Simulation-Aided Infrared Thermography with Faster R-CNN-ECA Model and LRTDTV Denoising Method: the Case-Study of Ancient Polyptychs

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

Simulation-aided infrared thermography with Faster R-CNN-ECA model and LRTDTV denoising method: The case-study of ancient polyptychs

Wang X.[1], Jiang G.[1,2][†,1], Mostacci M.[3], Sfarra S.[4], Zhang H.[2][†,2]

[1] School of Automation and Electrical Engineering,
Shenyang Ligong University,
110159, Shenyang, China
wangxin19990828@163.com
[†,1]2502818383@qq.com
[2] Centre for Composite Materials and Structures (CCMS),
Harbin Institute of Technology,
150001, Harbin, China
[†,2] hai.zhang@hit.edu.cn
[3] Professional Restorer, Via Muranuove 64, 67043 Celano, Italy
miranda.mostacci.90@gmail.com
[4] Department of Industrial and Information Engineering and Economics,
University of L'Aquila,
67100, L'Aquila, Italy
stefano.sfarra@univaq.it

Abstract. In this study, we investigate how to automatically and efficiently detect defects in ancient polyptychs by infrared thermography, combined with numerical simulation, deep learning networks and machine learning algorithms. Through an innovative improved Faster R-CNN model and LRTDTV denoising method, the recognition of surface and internal defects of ancient artworks is effectively improved. This enhanced Faster R-CNN model incorporates an effective channel attention mechanism in the feature extraction stage, significantly boosting the model's performance in recognizing small defects. Comparisons with the original Faster R-CNN model show that the average precision at an intersection over union of 0.5 has increased to 87.3% for the improved model. Notably, the average precision for detecting small defects has risen to 54.8%. The experimental results verify the practicality and efficiency of the method in cultural heritage conservation, which helps to maximize the conservation 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 objects while avoiding secondary damage to the samples, providing effective technical support for cultural heritage conservation.

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

1   Introduction

Cultural heritage, because of its unique historical and cultural value, has become an important force for social progress and civilization. The preservation of the cultures of diverse ethnic groups and regions is fundamental to the harmonious coexistence of humanity, and safeguarding cultural heritage is vital for maintaining cultural diversity. As a form of cultural heritage, polyptychs hold significant historical and artistic value. However, over time, they inevitably develop defects such as cracks and holes. Therefore, it is particularly important to detect and repair these defects in a timely manner.

Non-destructive testing (NDT) [1] techniques have received widespread attention due to their non-destructive nature, safety and adaptability to structures. Currently, the main methods include infrared photography, digital photography, ultraviolet imaging, X-ray and acoustic emission. In recent years, infrared thermography (IRT) has emerged as a widely used tool for artifact detection due to its non-invasiveness, immediacy, and high imaging quality [2]. It effectively detects defects by visualizing the temperature distribution on the surface of materials through thermal imaging

Although IRT performs well in detecting near-surface and sub-surface defects, indepth detection is still a challenge as strong thermal excitation may damage the inspected artefacts. In addition, traditional human inspection is gradually becoming infeasible due to time and cost constraints, making the shift to automated defect recognition particularly important.

To address these issues, this study employs a numerical modelling approach to obtain quantitative and reproducible results to guide the actual testing procedure, and introduces deep learning networks and machine learning algorithms to build an automatic defect detection system. Numerical modelling not only simulates the experimental environment to avoid damage to the artwork, but also provides sufficient training and test data for the defect detection system, thus speeding up the detection speed and identifying defects that are difficult to detect with the naked eye.

The test object of interest in this work is the ancient polyptychs. The original polyptych was painted in 1320 by Pietro Lorenzetti, as shown in Fig. 1 (a) The artwork is currently preserved in the Church of Santa Maria del Fiore in Arezzo, Italy. Fig 1 (b) shows a detail of most of the polyptych. It should be highlighted that this replica has been thoroughly described in [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 procedure was intended to approximate the replication of historical artistic methods to test the effectiveness of the defect detection system.
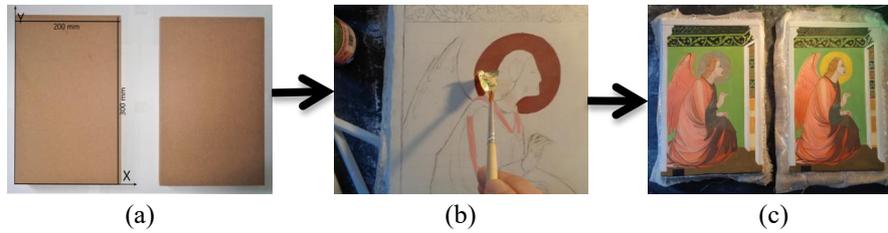

(a) (b) (c)

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 test sample to numerically simulate the temperature distribution on the sample surface. Firstly, the outline of the sample model was sketched using AutoCAD 2024 software and then imported into COMSOL Multiphysics 6.0 software and simulated for IRT experiments using the Solid Heat Transfer module. For more details on the modelling, see the author's 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. And 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 $y := \{Y^1, Y^2, Y^3, \dots, Y^B\}$, where $Y^i \in R^{H \times W}$ ($i = 1,2,3,\dots,B$) represents the ith 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:

$$y = X + N + S \tag{1}$$

where X is the noiseless image of our data, N is Gaussian noise, and 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:

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

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

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

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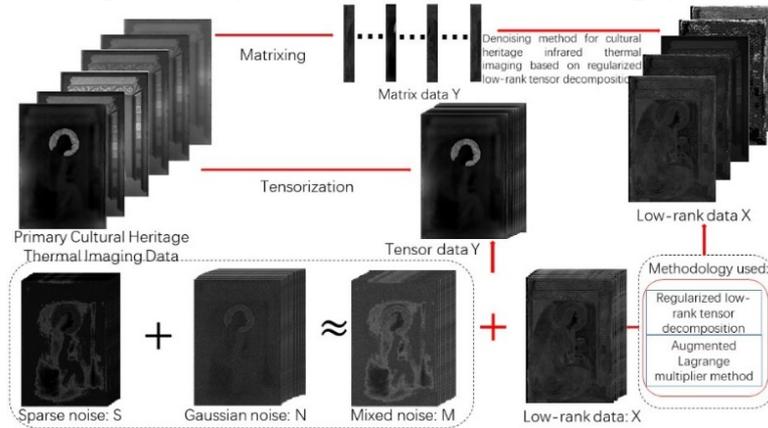

Fig. 3. Schematic diagram of the LRTDTV denoising method.

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

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

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:

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

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 R-CNN model see Fig. 4.
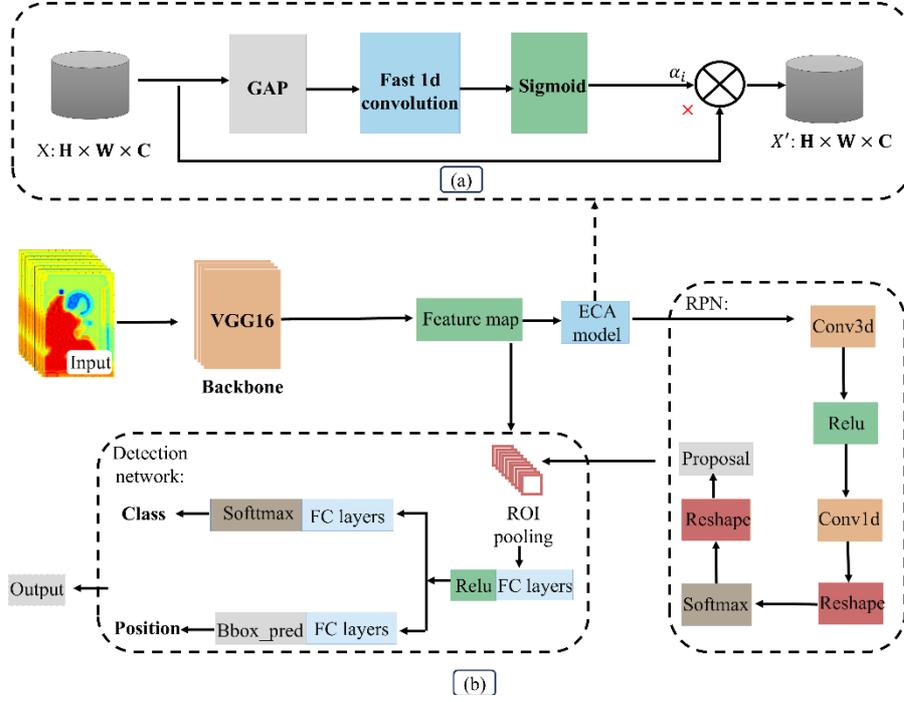

Fig. 4. (a) ECA module structure diagram, (b) Structural diagram of the improved Faster R-CNN network model.

4   Experimental results and analysis

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 R-CNN 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	81.6	70.2	46.3	64.7	50.1	54.8	56.4	0.31
Improved model	87.3	72.2	54.8	66.5	52.2	57.7	57.8	0.33

To get a more intuitive feel for the enhancement of machine learning algorithms and deep learning networks for IRT defect detection, we show the experimental results for sample A and sample B in Fig. 5.
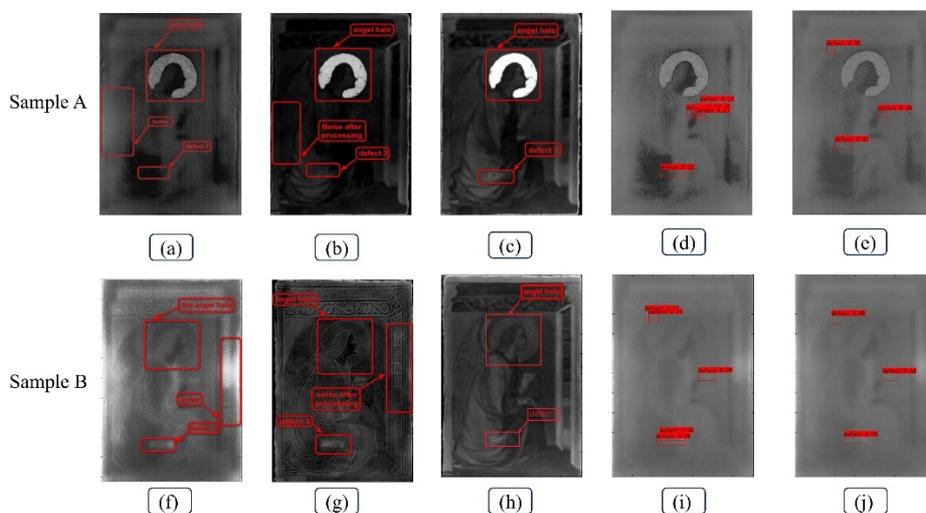

Fig. 5. IRT experimental results: (a) the raw image of sample A, (b) the image of sample A after Fourier transform, (c) the image of sample A after LRTDTV de-noise and Fourier transform, (d) automatic detection results after applying the Faster R-CNN model of Sample A, (e) automatic detection results after applying the improved Faster R-CNN model for sample A, (f) the raw image of sample B, (g) the image of sample B after Fourier transform, (h) the image sample B after LRTDTV de-noise and Fourier transform,(i) automatic detection results after

applying the Faster R-CNN model of sample B, and (j) automatic detection results after applying the improved Faster R-CNN model for sample B.

By observing the results of this experiment, we can find the following points: comparing Fig. 5(d) with Fig. 5(e), Fig. 5(i) with Fig. 5(j), the automatic defect detection network based on the improved Faster R-CNN reduces the error rate and improves the accuracy, and it can better detect tiny defects that are difficult to be detected by human. Comparison of Fig. 5(f), Fig. 5(g) and Fig. 5(h) shows that the LRTDTV denoising method can effectively remove the noise interference in the infrared thermal images.

5 Conclusions

In this study, IRT was used to perform NDT of ancient polyptychs, with care taken to avoid secondary damage. To protect the ancient polyptychs from secondary damage, we used numerical simulation to match the experimental surface temperature, and employed the LRTDTV decomposition model to effectively remove the noise from the thermal images, thus facilitating the detection of sample defects by IRT. In addition, an improved Faster R-CNN model is proposed, which utilizes VGG16 for feature extraction and introduces an ECA mechanism after feature gold extraction, which significantly improves the defect detection of ancient artworks. After comparative experiments, the results of the study demonstrate the efficiency and practicality of the model, which provides strong technical support and greatly improves the accuracy and speed of IRT in cultural heritage conservation.

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).