had an impressive resolution of 26.345 × 15.486 pixels. These high resolution allowed a good visibility of all the physical deterioration details (cracks, blisters, losses) which made these large images the perfect candidates for this approach. These orthophoto mosaics were segmented in much smaller sections that were to be fed in the processing model.

Our approach is using a standard 16-layer VGG-16 model with the last 8 trainable layers and softmax activation for the classification layer raising a flag for the classification. Using the standard VGG-16 encoder means that the input size is fixed at 224 × 224 for each of the RBG channels thus the input size will be (3, 224, 224).

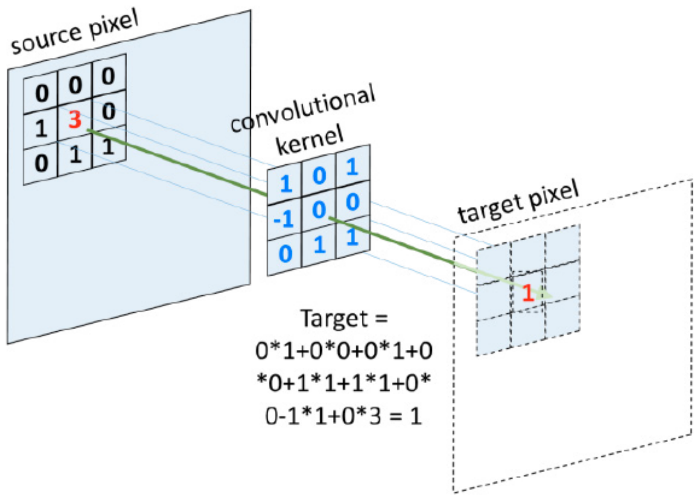

Fig. 1 – Convolutional layer morphology [18].

A convolutional layer performs a mathematical convolution operation described in eq. 1 with the input data (*x*) and a convolution filter – also known as kernel – (*w*). The operation is performed within a window that is sliding across the input resulting a feature map that is integrated in the end to produce the layer output.

$$y = x * w \Rightarrow y[i] = \int_{-\infty}^{+\infty} x[i-j] * w[j] \mathrm{d}j . \tag{1}$$

Softmax layers uses the model with the same name where a probability distribution is assigned to the output layer and it's fitted to the best results providing the classification layer.

The pooling layers allow for layer composition according to a kernel while keeping the information contained within the level to the largest amount possible.

The obtained model uses on input layer with a shape of (3,224,224) followed by two convolutional layers and a pooling layer followed by the same segment again. The classifier part of the model consists of three convolutional layers with a pooling layer repeated three times before the averaging level and the output layer. The model was compiled with categorical cross entropy loss function and the stochastic gradient descent (SGD) optimizer that improves accuracy.

While using high resolution images with a VGG network we have to adapt the input image to the network requirements. This process consists of image numerical segmentation with a moving window while the outer margins are usually discarded to multiple of window size (the input image has to be decomposed into its parts with the image size). Using this technique usually takes resources and computing time leading to less than real-time results.

Numerical segmentation consists of subsampling images from an array of pixels for each of the RGB layers. The image is transformed into a set of three arrays with pixel values that are traversed using a moving window without superimposing it over the last considered one.

The complex morphology proposed for the convolutional network reduces the input size using 5 pooling layers and leads to a high accuracy for our training and testing dataset.

The model was fitted to a dataset of 2.148 pictures scaled according to the model input size and leading to an 80% average accuracy for a training session consisting of 20 epochs and a batch size of 32 covering the entire training dataset. The initial dataset was divided into 80% training dataset and 20% testing dataset. The confusion matrix for our setup is presented into the Fig. 3.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 224, 224, 3)	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
global_average_pooling2d_1 (	(None, 512)	0
dense_1 (Dense)	(None, 3)	1539

Total params: 14,716,227
Trainable params: 12,980,739
Non-trainable params: 1,735,488

Fig. 2 – Representation of the implemented model.

3. RESULTS AND DISCUSSIONS

Model training results are presented in Fig. 3 for accuracy and loss. Although this model provides a decent estimator for degradation it requires a high amount of hardware and software resources just to be employed. Retraining the model requires even more resource.

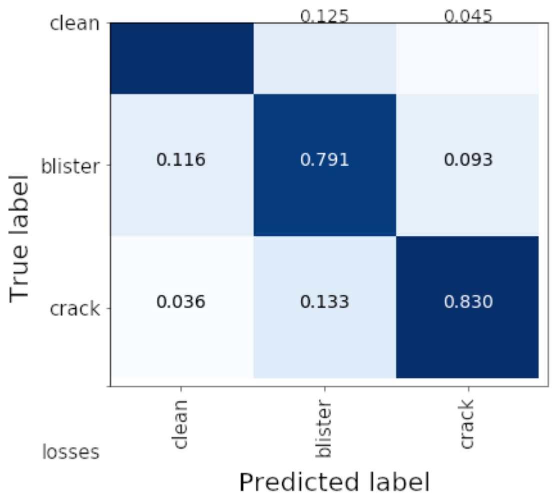

Fig. 3 – Confusion Matrix for the testing dataset.

The accuracy plot presented in Fig. 4 shows an accuracy degradation in the 12.5–15 epochs (that is also present in the loss plot) where specific features for the training set are considerably different with the rest of the image thus showing the specific issue with cultural heritage objects where a large palette of features are harder to predict over the feature rich in diversity portions of them.

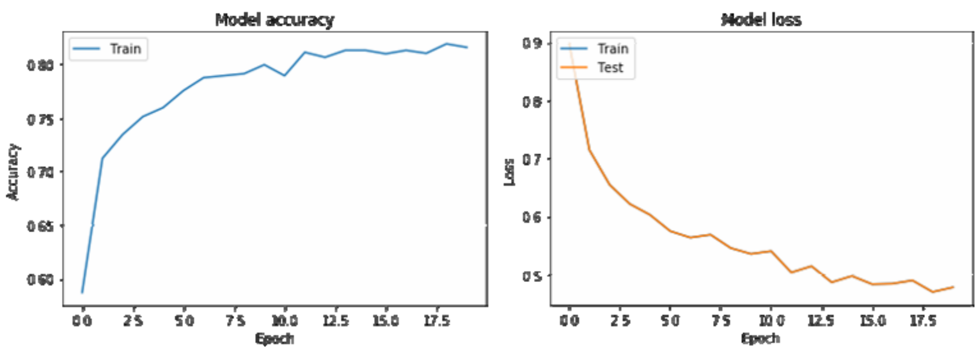

Fig. 4 – Model accuracy (on the left) *vs* the model loss for the testing dataset.

Table 1

Training results model

	Precision	Recall	Fl-score	Support
0	0.94	0.83	0.88	287
1	0.54	0.79	0.64	66
2	0.87	0.83	0.85	165
Accuracy			0.82	538
Macro avg.	0.78	0.82	0.79	538
Weighted avg.	0.85	0.82	0.83	538

Sample activation matrix are presented in the figure below for specific clean and degradation cases.

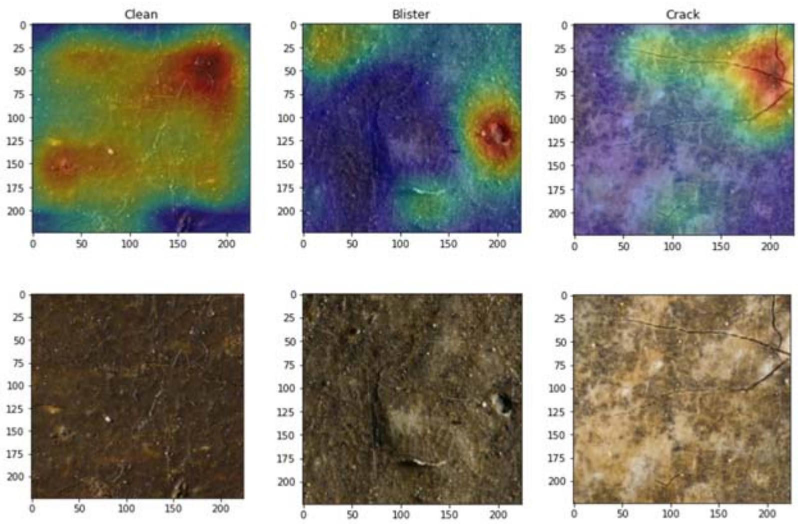

Fig. 5 – Activation map for several degradation image segments (upper row) *vs* the same segments without the activation map (lower row).

In order to be able to improve accuracy, hardware and software dependencies and the applicability of the model it needs to be changed to an encoder-decoder patters that usually implements a FCN (Fully Convolutional Network) design with interconnected layers allowing for certain attributes to be better passed on between the encoder layer and the decoder layer.

4. CONCLUSION

Deep Learning techniques prove to be a valuable method for knowledge management, such algorithms allow for training specific models to correctly assess degradation classes and save valuable time during the estimation of such degradations in standard image sources (taken with phones or other portable devices). Our model uses a specific approach for cultural heritage item assessment for elements with an increased visual complexity that can be quite difficult to assess using other analytical methods.

Our chosen model with pre-trained input layers allowed us to use a small dataset (774 blister images, 464 images with cracks, 46 images with losses and 1402 clean images) with respect to other models found in literature.

During the model development specific resource management techniques had to be applied (aggressive garbage collectors, increased memory and CPU resources and even using the GPU to perform all the steps necessary for model development). Using M.L. technics, once such a model is developed and the correct neural network coefficients are established, the model can be run on standard or even low resource devices as it is not so resource intensive as the training side.

Our model uses standard techniques based upon the Tensorflow® engine allowing for rapid spread of such models for standard degradation assessment techniques.

Future developments will follow with emphasis on model development and accuracy improvements while reducing the overall model complexity (the complexity comes mostly from the VGG16-VGG19 base model [23]). Once the complexity is reduced as much as possible, the model can be implemented in computer vision devices allowing for augmented reality applications to be made available to most of the off-the-shelf devices.

Acknowledgements. This study is financed by the Ministry of Research and Innovation through “Program Nucleu” PN 19-18.01.01: OPTRONICA VI.

REFERENCES

1. ICOM-CC, *Terminology to characterize the conservation of tangible cultural heritage*, 15th Triennial Conference, New Delhi, 22–26 September 2008.
2. Robson, S., Bucklow, S., Woodhouse, N., Papadaki, H., *Periodic photogrammetric monitoring and surface reconstruction of a historical wood panel painting for restoration purposes*, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci –**XXXV** Part B5, 395–400 (2004).
3. Abate, D., *Documentation of paintings restoration through photogrammetry and change detection algorithms*, Herit. Sci. **7**, 13 (2019).
4. Thickett D., *Vibration damage levels for museum objects*. ICOM Committee for Conservation: Preprints of the 13th Triennial Meeting, ed., Vontobel R., London, James and James **1**, 90–95, 2002.
5. Łukomski, Michał, *Painted wood. What makes the paint crack?,* Journal of Cultural Heritage **13**, S90–S93 (2012).
6. Tsiranidou, E., Bernikola, E., Vivi, T., Fankhauser, T., Läuchli, M., Palmbach, C., Bäschlin, N., *Holographic monitoring of transportation effects on canvas paintings*, SPIE Newsroom, 2011.
7. Fukushima, K., *Neocognitron*. Scholarpedia **2** (1), 1717 (2007).
8. Silva, W.R.L., Lucena, D.S., *Concrete Cracks Detection Based on Deep Learning Image Classification*, Proceedings **2**, 489 (2018).

9. Shihavuddin, A., Chen, X., Fedorov, V., Nymark Christensen, A., Andre Brogaard Riis, N., Branner, K., Bjorholm Dahl, A., Reinhold Paulsen, R., *Wind Turbine Surface Damage Detection by Deep Learning Aided Drone Inspection Analysis*, Energies **12**, 676 (2019).
10. Borin, P., Cavazzini, F., *Condition assessment of RC bridges. Integrating machine learning, photogrammetry and BIM*, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. **XLII-2**/W15, 201–208 (2019).
11. Young-Jin, C., Wooram, C., *Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks*, Computer-Aided Civil and Infrastructure Engineering, 1–18, 2017.
12. Pezzica, C., Schroeter, J., Prizeman, O. E., Jones, C. B., and Rosin, P. L., *Between images and built form: automating the recognition of standardised building components using deep learning*, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. **IV-2**/W6, 123–132 (2019).
13. Condorelli, F., Rinaudo, F., Salvadore, F., and Tagliaventi, S., *Architectural heritage recognition in historical film footage using neural networks*, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. **XLII-2**/W15, 343–350 (2019).
14. F. Calivá et al., *A Deep Learning Approach to Anomaly Detection in Nuclear Reactors*, International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, 2018, pp. 1–8.
15. R. Sizyakin, B. Cornelis, L. Meeu, M. Martens, V. Voronin and A. Pizurica, 2018, *A deep learning approach to crack detection in panel paintings*, in IP4AI, Ghent, Belgium, 2018, pp. 40–42.
16. Sidorov, A., Hardeberg, J. Y., *Craquelure as a Graph: Application of Image Processing and Graph Neural Networks to the Description of Fracture Patterns*, The IEEE International Conference on Computer Vision (ICCV), 2019.
17. Y. Fu, C. Aldrich, *Flotation froth image recognition with convolutional neural networks*, Minerals Engineering **132**, pp. 183–190, 2019.
18. Y. Fu, C. Aldrich, *Using Convolutional Neural Networks to Develop State-of-the-Art Flotation Froth Image Sensors*, IFAC-PapersOnLine **51** (21), 152–157 (2018).
19. Long, J., Shelhamer, E., Darrell, T., *Fully convolutional networks for semantic Segmentation*, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, pp. 3431–3440.
20. Chollet, F. et al., Keras. Retrieved from https://github.com/fchollet/keras., 2015.
21. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Zheng, X. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*. Retrieved from http://arxiv.org/abs/1603.04467, 2016.
22. Anghelută, L. M., Rădvan, R, *Macro photogrammetry for the damage assessment of artwork painted surfaces*, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. **XLII-2**/W15, 101–107 (2019).
23. Ciresan, Dan C. and Meier, Ueli and Masci, Jonathan and Gambardella, Luca M. and Schmidhuber, Jürgen, *Flexible, High Performance Convolutional Neural Networks for Image Classification*, Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, Vol. 2, IJCAI'11, 2011, pp. 1237–1242.