

Contrastive Texture Metric Learning: An Attention-Guided Siamese Model for Fabric Inspection

Feyzanur Ozden¹ and Canan Tastimur¹

¹Department of Computer Engineering, Faculty of Engineering Architecture, Erzincan Binali Yildirim University, Erzincan 24100, Turkey
Corresponding Author: ctastimur@erzincan.edu.tr

ABSTRACT

In smart manufacturing systems, it is crucial to diagnose defects as products pass along conveyor belts using automated computer vision systems. Compared to traditional manual defect diagnosis methods, this approach reduces time loss, high costs, and operator errors. This study proposes a fast Siamese network supported by contrastive learning and attention mechanisms for the detection and classification of fabric defects. A Channel Attention module is integrated during the feature extraction phase to ensure the network focuses on key regions. Thanks to the Siamese structure, samples belonging to the same class are brought closer together in the feature space, while different classes are distanced from each other. This aims to enable the model to more accurately distinguish small defects, especially between similar fabrics. During the training process, contrastive loss and cross-entropy loss were optimized together. In addition, computational costs were reduced using data augmentation methods and mixed precision training strategies. Experimental studies showed that the proposed model achieved 97% test accuracy. Grad-CAM visualization results showed that the model successfully localized defective regions. In conclusion, the proposed method offers a high-performance and practical solution for real-time textile quality control applications.

Date of Submission: 22-06-2026

Date of acceptance: 03-07-2026

I. INTRODUCTION

The textile industry is one of the most important manufacturing sectors, supplying essential products for apparel, automotive interiors, medical textiles, technical fabrics, and industrial applications. Fabric production plays a critical role in modern manufacturing; however, various defects may occur during weaving, knitting, dyeing, or finishing processes due to machine malfunctions, yarn inconsistencies, or environmental factors. Ensuring product quality in textile manufacturing is one of the most challenging issues in contemporary production systems, as even minor surface defects can significantly affect product appearance, durability, and customer satisfaction. With the rapid development of Industry 4.0 technologies, the importance of automation and intelligent quality assurance systems has increased across all stages of textile production.

Quality control is particularly crucial in textile manufacturing because defects such as holes, stains, broken ends, thick or thin yarns, misweaves, scratches, and surface irregularities can reduce the commercial value and functional reliability of fabrics. The textile production process is highly sensitive to such defects, which may negatively impact both aesthetic quality and mechanical performance. Traditionally, quality inspection in textile factories is performed manually by human operators. However, manual inspection is highly dependent on operator experience, concentration, and environmental conditions, which can lead to inconsistent evaluations and overlooked defects. In addition, conventional inspection methods are labor-intensive, time-consuming, and costly, especially in high-speed production environments.

Traditional quality control approaches are often insufficient for modern textile manufacturing systems that require high precision and continuous monitoring. As production capacities increase and quality standards become stricter, relying solely on manual inspection can result in higher operational costs, reduced efficiency, and increased product rejection rates. Therefore, the integration of automated and intelligent inspection systems has become an essential requirement for ensuring reliability, consistency, and productivity in textile manufacturing processes. In recent years, significant attention has been directed toward the development of computer vision and deep learning-based methods for automated fabric defect detection and classification. Machine learning approaches have demonstrated promising performance in identifying complex textile surface defects with high accuracy and speed. Consequently, many studies in the literature have focused on intelligent defect detection systems to improve textile quality control and support fully automated production environments.

Talab and Shakoor [1] used Local Binary Pattern mapping to detect defects in patterned fabrics. While traditional local binary pattern approaches cannot discriminately subclass irregular patterns without information loss, this approach was able to produce more distinctive features. It is particularly robust against large adjacency areas and rotational deformations. The dataset used is a patterned fabric dataset previously used in the literature. This dataset is divided into two main categories: defective and non-defective. The highest accuracy rate was obtained as 88.25%. Chitra et al. [2] performed various feature extraction filtering operations (such as blurring, texture highlighting, and edge detection) on the fabric dataset using five different machine learning approaches: XGBoost, Random Forest, Naive Bayes, CatBoost, and SVM. The highest success rate was obtained with CatBoost, with an accuracy rate of 72%, similarly performing defective and non-defective binary classification. The lowest accuracy was achieved with Naive Bayes at 55% accuracy. When filtering was applied, an accuracy rate of 68% was achieved. While filtering increased the success rate in some algorithms, it had little effect in others.

Kailas et al. [3] went beyond transfer learning approaches and adapted pre-trained models to the fabric dataset, achieving higher accuracy with limited data. Fine-tuning was applied to the Standard InceptionV3 architecture to learn specific patterns and textures in the fabric dataset. By performing binary classification of defective and healthy, they achieved an accuracy of 97%, with 10,820 healthy and 3,000 defective. Jerusha and Kumar [4] proposed a hybrid CNN-LSTM model based on Textile Generative Adversarial Networks for the classification of textile surface defects and also achieved 94% accuracy. The fabric dataset was first preprocessed with Coherence diffusion and log Gabor filters, and thanks to the proposed model, four classes of fabrics were classified: color changing fabric, electronic information fabric, temperature control fabric, and breathable fabric.

Geze and Akbas [5] based their classification of fabric defects on the transfer learning approach. In this study, similar to our study, two classes of defective and healthy fabrics were classified. As a result of using VGG16 and InceptionV3 models, accuracy ratios of 89% and 86% were obtained, respectively. In addition, 95% accuracy was achieved as a result of the hybrid use of a CNN network from scratch and ResNet50 models. Duan et al. [6] applied mining to determine the labels in the classification of fabric defects. In defect classification, more general categories were grouped and additional auxiliary labels were assigned. By collecting data from two different fabric datasets, a new fabric dataset with 24 classes was obtained, demonstrating that the proposed method consistently increased both the classification accuracy and the mAP@0.5 value by more than 1%. An accuracy of 80.84% was obtained on the 24-class dataset.

Feng et al. [7] used an AlexNet-based binary hash classification mechanism to find defects in a large number of fabric images. They used Log-ReLU activation instead of the traditional activation function. A hidden layer was added to learn the binary hash codes, and the parameters in the network layers were optimized. An accuracy of 94.80% was obtained, with an improvement of approximately 10% compared to the classical activation function. Pourkaramdel et al. [8] achieved an accuracy rate of 97.2% in this study, using autocorrelation to learn the repeating dimensions in the fabric to identify defects on the fabric surface. Differences in feature maps between windows in the entire image were measured using Completed Local Quartet Patterns. The maximum value obtained was determined as the threshold value.

Tajeripour et al. [9] provided classification of fabric surface defects using Modified Local Binary Patterns. A reference histogram feature map was created for pixels on healthy fabric surfaces using the LBP operator. The method was also tested for its robustness in the gray level changes in fabric images. Li et al. [10] similarly used the LBP method for feature extraction from fabric images and a Gray Level Co-occurrence Matrix to learn textural relationships. They classified the obtained features using a Support Vector Machine to distinguish between defective and healthy fabrics. Testing was completed on 15 different fabrics with an accuracy of 87.77%. Reddy et al. [11] applied a feature extraction process for defect and similarity classification on fabric surfaces. They used Gray Level Co-occurrence Matrix and Binary Pattern Based Features for feature extraction. For the classification phase, they used K-nearest neighbor and SVM and achieved 100% accuracy with both approaches.

Roomi and Saranya [12] used GLCM and Binary Level Co-occurrence Matrix for feature extraction from fabric images. Then, they used the obtained features to classify fabric defects with a Bayesian classifier. They achieved 92.75% accuracy with the GLCM+Bayes classifier and 95.5% accuracy with the BLCM+Bayes classifier on 120 fabric samples. Mewada et al. [13] used a pre-network trained with ResNet50 and Cuckoo Search Optimization together to detect small and complex defect samples on the fabric surface. In this study, which achieved an accuracy rate of 95.36%, experiments were carried out on two different fabric datasets.

Huang et al. [14] used the DenseNet121 network in this study, adopting a transfer learning approach. In this study, which achieved a validation accuracy of 97.34%, defect classification was performed on two different fabric types, woven and knitted. Mao et al. [15] achieved an accuracy rate of 85.6% in detecting fabric defects, providing approximately 2.5% improvement compared to the YOLOv9 model using the YOLOv10

architecture. They enabled both defect detection and classification with a dataset containing five different defects on the fabric surface.

Table 1: Comparison of fabric defect detection and classification methods reported in the literature.

Ref	Metot	Accuracy (%)
[1]	LBP Mapping	88.25
[2]	CatBoost	72.00
[3]	Fine-tuned InceptionV3	97.00
[4]	CNN-LSTM + TGAN	94.00
[5]	VGG16	89.00
	InceptionV3	86.00
	CNN + ResNet50	95.00
[6]	Hierarchical Label Mining	80.84
[7]	Log-AlexNet	94.80
[8]	CLQP	97.20
[10]	LBP + GLCM + SVM	87.77
[11]	GLCM + Binary Pattern + KNN/SVM	100.00
[12]	GLCM + Bayes	92.75
	BLCM + Bayes	95.50
[13]	ResNet50 + CSO	95.36
[14]	DenseNet121	97.34
[15]	YOLOv10	85.60

In this study, deep feature extraction, Siamese network, and attention mechanism are used together for fabric surface defect classification and detection. The learned representation space is made more meaningful by utilizing a contrastive learning mechanism to learn similarity relationships between defective and non-defective fabric samples. Multilayer CNN blocks are used for deep feature extraction. The feature representation maps of defect regions on the fabric surface are given greater importance using a channel attention mechanism. The proposed architecture is computationally inexpensive because it uses a lightweight CNN architecture instead of large-scale pre-trained models. Furthermore, Grad-CAM based analyses allow for a detailed examination of which image regions the model uses to generate decisions, demonstrating that the network truly focuses on defect regions. In this proposed study, all experimental work was performed on a binary fabric dataset with both healthy and defective fabrics.

II. MATERIAL AND METHODS

A hybrid model of Siamese CNN and channel attention is proposed for the detection and classification of industrial fabric defects. This proposed architecture aims to learn the discrimination between healthy and defective fabric images, as well as the similarity between these visual features. While traditional CNN-based approaches focus only on supervised classification, this study achieves the learning and acquisition of more discriminatory features thanks to the Siamese network structure and contrastive loss. This results in stronger feature representations, especially in limited datasets and textile images with high intraclass variation. Figure 1 shows the overview of the proposed model.

2.1. Dataset Preparation and Preprocessing

The dataset containing defective and healthy fabric images is binary-classified. The dataset used is divided into 80% training, 10% validation, and 10% testing to prevent data leakage. The images in the dataset have undergone several preprocessing steps before being fed into the proposed network. First, the images are resized to 128x128 and the pixel values are normalized between [0,1] by dividing by 255. To reduce overfitting, data augmentation is applied to the images during the training process using Random horizontal flipping and Random vertical flipping. Sample images from the dataset used are shared in Figure 2.

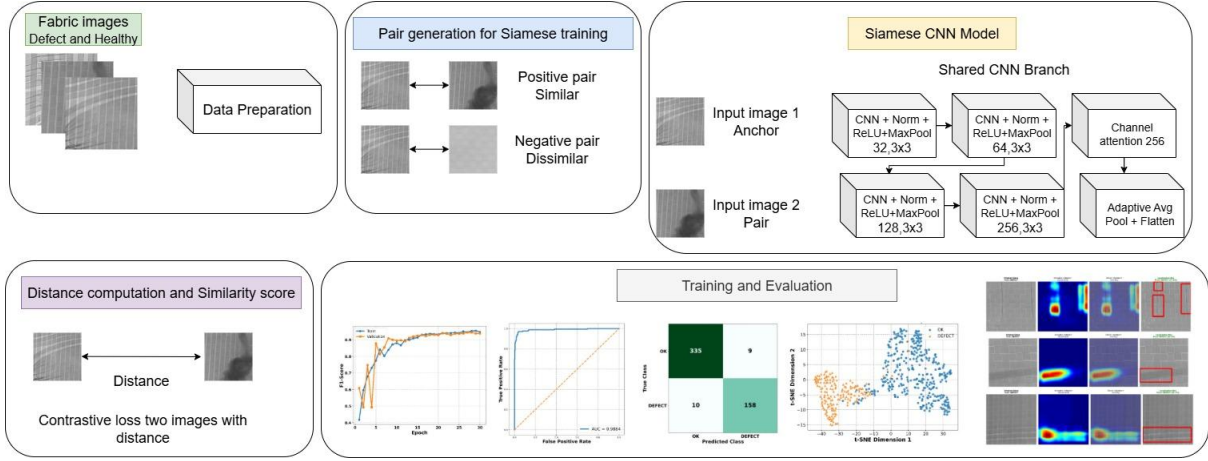


Figure 1: The flow of the methodology

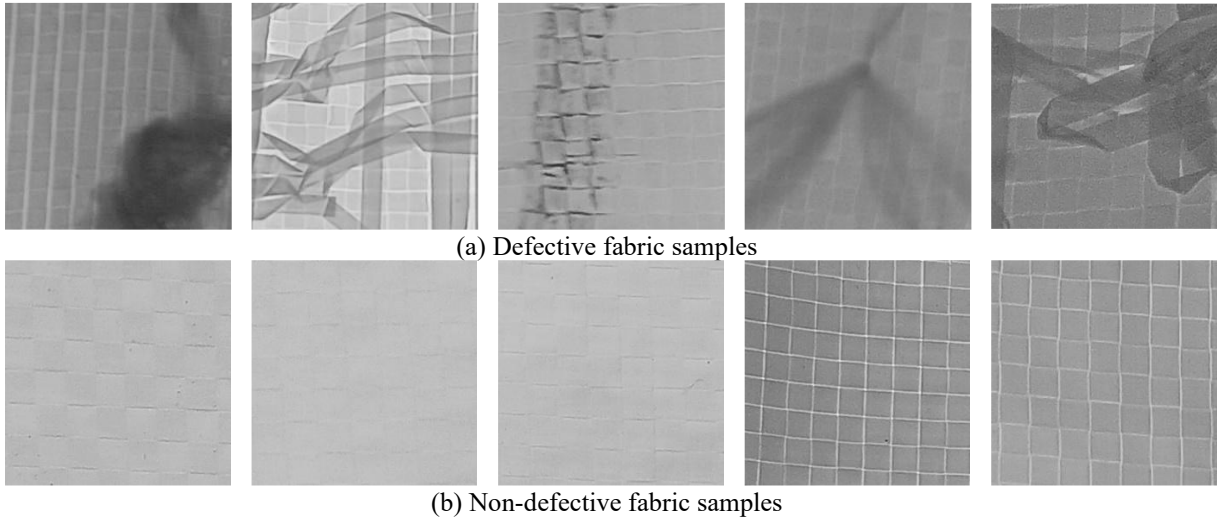


Figure 2: Example images of the fabric dataset

2.2. Siamese Pair Generation

In classical defect classification approaches, models only learn the relationship between the class label and the image, and each image is treated as an independent sample. Detection of subtle textural differences in fabric is not always possible. Furthermore, the most significant problem is that the presence of different samples belonging to the same class negatively impacts model performance. Therefore, similarity-based metric learning algorithms that consider similarities and differences between samples have gained popularity.

In this study, unlike the classical CNN model, Siamese-based learning was applied. The model calculates similarity between different images by calculating similarity from a single sample. During the training phase, the model treats each image in the dataset as an anchor image. To compare this image, pairs are created with positive samples from the same class or negative samples from different classes, and similarity or difference values are calculated. Positive sample pairs indicate two different images without defects or two different fabric images with defects. Negative sample pairs consist of defective and healthy fabric samples. The aim of this network model is to increase the distance between the feature vectors of different classes, i.e., between faulty and healthy examples. As the distance between faulty and healthy feature maps increases in this feature space, the discriminatory power between the classes also increases.

2.3. Channel Attention-Enhanced Feature Extraction Network

The fact that the defects on the textile fabric surfaces are very small and complex reveals that this is important in the feature extraction phase of the model. Thanks to this structure consisting of four consecutive convolutional layers for the feature extraction phase, structural differences in the image can be detected and the distortion process becomes more stable. While features extracted with only convolutional layers may not give sufficient importance to defective regions, the importance level of each channel in the feature map is automatically learned thanks to the channel attention mechanism. The obtained channel descriptors are then passed through two fully connected layers to calculate the channel importance weights. In this way, the

proposed model will give less weight to the background image and more weight to the defective regions in the fabric. In particular, it will be possible to detect defective regions with small dimensions more clearly.

2.4. Embedding Representation Learning

It has low dimensionality and high discrimination with the embedding space where the distinctive features from the fabric images are represented. The presence of samples belonging to the same class in different ways, or variations in brightness and texture in the image, suggest that the image may belong to a different class. By learning the relative relationships between images through embedding learning, it is possible to create stronger and more stable feature representations. The embedding feature maps obtained from the image are passed through the siamese architecture with (1), (2), (3) and two separate embedding vectors are obtained as a result of two identical branches. The similarity relationship between these pairs of images is calculated with Euclidean distance as in (4).

$$e_i = f(x_i, \theta) \tag{1}$$

$$e_1 = f(x_1, \theta) \tag{2}$$

$$e_2 = f(x_2, \theta) \tag{3}$$

$$D(e_1, e_2) = \sqrt{\sum_{i=1}^n (e_{1i}, e_{2i})^2} \tag{4}$$

Here, $D(e_1, e_2)$ represents the feature distance between two images, n represents the embedding dimension, and e_{1i} and e_{2i} represent the respective feature components. The conclusion can be interpreted as follows: If fabric images belonging to the same class are being compared, the two embeddings will be close to each other in space; conversely, if two fabric samples belonging to different classes are being compared, they will be far apart.

2.5. Contrastive Loss Optimization

In the fabric classification problem, the fact that samples belonging to the same class can sometimes be visually quite different from each other is a factor that makes this problem difficult. In two different fabric defect images, the size, location, and appearance of the defect can change. In such complex samples, networks that perform learning based only on class labels may not achieve sufficient discrimination power. The contrastive learning approach, on the other hand, creates a more meaningful and generalizable feature space by optimizing the distance relationships between samples. In each pair of images, pairs belonging to the same class are labeled as positive and 0; samples belonging to different classes are labeled as negative and 1 during the training phase. Our aim is to reduce the distance between samples belonging to the same class and to maximize the distance between negative samples. The contrastive loss function is given in (5) based on the Euclidean distance in (4).

$$L = yD^2 + (1 - y)max(0, m - D)^2 \tag{5}$$

Here, L represents the contrast loss value, y is the pair label, with a value of 1 for positive pairs and 0 for negative pairs, D is the Euclidean distance between the embedding vectors, m is the margin parameter. This function exhibits different behavior for two different cases. The y -value is 0 between positive pairs and 1 between negative pairs. Thanks to this loss function, it produces very useful results even in datasets with classified data samples.

III. EXPERIMENTAL RESULTS

In this study, a detailed experimental study was conducted on our fabric dataset to demonstrate the effectiveness and success of the proposed model. Figure 3 shows the accuracy, training accuracy, and validation accuracy of the proposed architecture, revealing its classification performance. Similarly, comparisons of training and validation for precision, recall, and f1 score are given in Figures 4, 5, and 6, respectively. Table 2 presents the metric results of these graphs. Table 2 also includes the test performance results of the proposed model. Looking at these results, it is observed that it surpasses many models in the literature.

Figure 7 shows the ROC (Receiver Operating Characteristic) curve, and Figure 8 shows the confusion matrix, revealing the success rate at which the model distinguishes between different samples. When the Confusion Mat was examined, it was observed that 9 healthy fabric samples were incorrectly labeled as

defective, and 10 defective fabrics were labeled as normal. However, it correctly labeled 335 normal samples and 158 defective samples. It was noticed that these samples contained finely detailed defects. Figure 9 shows a comparison of the scattered feature space of healthy and defective fabric samples in the dataset with the model's performance in the classification result. As can be seen from this visual, the model can distinguish samples belonging to different classes with a high performance.

Figure 10 shows how the model, in its proposed fabric surface defect detection process, focuses on specific areas to identify defective regions. Examination of these images reveals that the model has a greater detection capability by paying close attention to and identifying defective areas.

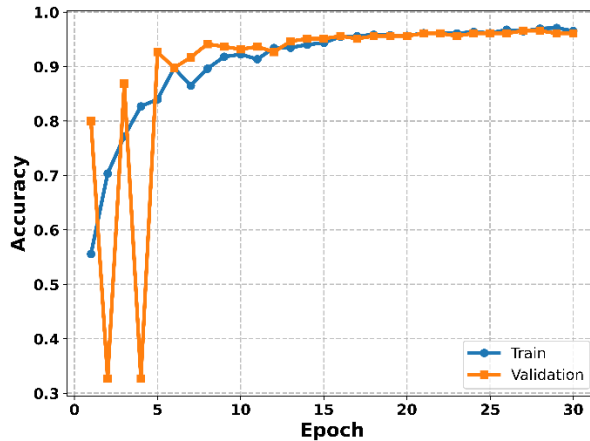


Figure 3: The training and validation accuracy curves of the our model

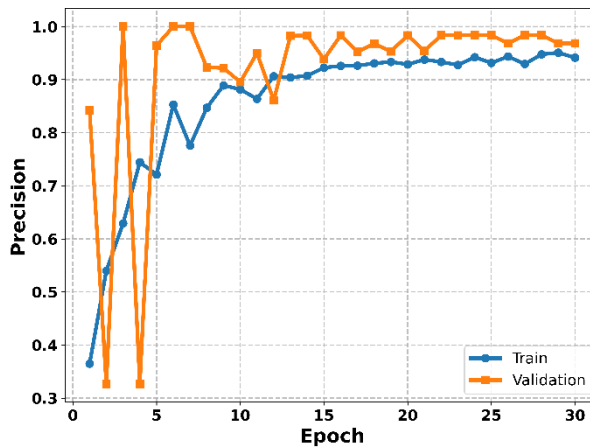


Figure 4: The training and validation precision curves of the our model

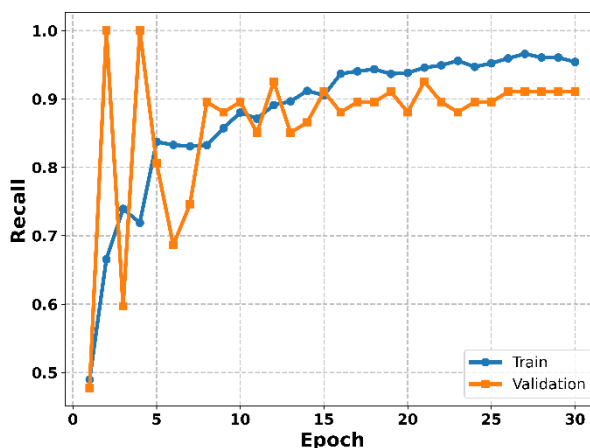


Figure 6: The training and validation recall curves of the our model

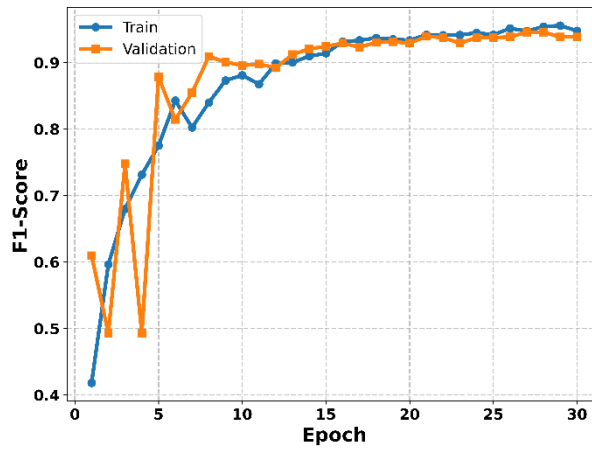


Figure 7: The training and validation f1-score curves of the our model

Table 2: Comparative analysis of training, validation, and test performance of the proposed model

Metric	Train	Validation	Test
Accuracy	0.9654	0.9610	0.9629
Precision	0.9413	0.9683	0.9461
Recall	0.9541	0.9104	0.9405
F1-Score	0.9476	0.9385	0.9433

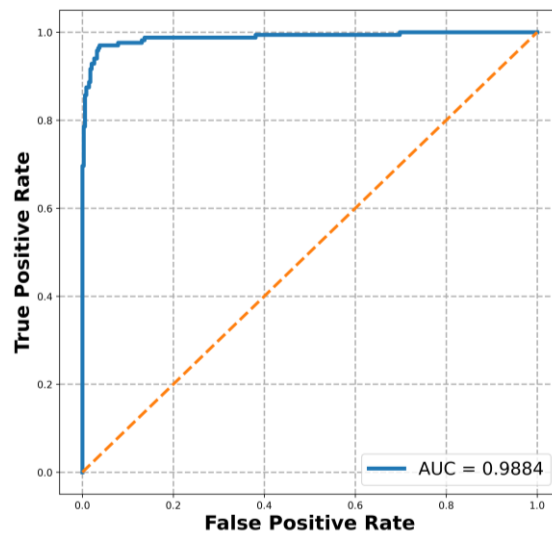


Figure 7: The Roc curve of the our model

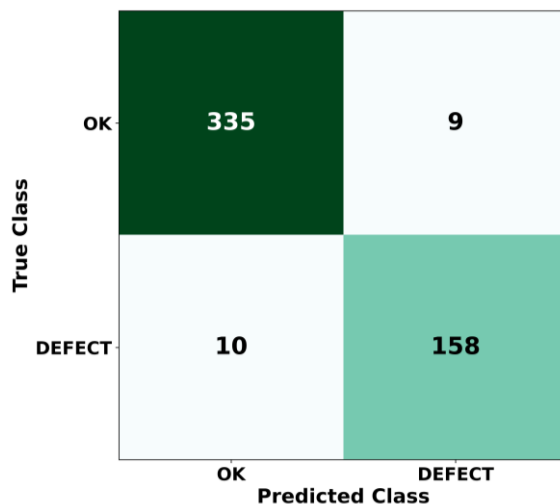


Figure 8: The confusion matrix of the proposed model

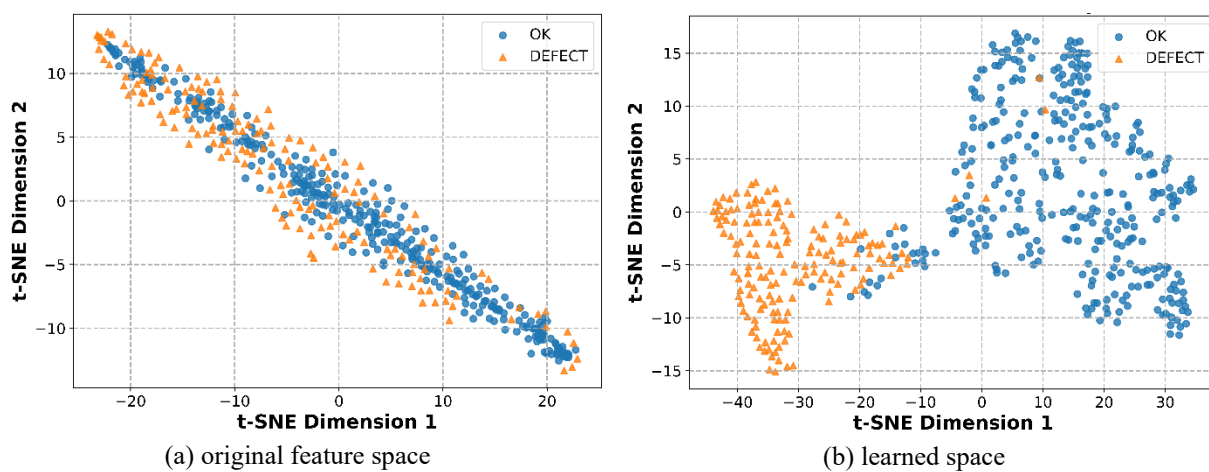
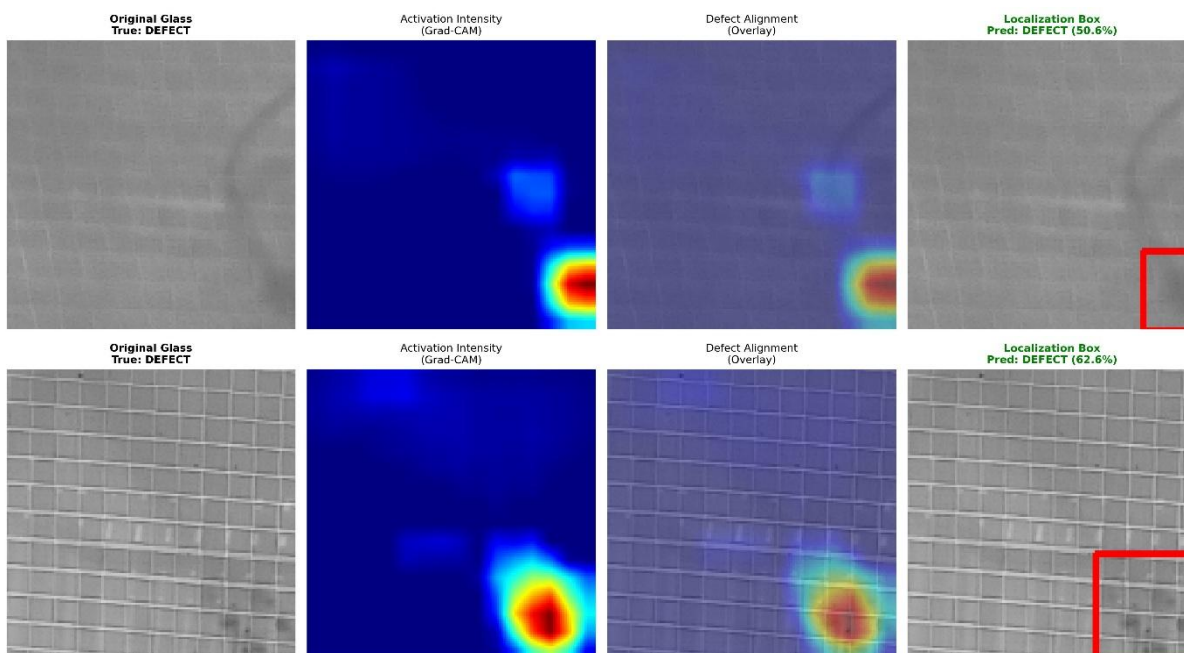
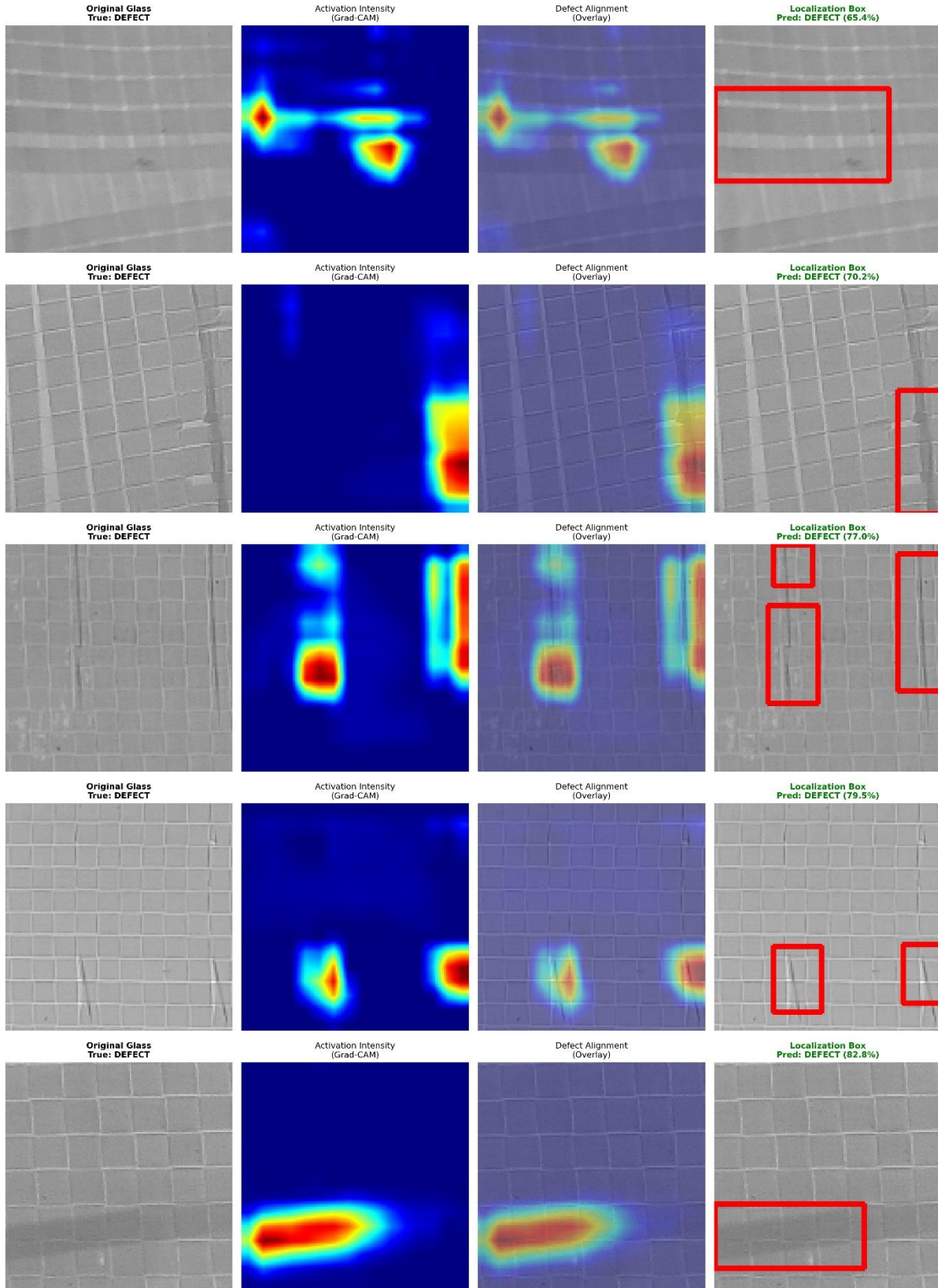


Figure 9: The t-SNE visualization of the proposed model





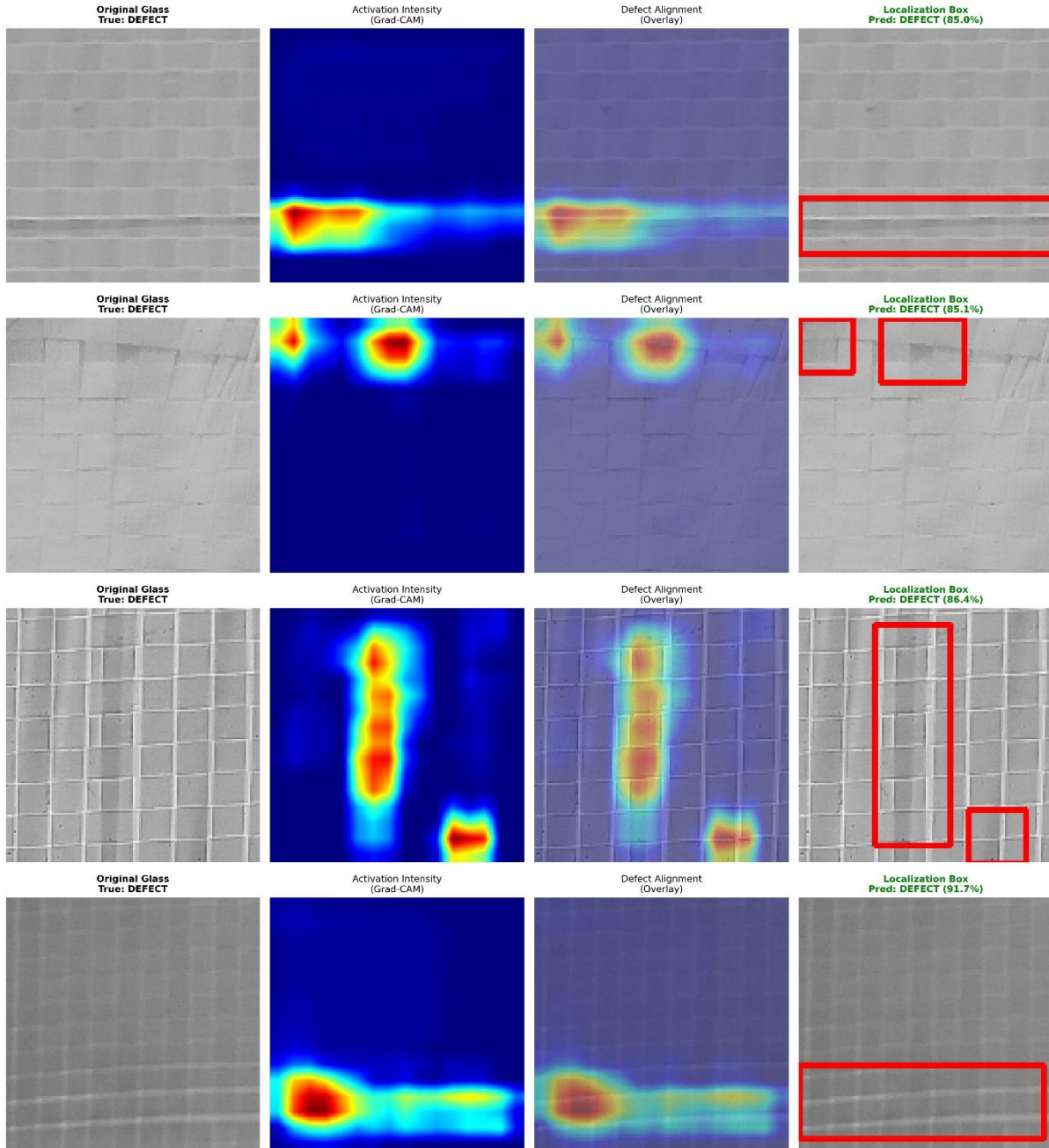


Figure 10: The proposed defect detection results

IV. CONCLUSIONS

This study proposes a hybrid model based on a Siamese network and an attention mechanism for detecting and classifying defects on industrial fabric surfaces. In this study, which uses binary classification (normal and defective), the similarities and differences between fabric patterns are calculated, and a contrast-learning-based Siamese network model is used to distinguish between defective and sound fabrics with high accuracy. The model is also designed to be lightweight to increase its usability in industrial processes. The proposed model is comprehensively evaluated in terms of classification performance and its ability to detect defective areas on the fabric surface, and all results are presented. It has been shown that the model focuses more on defective areas in the input fabric images. This innovative approach, with its high accuracy and low computational cost, promises to be usable in quality control processes in the textile industry. Future studies aim to access more complex defective fabric data, enabling the classification of multiple defects and focusing more on finer, more detailed defects.

REFERENCES

- [1]. A. R. R. Talab and M. H. Shakoor, "Fabric classification using new mapping of local binary pattern," *2018 International Conference on Intelligent Systems and Computer Vision (ISCV)*, Fez, 2018, pp. 1-4, doi: 10.1109/ISACV.2018.8354026.
- [2]. Chitra G M, Shylaja S S, Vijay Ram E, Naresh Sr, Saakshi S, Mohnish S C. Analysis of Image Filters for Binary Fabric Classification, *International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)* Volume 11 Issue 06 June 2024, pp: 35-39, 2024.
- [3]. Kailas, A. S., Yemekar, A., & Ugarakhod, R. N. (2025, December). Enincv3-Enhanced Deep Learning Model for Binary Image Classification of Fabric Defects. In *2025 IEEE 17th International Conference on Computational Intelligence and Communication Networks (CICN)* (pp. 454-459). IEEE.
- [4]. Jerusha, J. A., & Kumar, C. A. (2025). Fabric defect detection: a hybrid CNN-LSTM approach using TGANet for improved classification and traceability. *Industria Textila*, 76(3), 387-396.
- [5]. Geze, R. A., & Akbaş, A. (2023). Detection and classification of fabric defects using deep learning algorithms. *Politeknik Dergisi*, 27(1), 371-378.
- [6]. Duan, X., Lu, W., Cai, W., Tong, F., Mao, Y., & Xu, W. (2026). Improving Fabric Defect Classification via Automatic Mining and Leveraging of Category Hierarchy. *IEEE Transactions on Instrumentation and Measurement*.
- [7]. Feng, M., Chen, D., & Wang, X. (2018, October). Classification of Fabric Patterns Image Based on Improved Log-AlexNet. In *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)* (pp. 1-6). IEEE.
- [8]. Pourkaramdel, Z., Fekri-Ershad, S., & Nanni, L. (2022). Fabric defect detection based on completed local quartet patterns and majority decision algorithm. *Expert Systems with Applications*, 198, 116827.
- [9]. Tajeripour, F., Kabir, E., & Sheikhi, A. (2007). Fabric defect detection using modified local binary patterns. *EURASIP Journal on Advances in Signal Processing*, 2008(1), 783898.
- [10]. Li, P. F., Wang, J., Zhang, H. H., & Jing, J. F. (2012, March). Automatic woven fabric classification based on support vector machine. In *International Conference on Automatic Control and Artificial Intelligence (ACAI 2012)* (pp. 581-584). Stevenage UK: IET.
- [11]. Reddy, R. O. K., Reddy, B. E., & Reddy, E. K. (2013). Classifying similarity and defect fabric textures based on GLCM and binary pattern schemes. *International Journal of Information Engineering and Electronic Business*, 5(5), 25.
- [12]. Roomi, M., and Saranya, S. (2012). Bayesian classification of fabrics using binary co-occurrence matrix. *International Journal of Information Sciences and Techniques (IJIST)* Vol.2, No.2
- [13]. Mewada, H., Pires, I. M., Engineer, P., & Patel, A. V. (2024). Fabric surface defect classification and systematic analysis using a cuckoo search optimized deep residual network. *Engineering Science and Technology, an International Journal*, 53, 101681.
- [14]. Huang, Y., Leonas, K., & Quddus, M. A. (2026). Deep Learning-Based Woven and Knit Fabric Classification Toward Circular Textile Systems. *Circular Economy and Sustainability*, 6(3), 234.
- [15]. Mao, M., Lee, A., & Hong, M. (2024). Efficient fabric classification and object detection using YOLOv10. *Electronics*, 13(19), 3840.