
| RESEARCH ARTICLE

Automated Textile Quality Control Using Deep Learning: Multi-Class Classification of Fabric Surface Defects

Nihaz Rahman¹, Antor Mahamudul Hashan²✉, Md. Abdur Rahim³, Dhananjoy Biswas⁴

¹A. N. Kosygin Russian State Textile University, Textile Technology and Design, Moscow, Russia.

²University of Science and Technology MISIS, Senior Lecturer, Department of Computer-Aided Design and Engineering, Russia.

^{3,4}University of Science and Technology MISIS, Department of Computer-Aided Design and Engineering, Moscow, Russia.

Corresponding Author: Antor Mahamudul Hashan, **E-mail:** hashan.antor@gmail.com

| ABSTRACT

Fabric defects directly influence the surface appearance and overall quality of textile products, which makes their reliable detection and classification an essential component of quality control in the textile industry. This study applies deep learning, specifically transfer learning with a pretrained convolutional neural network, to automatically recognize six fabric image categories: vertical defects, defect-free fabric, holes, horizontal defects, lines, and stains. A dataset of 2,739 fabric images was divided into 1,369 images for training and 1,370 images for validation. Prior to model input, all images were converted to three-channel colour format, resized, centre-cropped to 224 × 224 pixels, and normalized. An 18-layer residual neural network (ResNet18), pretrained on a large image corpus, was fine-tuned for five epochs using the Adam optimizer with a learning rate of 0.001 and a batch size of 32. Performance was evaluated on the validation partition, which was also reused for final testing, yielding an overall accuracy of 91.17%, a macro-averaged F1 score of 81.10%, and a support-weighted F1 score of 91.12%. Classification results differed noticeably across defect categories: defect-free images and stains achieved F1 scores of 98.12% and 92.27%, respectively, whereas holes proved more difficult to recognize, reaching an F1 score of 68.31% and a recall of 60.58%. These outcomes demonstrate that a fine-tuned deep learning model can effectively distinguish multiple categories of fabric surface condition within the evaluated dataset, while also exposing uneven recognition performance for certain defect types, particularly those with subtle or irregular visual patterns. By linking established textile defect categories with quantitative deep learning evaluation metrics, this work offers a preliminary baseline for automated, CNN-driven fabric quality assessment. The approach is relevant to textile manufacturing organizations, including those in Russia, seeking to integrate intelligent visual inspection systems into their production and quality assurance workflows.

| KEYWORDS

Deep learning, Computer vision, Fabric defect classification, Textile quality control, Visual inspection.

| ARTICLE INFORMATION

ACCEPTED: 19 September 2025

PUBLISHED: 27 October 2025

DOI: 10.32996/jcsts.2026.8.12.1

1. Introduction

Fabric inspection is a fundamental component of textile quality control, since surface defects such as holes, stains, and directional irregularities can adversely affect both product appearance and functional performance [1]. In image-based inspection, a class label identifies an observed surface condition, but it does not necessarily determine the underlying manufacturing mechanism. For this reason, classification outcomes should be interpreted in relation to the textile condition represented by each label rather than judged exclusively by aggregate accuracy. In the present study, fabric inspection is treated as the application context; however, the dataset does not provide verified information on fibre composition or fabric construction.

Deep learning has become an effective framework for image classification, particularly when combined with transfer learning from pretrained convolutional neural networks [2]. Residual networks are especially suitable for this setting because

shortcut connections improve optimization and facilitate feature reuse during fine-tuning [3]. Accordingly, this study investigates whether a pretrained ResNet18 model can provide consistent recognition of six fabric surface categories: vertical defects, defect-free fabric, holes, horizontal defects, lines, and stains. Because the dataset is imbalanced, overall accuracy alone may obscure performance differences across classes, making class-wise evaluation essential for meaningful assessment. This is especially important in textile inspection, where correct recognition of normal material and correct identification of rare defect classes represent different but equally relevant quality-control requirements [4].

The contribution of this paper is an application-oriented analysis of a six-class ResNet18 transfer-learning baseline for fabric defect classification. Rather than proposing a new architecture, loss function, or dataset, the study interprets the recorded model outputs in terms of textile surface conditions and examines variation in recognition performance across categories. This perspective highlights the distinction between detecting a visible defect and inferring its textile cause, while also providing a preliminary benchmark for automated fabric quality assessment within the limits of the available dataset and evaluation design.

2. Related Work

Textile defect classification has attracted increasing attention as deep learning methods have matured for image recognition and industrial inspection. Residual learning, introduced by He et al. [5], established a widely used approach for training deep neural networks through shortcut connections that facilitate optimization in image recognition tasks. In textile inspection, this development has supported the adoption of convolutional architectures for automated defect recognition, particularly in settings where classification performance depends on subtle texture variation and limited labelled data.

Transfer learning has been especially prominent in this area. Beljadid et al. [6] applied GoogLeNet to textile texture defect classification on the TILDA database and demonstrated that pretrained models can be effectively adapted to fabric inspection tasks. Their results also underscore an important methodological point : direct numerical comparison across studies is generally inappropriate unless the datasets, label definitions, and evaluation protocols are closely aligned. Accordingly, the present work should be interpreted as a dataset-specific six-class baseline rather than as a direct competitor to previously reported results.

Related studies have also highlighted the importance of the production context in defining the inspection problem. Chakraborty et al. developed a printed-fabric database and employed a convolutional neural network to analyse print-related abnormalities, illustrating how the formulation of the task depends on the material type and manufacturing process [7]. In a different direction, Zhou et al. [8] proposed a one-class defect detection framework based on defect-free fabric images, combining Gabor feature extraction, autoencoder-based representation learning, and nearest-neighbour density estimation to identify abnormal regions. Although these methods differ from supervised multi-class classification, they collectively demonstrate the range of deep learning strategies available for textile inspection.

Benchmark datasets and label interpretation are also central to the literature. The AITEX Fabric Image Database provides a documented resource for developing and evaluating defect detection methods [9]. CottonWorks similarly organises examples according to fabric construction, defect appearance, and, in some cases, production mechanism [10]. However, labels such as vertical, horizontal, and lines may reflect visible image patterns rather than verified textile fault mechanisms. Their interpretation therefore requires explicit label definitions and, where possible, information on fabric orientation and construction.

Overall, the literature indicates that deep learning can be effective for textile defect inspection, but the meaning of reported results depends strongly on dataset composition, class structure, and evaluation design. In contrast to prior studies that emphasise backbone comparison, anomaly detection, or dataset-specific defect analysis, the present work offers a six-class ResNet18 transfer-learning baseline with class-wise evaluation and textile-oriented interpretation. Its contribution lies not in proposing a new architecture or dataset, but in relating recognition performance across categories to fabric surface conditions and in distinguishing between the identification of a visible defect and inference about its manufacturing cause.

3. Methods

This section describes the dataset used in the study, the preprocessing pipeline applied to the fabric images, the ResNet18 transfer-learning configuration, and the training and evaluation procedures used for six-class fabric surface classification. The experimental setup follows a supervised image-classification workflow in which the model is fine-tuned on labelled fabric images and assessed using class-wise and aggregate performance metrics.

3.1 Fabric Image Dataset And Label Encoding

The study uses the publicly available Fabric Defects Dataset [11], comprising 2,739 images distributed across six categories: 1,666 defect-free images and 1,073 defected images (281 holes, 398 stains, 157 lines, 136 horizontal defects, and 101 vertical defects). In the ImageFolder directory structure, the class indices are assigned as follows: 0 = Vertical, 1 = defect free, 2 = hole, 3 = horizontal, 4 = lines, and 5 = stain, with capitalization retained from the source folder names. The dataset is strongly imbalanced. Defect-free images account for approximately 61% of the full collection, whereas vertical defects represent only about 3.7%. This imbalance is preserved in the train/validation split, so the validation partition reflects the same class distribution. Because the dataset does not provide verified composition or construction metadata, cotton is treated here as the

intended application context rather than a confirmed attribute of every image. In addition, the physical distinction between the vertical, horizontal, and lines categories is not documented in the source material. The validation partition contains 1,370 images, and its class distribution is reported in Table 1.

Table 1. Class distribution of the validation partition.

Class	Images	Share
Vertical	49	3.58%
Defect free	838	61.17%
Hole	137	10.00%
Horizontal	63	4.60%
Lines	84	6.13%
Stain	199	14.53%

To provide a clearer visual overview of the dataset and to illustrate the appearance of the six fabric surface categories, representative sample images from each class are shown in Figure 1, including vertical defects, defect-free fabric, holes, horizontal defects, lines, and stains.

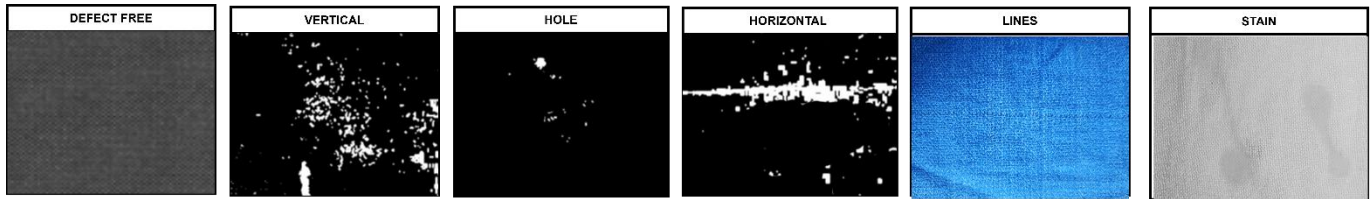


Figure 1. Representative examples from the Fabric Defects Dataset used in this study. Each panel shows one of the six classes: vertical, defect free, hole, horizontal, lines, and stain.

3.2 Image preprocessing

Each image is opened with the Python Imaging Library (PIL) and converted to RGB [12]. The active DataModule [13] transformation resizes the shorter side to 224 pixels while preserving the aspect ratio, applies a 224×224 central crop, converts the image to a tensor, and normalizes it channelwise using the ImageNet statistics [14]: means (0.485, 0.456, 0.406) and standard deviations (0.229, 0.224, 0.225). For each channel c , normalization is defined as:

$$x'_c = \frac{x_c - \mu_c}{\sigma_c}$$

Where x'_c is the tensor value before normalization, and μ_c and σ_c are the specified channel mean and standard deviation. Centre cropping guarantees a uniform input shape but may exclude off-centre defects; the available outputs do not establish whether this occurred in any sample. An earlier standalone transformation included random rotation and horizontal flipping, but the pipeline used for training does not apply these operations.

3.3 Model Architecture And Optimization

A pretrained ResNet18 model was implemented using the timm library [15] and adapted for six-class fabric image classification by replacing the final fully connected layer with a six-way output layer. The network accepts RGB images of size 224×224 pixels and processes them through residual blocks with shortcut connections, followed by global average pooling and a linear classifier that produces six output logits. The overall architecture is illustrated in Figure 2, which shows the feature-extraction backbone and the final classification head used in this study. The model contains approximately 11.2 million parameters, all of which were left trainable during fine-tuning.



Figure 2. ResNet18 architecture for six-class fabric image classification.

Training was performed using the multiclass cross-entropy loss function [16] computed from the six output logits and the corresponding class labels. Optimization employed the Adam algorithm with a learning rate of 0.001, a batch size of 32, and a

total of five epochs. The optimization workflow is summarized in Figure 3, where gradients are propagated backward through the network and Adam updates the model parameters during each training step.

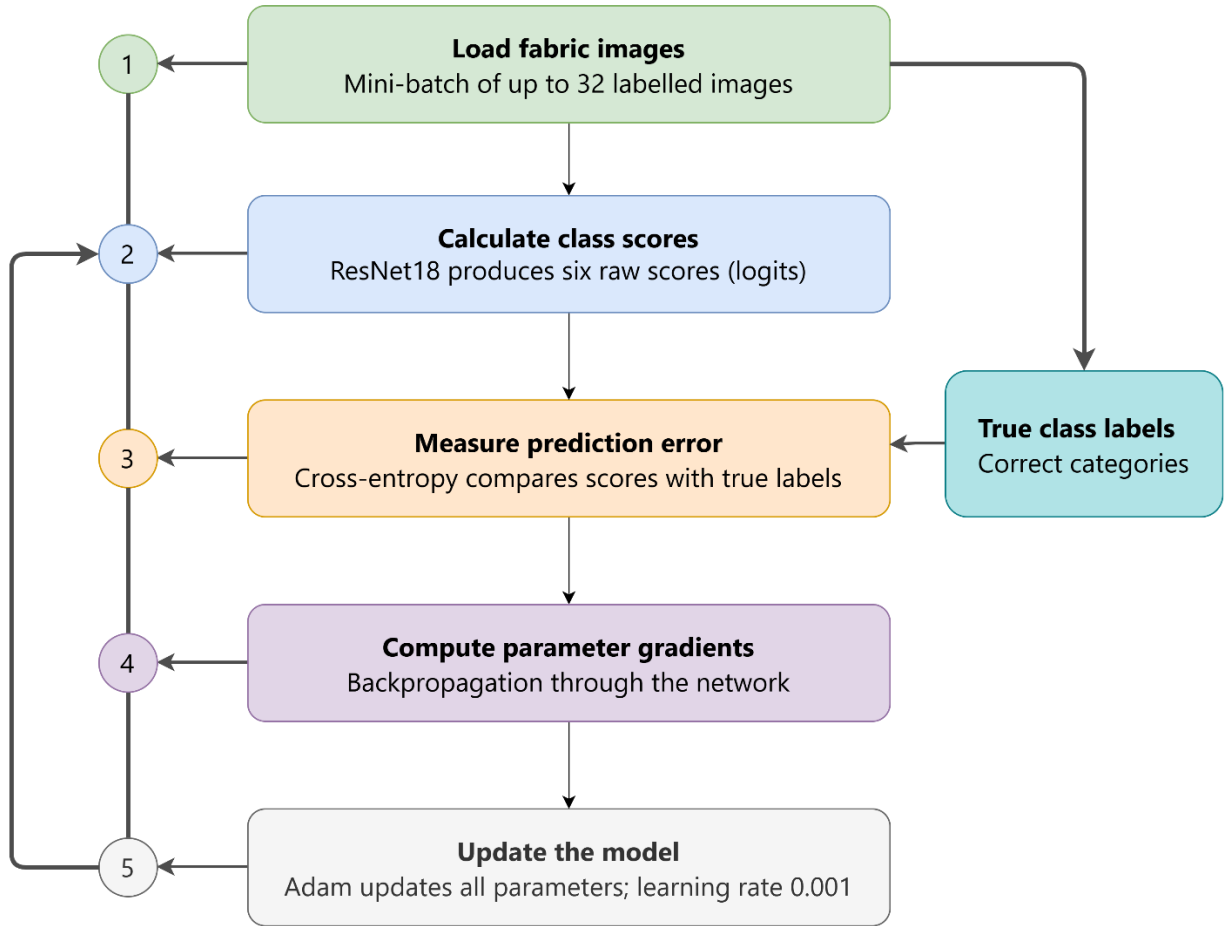


Figure 3. Training loop using cross-entropy loss and Adam optimization.

Table 2 provides a compact summary of the implemented architecture and training configuration, including the backbone, input format, optimizer, learning rate, loss function, number of epochs, and fine-tuning strategy.

Table 2. Implemented model and training configuration.

Item	Recorded configuration
Backbone	Pretrained ResNet18 through timm
Output classes	6
Input	RGB, 224 × 224 pixels
Training / validation	1,369 / 1,370 images
Batch size	32
Optimizer	Adam
Learning rate	0.001
Loss	Multiclass cross-entropy without explicit class weights
Epochs	5
Fine-tuning	All model parameters trainable
Active augmentation	None
Hardware	CPU; no GPU used
Lightning version	2.5.2
Independent test partition	Absent; validation partition reused

Evaluation was carried out on the validation partition, which consisted of 1,370 images and was also reused as the final test set. Performance was reported using accuracy together with class-wise precision, recall, and F1 score, as well as macro-averaged and support-weighted summaries [17]. In this single-label classification setting, macro recall is equivalent to balanced accuracy.

3.4 Results And Discussion

The classifier was trained using PyTorch and timm, with Lightning 2.5.2 managing the five-epoch training process on a CPU [12]. Evaluation was performed on the reused validation partition, and the classification report was generated from the restored checkpoint with gradient computation disabled. The model achieved an overall accuracy of 91.17%, a macro-averaged F1 score of 81.10%, a support-weighted F1 score of 91.12%, and a cross-entropy loss of 0.2150. Because defect-free images accounted for 61.17% of the evaluation set, overall accuracy alone does not fully describe class-level performance, and the lower macro F1 indicates that recognition quality varied across the six categories.

As shown in Figure 3, the class-wise precision, recall, and F1 scores differ substantially across categories. Defect-free fabric and stains achieved the strongest results, with F1 scores of 98.12% and 92.27%, respectively. In contrast, holes were the most difficult category to recognize, with an F1 score of 68.31% and a recall of 60.58%, corresponding to approximately 83 correctly classified images out of 137 hole samples. The remaining hole images were assigned to other categories, although the available report does not allow the exact confusion pattern to be reconstructed. Lines achieved a high recall of 94.05% but a lower precision of 64.75%, indicating that this label was frequently assigned to images belonging to other categories. Vertical and horizontal defects obtained intermediate performance, with F1 scores of 77.42% and 73.77%, respectively.

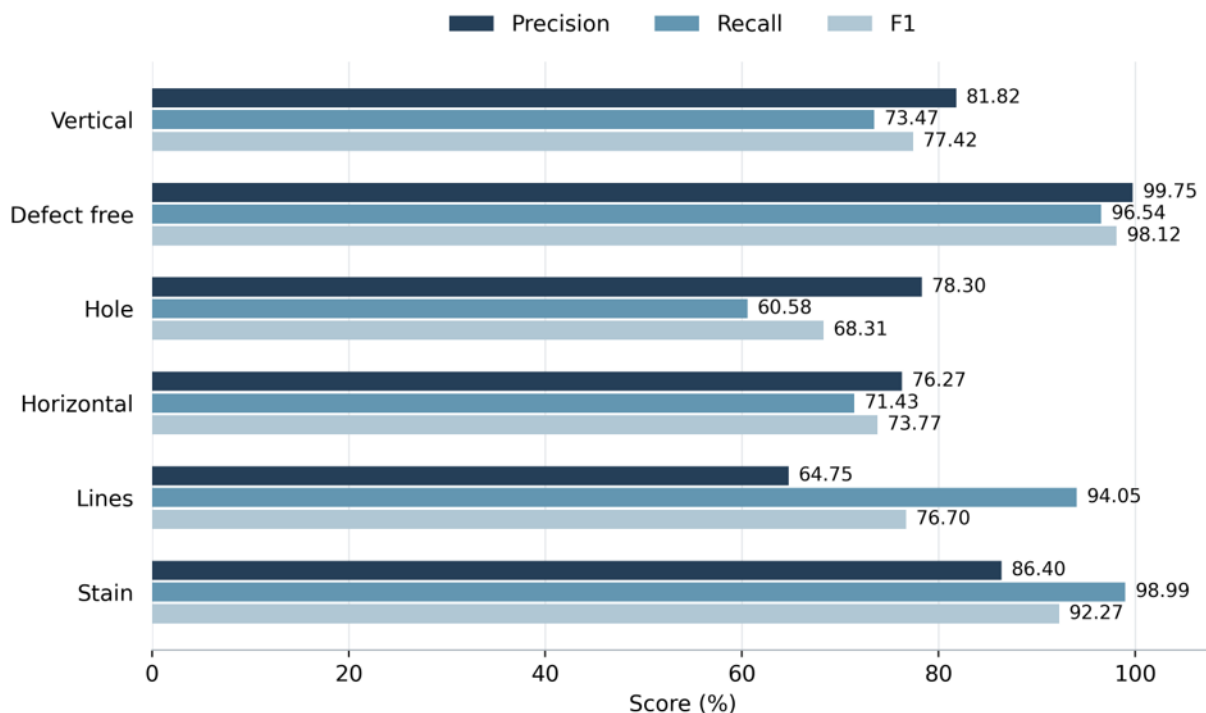


Figure 4. Class-wise precision, recall and F1 on the reused validation partition of 1,370 images. Scores are expressed as percentages.

These results suggest that the model is more reliable for defect-free fabric and stain recognition than for hole classification. Such variation is expected in an imbalanced multi-class setting, where visually distinctive categories are easier to learn than subtler or more irregular defect types. The results should therefore be interpreted as a class-wise baseline rather than as evidence of superiority over other textile inspection methods. For textile quality control, this distinction is important because recognition of a visible surface condition is not the same as inference about its manufacturing cause.

Transfer learning for fabric classification has also been investigated by Beljadid et al. [6], while Chakraborty et al. [7] examined printed-fabric defects and Zhou et al. [8] used one-class abnormality detection. However, differences in datasets, label definitions, and evaluation procedures prevent a direct numerical ranking against these studies. The present experiment therefore provides a six-class baseline without demonstrating superiority over competing methods. The weaker recognition of holes deserves particular attention in cotton fabric inspection, although the dataset composition is unverified. In practical terms, the category predictions may support operator review, but the experiment does not establish roll-level defect grading or deployment performance in industrial settings.

4. Conclusion

This study combines textile fabric assessment and machine learning through a six-class ResNet18 transfer-learning experiment. The recorded model was fine-tuned on 1,369 images and evaluated on a reused validation partition of 1,370 images. It achieved 91.17% accuracy, 81.10% macro F1 and 91.12% weighted F1. Defect-free fabric and stains obtained the strongest F1 scores, whereas hole recognition was substantially weaker, with F1 of 68.31% and recall of 60.58%. Line classification showed high recall alongside lower precision. The findings demonstrate that overall classification accuracy is insufficient to describe recognition of every fabric defect category. The study contributes an interpretation that connects machine learning metrics with textile surface conditions and distinguishes identification of a visible fault from diagnosis of its manufacturing cause. Cotton fabric inspection provides the principal material context, while unverified composition records prevent a cotton-specific performance claim. The evidence supports a preliminary baseline relevant to textile quality assessment and intended inspection applications in textile organizations, including those in Russia, within the limits of the recorded evaluation design.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

- [1] R. Tanjim *et al.*, "Computer Vision-Driven Textile Quality Assessment: An EfficientNet Approach for Multi-Class Fabric Defect Recognition," in *2025 28th International Conference on Computer and Information Technology (ICCIT)*, Cox's Bazar, Bangladesh: IEEE, Dec. 2025, pp. 2074–2079. doi: 10.1109/ICCIT68739.2025.11491584.
- [2] A. M. Hashan, S. M. T. Rahman, R. M. R. U. Islam, K. Avinash, S. Shekhor, and S. M. Iftakhairul, "Smart Horticulture Based on Image Processing: Guava Fruit Disease Identification," in *2023 IEEE 21st Student Conference on Research and Development (SCoReD)*, Kuala Lumpur, Malaysia: IEEE, Dec. 2023, pp. 270–274. doi: 10.1109/SCoReD60679.2023.10563573.
- [3] M. Shafiq and Z. Gu, "Deep Residual Learning for Image Recognition: A Survey," *Applied Sciences*, vol. 12, no. 18, p. 8972, Sep. 2022, doi: 10.3390/app12188972.
- [4] N. Ingle and W. J. Jasper, "A review of the evolution and concepts of deep learning and AI in the textile industry," *Textile Research Journal*, vol. 95, no. 13–14, pp. 1709–1737, Jul. 2025, doi: 10.1177/00405175241310632.
- [5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.
- [6] A. Beljadid, A. Tannouche, and A. Balouki, "Fabric defect classification using transfer learning and deep learning," *IJ-AI*, vol. 12, no. 3, p. 1378, Sep. 2023, doi: 10.11591/ijai.v12.i3.pp1378-1385.
- [7] S. Chakraborty, M. Moore, and L. Parrillo-Chapman, "Automatic defect detection for fabric printing using a deep convolutional neural network," *International Journal of Fashion Design, Technology and Education*, vol. 15, no. 2, pp. 142–157, May 2022, doi: 10.1080/17543266.2021.1925355.
- [8] H. Zhou, Y. Chen, D. Troendle, and B. Jang, "One-Class Model for Fabric Defect Detection," 2022, *arXiv*. doi: 10.48550/ARXIV.2204.09648.
- [9] J. Silvestre-Blanes, T. Alberó-Alberó, I. Miralles, R. Pérez-Llorens, and J. Moreno, "A Public Fabric Database for Defect Detection Methods and Results," *Autex Research Journal*, vol. 19, no. 4, pp. 363–374, Dec. 2019, doi: 10.2478/aut-2019-0035.
- [10] W. H. Fraser, "Statement by the Proprietors of Cotton Works in Glasgow and the Vicinity: Case of the Operative Cotton - Spinners, in Answer to That Statement: Reply by the Proprietors: and An Introduction and Appen dix," in *British Trade Unions, 1707–1918, Part I, Volume 2*, Routledge, 2007.
- [11] F. He, B. Xin, Z. Zhan, W. Yang, and J. Liu, "A review of fabric defect detection in textile manufacturing," *The Journal of The Textile Institute*, vol. 117, no. 2, pp. 327–345, Feb. 2026, doi: 10.1080/00405000.2025.2502188.
- [12] K. Mohaideen Abdul Kadhar and G. Anand, "Python Libraries for Image Processing," in *Industrial Vision Systems with Raspberry Pi*, in *Maker Innovations Series.*, Berkeley, CA: Apress, 2024, pp. 33–72. doi: 10.1007/979-8-8688-0097-9_3.
- [13] S. dos S. Rahoo, "Image restoration with neural networks," *Restauração de imagens utilizando redes neurais*, 2023, Accessed: Oct. 18, 2026. [Online]. Available: <https://lume.ufrgs.br/handle/10183/271387>
- [14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
- [15] X. Zhao, Y. Hong, X. Dai, and Y. Li, "Optimized ResNet18 for few-shot fabric texture image recognition: enhancing accuracy in digital textile applications," *The Journal of The Textile Institute*, vol. 117, no. 10, pp. 3019–3029, Oct. 2026, doi: 10.1080/00405000.2026.2680843.
- [16] J. Terven, D.-M. Cordova-Esparza, J.-A. Romero-González, A. Ramírez-Pedraza, and E. A. Chávez-Urbiola, "A comprehensive survey of loss functions and metrics in deep learning," *Artif Intell Rev*, vol. 58, no. 7, p. 195, Apr. 2025, doi: 10.1007/s10462-025-11198-7.
- [17] A. Mahamudul Hashan, R. Md Rakib Ul Islam, and K. Avinash, "Apple Leaf Disease Classification Using Image Dataset: a Multilayer Convolutional Neural Network Approach," *IA*, vol. 21, no. 4, pp. 710–728, Jun. 2022, doi: 10.15622/ia.21.4.3.