International Symposium on the Conservation of Monuments in the Mediterranean Basin

(2024)

Proceedings of the 11th MONUBASIN (2024)

Simulation-aided infrared thermography with a new efficient channel attention mechanism aided Faster R-CNN model and decomposition-based noise reduction for detecting defects in ancient polyptychs

Xin Wang, Guimin Jiang, Miranda Mostacci, Stefano Sfarra, Hai Zhang

doi: 10.12681/monubasin.8345

Simulation-aided infrared thermography with a new efficient channel attention mechanism aided Faster R-CNN model and decomposition-based noise reduction for detecting defects in ancient polyptychs

Xin Wang, *School of Automation and Electrical Engineering*, Shenyang Ligong University, 110159, Shenyang, China
wangxin19990828@163.com

Guimin Jiang, *School of Automation and Electrical Engineering*, Shenyang Ligong University, 110159, Shenyang, China & *Centre for Composite Materials and Structures (CCMS*), Harbin Institute of Technology, 150001, Harbin, China
2502818383@qq.com

Miranda Mostacci, *Professional Restorer*, Via Muranuove 64, 67043 Celano, Italy
miranda.mostacci.90@gmail.com

Stefano Sfarra, *Department of Industrial and Information Engineering and Economics*, University of L'Aquila, 67100, L'Aquila, Italy
stefano.sfarra@univaq.it

Hai Zhang, *Centre for Composite Materials and Structures (CCMS)*, Harbin Institute of Technology, 150001, Harbin, China
hai.zhang@hit.edu.cn

Abstract. In this study, we investigate how to automatically and efficiently detect defects in ancient polyptychs by infrared thermography (IRT), combined with numerical simulation, deep learning networks and machine learning algorithms. Through an innovative improved Faster-RCNN model and LRTDTV denoising method, the recognition of surface and internal defects of ancient artworks is effectively improved. This improved Faster-RCNN model introduces an effective channel attention (ECA) mechanism in the feature extraction stage, which significantly improves the performance of the model in recognizing small defects, and a comparison with the original Faster-RCNN model reveals that the average detection accuracy_50 (AP_50) of the improved model is significantly improved to 86.9%. The average precision_small (AP_s) especially improved to 59.1% when detecting small-size defects. The experimental results verify the practicality and efficiency of the method in cultural heritage protection, which helps to maximize the protection and transmission of cultural heritage. In addition, the method in this study can achieve fast and accurate detection of defects in any type of cultural heritage object while avoiding secondary damage to the samples, providing effective technical support for cultural heritage protection.

Keywords: Numerical simulation, Machine learning algorithms, Faster R-CNN network, Attention mechanism, Defect detection, Deep learning.

1 Introduction

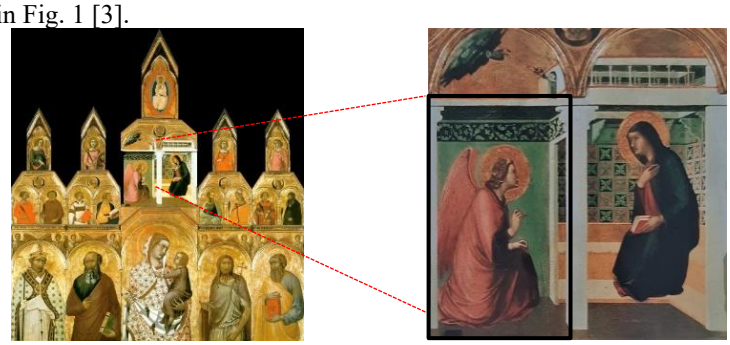
Cultural heritage, critical for civilization's progress, includes polyptychs, valued for their historical and artistic significance. These artworks often depict important cultural narratives but are prone to damage like cracks or splits over time. Timely and accurate detection of these defects is essential to preserve their integrity and continue their cultural transmission.

Non-destructive testing (NDT) [1] is essential in various industries for ensuring material safety and integrity without damage. NDT's especially crucial in cultural heritage to protect irreplaceable artifacts during restoration. Currently, the main NDT methods include infrared (IR) photography, ultraviolet imaging, X-ray photography, acoustic emission, and terahertz.

Infrared thermography (IRT) has become a key tool in the NDT of materials due to its non-invasive, immediate and superior imaging. Particularly for artwork detection [2], IRT is popular for assessing artefact defects due to its ability to scan large areas quickly, with high resolution and without contact. IRT operates in two phases: the first phase involves recording temperature changes on the surface, and the second phase identifies defects by differential analysis of thermophysical properties, resulting in a clear infrared contrast. However, if the heat source is not calibrated in the first stage, there is a risk that the IRT of the artwork may fade. Furthermore, the second stage detects the artwork by analyzing the reflected near-infrared and short-wave IR spectra, but this method is difficult to deal with deep internal defects and requires a visual inspection, leading to possible errors in defect assessment.

To prevent thermal discoloration in IRT used on cultural heritage, two strategies are effective: reducing input energy and employing numerical modelling to optimize and standardize testing procedures. This study focuses on numerical simulation, which enhances IRT system design by predicting outcomes and understanding heat mechanisms in complex materials. In order to avoid possible errors in the second-stage defect assessment, this study constructs an automatic defect detection system based on numerical simulation using deep learning networks and machine learning algorithms to reduce possible errors in actual defect detection and optimize the defect detection process.

This study focuses on a reproduction of a 1320 polyptych by Pietro Lorenzetti, preserved in Arezzo, Italy, as shown in Fig. 1 [3].



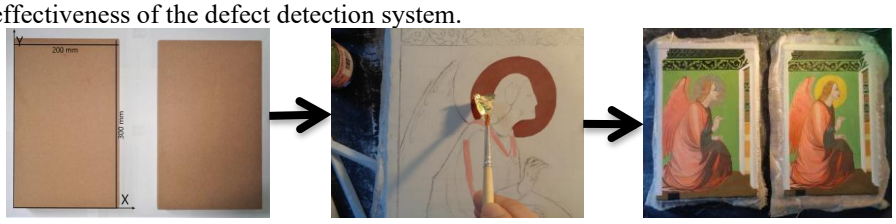
(a) (b)

Fig. 1. (a) A photograph of the polyptych, (b) a zoomed view on the reproduced part.

2 Description of the samples under test and numerical simulation setup

2.1 Description of the tested sample

In order to validate a numerical simulation-assisted approach to detecting defects in polyptychs via IRT, two simulated polyptychs were created on wood panels using a 14th-century pen-and-pencil technique. These boards were simulated with rabbit skin glue and Teflon inserts to simulate defects, then coated with multiple layers of *gesso di Bologna* and rabbit skin glue, and finally layered with tempera paint to complete the final artwork, as shown in Fig. 2. For more details on the production of the replica, see the author's article [3]. This setup was intended to approximate the replication of historical, artistic methods to test the effectiveness of the defect detection system.



(a) (b) (c)

Fig.2. Brief description of painting samples: (a) the boards are used as support, (b) using a damp cotton ball to adhere the gold leaf, and (c) the fabrication of the samples is completed.

2.2 Geometric modeling instructions

This section describes the process of constructing a geometric model of the sample under test in order to numerically simulate the temperature distribution on the sample surface. In this study, the surface of the geometric model was constructed by first drawing the surface contours (*X* and *Y* axes) through *CAD* software and then stretching the depth of the work plane where the contours were completed (along the *Z* axis). However, the remaining layers of the model were constructed using the linear tetrahedral units of the *COMSOL Multiphysics* software. Finally, IRT experiments were simulated using the Solid Heat Transfer module of *Multiphysics 6.0.* For more details on the modelling, see the authors' paper [3].

3 Methodology

The aim of this research is to enable faster and more accurate automatic defect detection in IRT using numerical simulation, deep learning networks and machine learning algorithms. This section introduces the content of the method to achieve faster and more accurate automatic defect detection.

3.1 The total-variation regularized low-rank tensor decomposition denoising method (LRTDTV)

Here, the thermographic image restoration method uses LRTDTV model to reduce noise. The approach leverages Tucker decomposition and total variation regularization, focusing on the spatial and spectral smoothness. Given the non-convex nature of the problem, the augmented Lagrange multiplier (ALM) method is used for optimization.

Define a third-order tensor $\mathcal{Y}: = \{Y^1, Y^2, Y^3, \dots, Y^B\}$, where $Y^i \in R^{H\times W}$ $(i = 1,2,3, \dots, B)$ represents the i^{th} frame of a thermographic sequence, with B being the number of frames, and H and W being the height and width of the image, respectively. Our data can be considered a mixture of a noiseless image and two types of noise, represented as:

$$\mathcal{Y} = \mathcal{X} + \mathcal{N} + \mathcal{S} \tag{1}$$

where $\mathcal{X}$ is the noiseless image of our data, $\mathcal{N}$ is Gaussian noise, and $\mathcal{S}$ is sparse noise. For more details, check out the authors' article [3].

To eliminate noise in thermographic images, LRTDTV model is used; the visualization of the decomposition can be found in Fig. 3. The objective function is:

$$\begin{aligned} &\min_{X,N,S} \tau\|X\|_{\text{SSTV}} + \lambda\|S\|_1 + \beta\|N\|_F^2 \\ &\text{s.t.} \quad \mathcal{Y} = \mathcal{X} + \mathcal{N} + \mathcal{S} \\ &\mathcal{X} = \mathcal{C} \times_1 U_1 \times_2 U_2 \times_3 U_3 \\ &U_i^T U_i = I \quad (i = 1,2,3) \end{aligned} \tag{2}$$

where τ, $\lambda$and β are regularization parameters. The $\mathcal{C} \times_1 U_1 \times_2 U_2 \times_3 U_3$ represents Tucker decomposition, and $\|X\|_{\text{SSTV}}$ is the anisotropic Frobenius norm, exploiting the spatial-spectral continuity of thermographic images:

$$\|X\|_{\text{SSTV}} = \sum_{i,j,k} \omega_1 \left|x_{i,j,k} - x_{i,j,k-1}\right| + \omega_2 \left|x_{i,j,k}\right| + -x_{i,j-1,k}\omega_3 \left|x_{i,j,k} - x_{i-1,j,k}\right| \tag{3}$$

where $x_{i,j,k}$ is the $(i,j,k)^{\text{th}}$ entry of $\mathcal{X}$, ω_j $(j = 1,2,3)$ are the weights controlling regularization strength, and *k* represents the dimension of the thermographic data.

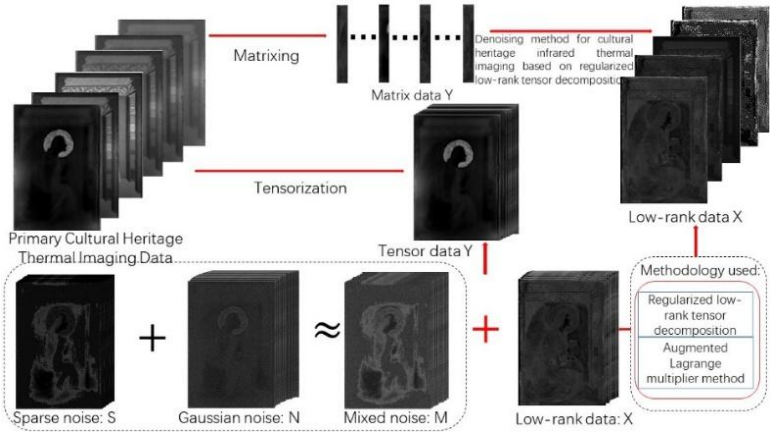


Fig. 3. Schematic diagram of the LRTDTV denoising method.

To solve Problem (2), one can introduce auxiliary variables to reformulate it into a simpler minimization problem:

$$\begin{aligned} &\min_{C,U_i,X,\mathcal{F},S,N} \tau\|\mathcal{F}\|_1 + \lambda\|\mathrm{S}\|_1 + \beta\|N\|_{\mathrm{F}}^2 \\ &s.t.\, y = X + S + N, X = Z, D_\omega(Z) = \mathcal{F}, \\ &X = C \times_1 U_1 \times_2 U_2 \times_3 U_3, U_i^T U_i = I \end{aligned} \tag{4}$$

where $D_\omega(\cdot) = [\omega_1 \times D_h(\cdot)\,; \omega_2 \times D_v(\cdot)\,; \omega_3 \times D_t(\cdot)]$ is the so-called weighted three-dimensional difference operator, and D_h, D_v, D_t are the first-order difference operators respect to three different directions. Since this is a non-convex optimization problem, the ALM method is used for optimization. Based on the ALM method, the problem can be transformed into minimizing the following augmented Lagrangian function:

$$\begin{aligned} &\mathrm{L}(X,S,N,Z,\mathcal{F},\Gamma_1,\Gamma_2,\Gamma_3) = \tau\|\mathcal{F}\|_1 + \lambda\|S\|_1 + \beta\|N\|_{\mathrm{F}}^2 \\ &\langle\Gamma_1, \mathrm{y} - X - S - N\rangle + \langle\Gamma_2, X - Z\rangle + \langle\Gamma_3, D_\omega(Z) - \mathcal{F}\rangle + \frac{\mu}{2}(\|y - X - N\|_{\mathrm{F}}^2 \\ &+ \|X - Z\|_{\mathrm{F}}^2 + \|D_\omega(Z) - \mathcal{F}\|_{\mathrm{F}}^2) \end{aligned} \tag{5}$$

where μ is the penalty parameter, and Γ_i $(i = 1,2,3)$ are the Lagrange multipliers. Enhancing the Lagrangian function requires optimization iterations, which requires iterative updating of each variable. For a detailed iterative procedure for the parameters U_i, X, Z, $\mathcal{F}$, S, N, and Γ_i, see the authors' separate manuscript [3].

These steps are repeated iteratively until convergence, using an adaptive approach for the penalty parameter μ. After completing this process, the noise-free image can be effectively separated from the noise components.

3.2 Improved Faster R-CNN internet

Ross B. Girshick introduced the Faster R-CNN in 2016 [4]. The model starts with normalizing data to handle various inputs and uses networks like VGG and ResNet for feature extraction. It features a region proposal network (RPN) that identifies potential defects, which are refined through non-maximum suppression (NMS) to highlight crucial areas. The selected regions undergo Region of Interest Pooling (RoI Pooling) and are processed by a fully connected layer for accurate defect classification and localization. The Faster R-CNN includes three main components: feature extraction, region proposal, and detection networks.

This study enhances the Faster R-CNN by incorporating the efficient channel attention (ECA) mechanism, which focuses on local interactions within each channel using a one-dimensional convolutional (C1D) layer instead of a fully connected one. This modification not only reduces the model's parameters but also boosts its accuracy and efficiency in localizing defects in complex thermal images.

ECA Mechanism Steps:

1. Adaptive Kernel Size Selection:

The kernel size K is dynamically determined based on the number of channels C:

$$C = \phi(K) = \gamma * K - b \tag{6}$$

Typically, channels C are powers of 2, so the kernel size K is set as:

$$K = \varphi(C) = \left|\frac{\log_2 C}{\gamma} + b\right|_{odd} \tag{7}$$

Here, $|\cdot|_{odd}$ rounds to the nearest odd number.

2. Generating Channel Attention Weights:

Apply Global Average Pooling (GAP) to input feature maps to get a $1 \times 1 \times C$ vector. Use a C1D on this vector to compute α_i:

$$\alpha_i = \sigma\big(C1D_k(y)\big) \tag{8}$$

The equation can be expressed as:

$$\alpha_i = \sigma\big(\textstyle\sum_{j=1}^{k} \omega_i^j\, y_i^j\big),\ y_i^j \epsilon \Omega_i^k \tag{9}$$

Here, ω_i^j are learning parameters and $\sigma(\cdot)$ is the sigmoid activation function.

3. Applying Channel Attention Weights:

The weights α_i are applied to the input feature map X:

$$X' = X \otimes \alpha_i \tag{10}$$

X' is the resulting feature map, and $\otimes$ denotes channel-wise multiplication.

This approach efficiently recalibrates features for improved defect detection. For an improved computational visualization of ECA for the Faster-RCNN model see Fig. 4.

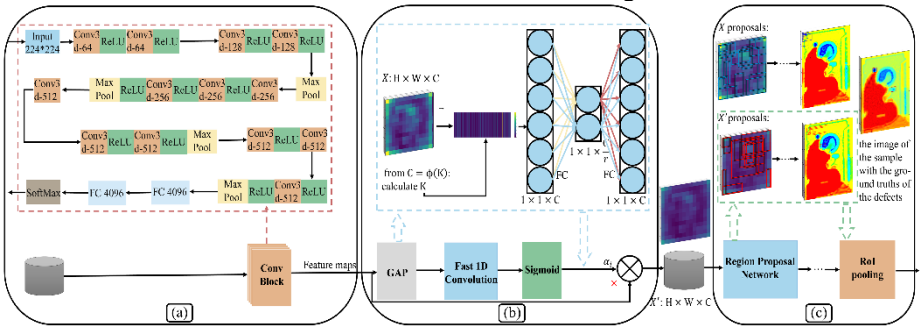


Fig. 4. (a) Original image input forward VGG16 network structure, (b) ECA module structure diagram, visualizing input feature map X and re-weighted output feature map X', (c) comparing the visualizations of X and X'proposals.

4 Experimental results and analysis

The aim of this study is to improve the speed and accuracy of automated defect detection in IRT experiments through numerical simulation, deep learning and machine learning methods. The study includes the production of IRT experimental datasets and numerical simulation datasets, performing model training and analyzing the results. First, we collected 28 thermal images of polyptychs using a 640x512 pixel IR camera. To increase the limited number of images, we extended the dataset to 812 images by numerical simulation, including 28 camera-collected images and 784 COMSOL-simulated images. 90% of the images were used for training and 10% for testing. Training was performed using a Faster-RCNN network, optimized by stochastic gradient descent (SGD) and VGG backbone.

After the completion of the network training experiment, this study also conducted the original Faster R-CNN network training experiment and simultaneously made a clear comparison between the original Faster R-CNN model and the improved Faster R-CNN model. In order to visually evaluate the advantages and disadvantages of these two models, the evaluation results of the two models are detailed in Table 1.

Table 1. Comparison of model testing performance metrics

model	AP_{50} *[%]*	AP_{75} *[%]*	AP_S *[%]*	AP_L *[%]*	AR_1 *[%]*	AR_{10} *[%]*	AR_{100} *[%]*	Detection time *[s]*
Faster R-CNN	**83.5**	**71.3**	**51.3**	**67.3**	**50.1**	**56.1**	**56.1**	**0.23**
Improved model	**86.9**	**74.1**	**59.1**	**69.3**	**57.8**	**62.4**	**62.4**	**0.22**

In order to have a more intuitive feeling of the enhancement of machine learning algorithms and deep learning networks for the detection of defects in infrared thermal imaging, we show four types of experimental results of infrared thermal imaging (see Fig. 5): the images in each row are, in order, the infrared thermal image of experimental sample A, the infrared thermal image of sample A after LRTDTV de-noise and Fourier transform, the infrared thermal image of experimental sample B, and the infrared thermal image of sample B after LRTDTV de-noise and Fourier transform.

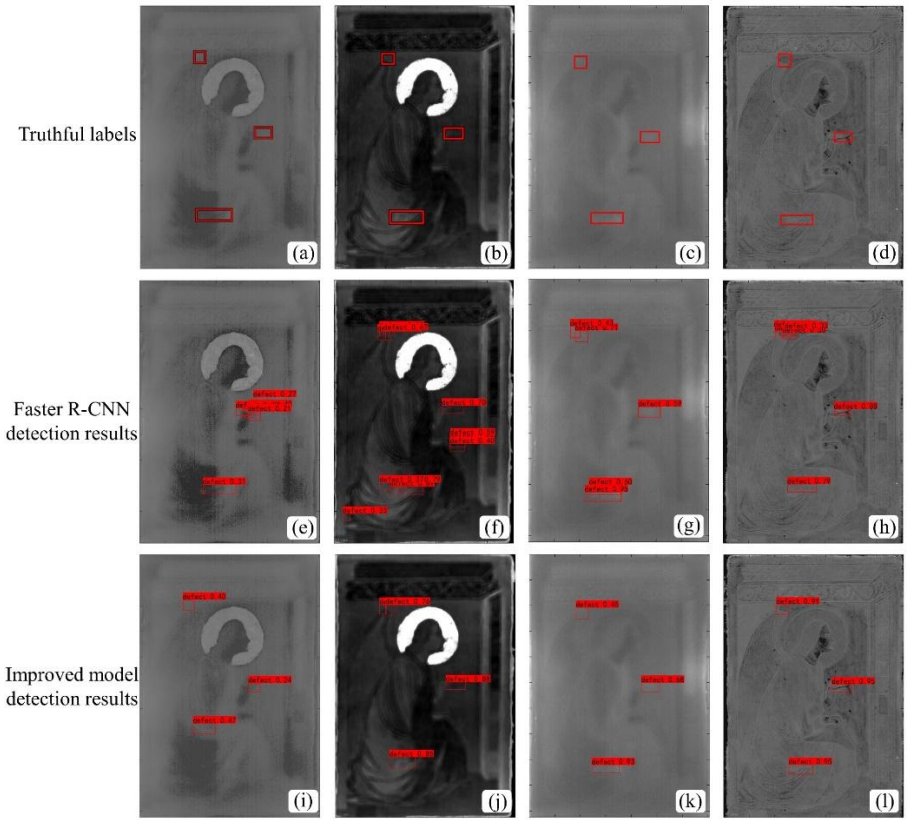


Fig. 5. IRT experimental results with real labels: (a) the raw sample A, (b) the sample A after LRTDTV de-noise and Fourier transform, (c) the raw sample B (d) the sample B after LRTDTV de-noise and Fourier transform. The defect detection results of IRT: (e) the raw sample A after applying the Faster R-CNN model, (f) the sample A after applying LRTDTV de-noise, Fourier transform, and the Faster R-CNN model, (g) the raw sample B after applying the Faster R-CNN model, (h) the sample B after applying LRTDTV de-noise, Fourier transform, and the Faster R-CNN model, (i) the raw sample A after applying the Improved Faster R-CNN model, (j) the raw sample A after applying LRTDTV de-noise, Fourier transform, and the improved Faster R-CNN model, (k) the raw sample B after applying the improved Faster R-CNN model and (l) the sample B after applying LRTDTV de-noise, Fourier transform, and the improved Faster R-CNN model.

By observing the results of this experiment, we can find the following: comparing Fig. 5(b), Fig. 5(f) and Fig. 5(j), it can be seen that the automatic defect detection network based on the improved Faster R-CNN can better detect the tiny defects that are difficult to be detected by human eyes. Comparing Fig. 5(c), Fig. 5(g), and Fig. 5(k) with Fig. 5(d), Fig. 5(h), and Fig. 5(l), it is shown that the LRTDTV de-noising method can effectively remove the noise interference in infrared thermal imaging.

5 Conclusions

In this study, IRT was used to perform non-invasive inspection of the ancient polyptychs and care was taken to avoid secondary damage. To protect the ancient polyptychs from secondary damage, we use numerical simulations to match the experimental surface temperatures and an LRTDTV decomposition model to efficiently remove Gaussian noise from the thermal images, thus facilitating the detection of defects in the samples by IRT. In addition, an improved Faster-RCNN model is proposed, which utilizes VGG16 for feature extraction and incorporates an ECA mechanism in the feature pyramid network (FPN), which greatly improves the defect detection capability of ancient artworks. After comparative experiments, the results of the study proved the efficiency and practicality of the model, which can provide strong technical support for cultural heritage protection. It greatly improves the accuracy and rapidity of IRT for cultural heritage protection and defect detection.

Acknowledgements
This work was supported by the National Key R&D Program of China (Grant No. 2023YFE0197800) and the Italian Ministry of University and Research (Grant No. PGR02016).

References

1. Hu, J., Zhang, H., Sfarra, S., Gargiulo, G., Avdelidis, N. P., Zhang, M., et al.: Non-destructive imaging of marqueteries based on a new infrared-terahertz fusion technique. Infrared Physics & Technology 125, 104277(2022).
2. Mix, P. E.: Introduction to nondestructive testing: a training guide. 2nd ed. John Wiley, Sons (2005).
3. Jiang, G., Wang, X., Hu, J., Wang, Y., Li, X., Yang, D., et al.: Simulation-aided infrared thermography with decomposition-based noise reduction for detecting defects in ancient polyptychs. Heritage Science 11, 223(2023).
4. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. In: IEEE transactions on pattern analysis and machine intelligence 39(6), 1137-1149(2016).