TEAM ID: LTVIP2025TMID41348

PROJECT NAME: *Woven Fabric Pattern Recognition and Classification Based on Deep Convolutional Neural Networks*

Abstract:

The weave pattern (texture) of woven fabric is considered to be an important factor of the design and production of high-quality fabric. Traditionally, the recognition of woven fabric has a lot of challenges due to its manual visual inspection. Moreover, the approaches based on early machine learning algorithms directly depend on handcrafted features, which are time-consuming and error-prone processes. Hence, an automated system is needed for classification of woven fabric to improve productivity. In this paper, we propose a deep learning model based on data augmentation and transfer learning approach for the classification and recognition of woven fabrics. The model uses the residual network (ResNet), where the fabric texture features are extracted and classified automatically in an end-to-end fashion. We evaluated the results of our model using evaluation metrics such as accuracy, balanced accuracy, and F1-score. The experimental results show that the proposed model is robust and achieves state-of-the-art accuracy even when the physical properties of the fabric are changed. We compared our results with other baseline approaches and a pretrained VGGNet deep learning model which showed that the proposed method achieved higher accuracy when rotational orientations in fabric and proper lighting effects were considered.

Keywords: pattern recognition; deep learning; ResNet-50; transfer learning; woven fabric

---

1. Introduction

Fabric is one of the vintage inventions of human beings [1], which has evolved from handmade woven textiles to the modern machine-based electronic textiles. In the textile manufacturing industry, weave pattern, which is the most important factor for woven fabrics, plays a significant role in design and redesign, textural analysis due to its structure, and appearance of fabrics [2–5]. It is essential to perform recognition of woven fabric pattern before it is processed further by weaving machines, in the manufacturing process. Presently, typically, woven fabric pattern recognition is dependent on manual operations using human eyes aided with equipment such as a microscope or a magnifying glass. Traditionally, this manual inspection is performed by an expert who requires expertise and experience. However, it is accompanied with several drawbacks such as extensive labor, inefficiency, and time-consuming, but also leads to subjective human factors, such as mental and physical stress, dizziness, and tiredness, etc., which ultimately affects the recognition results. Therefore, it is indispensable to develop an automated inspection system for the recognition of woven fabric patterns to produce high-quality products that meet the needs of customers.

In recent years, fabric pattern (texture) recognition has gained much attention and made great achievements [6–8]. Usually, methods for weave pattern recognition can be divided into two broad categories, i.e., texture-based statistical method and database/model-based method. The texture-based statistical method uses preprocessed images. Li et al. [9] proposed a method based on photometric

differential analysis, using histogram equalization and adaptive wiener filter to acquire information about fabric structure from different directions, and later, used adaptive mesh model to divide images into subimages for obtaining gray-scale features. The method was robust for extracting interlacing points in the weaving structure, but the work was limited to only simple woven fabrics. The recognition method based on database/model uses a recognition or classification algorithm to identify and match the fabrics weave patterns. Kuo et al. [10] and Pan et al. [11–13] evaluated backpropagation (BP) neural networks for the classification of the prerecognized weave pattern stored in a database of weave patterns based on a white-black co-occurrence matrix. Kuo and Kao [14] used co-occurrence matrix and the CIE-Lab color model to extract texture features, and later, a self-organizing map (SOM) network was implemented to classify the fabric weave patterns. However, the recognition accuracy and performance of these methods were dependent on the size of database. Fan et al. [15] presented a method for yarn segmentation, texture orientation, and texture recognition using K-means clustering method, gradient accumulation of gray levels, and Gabor filters. This method was not suitable for images that had an unclear texture as it neglected the importance of lighting effect while capturing the images. Schröder et al. [16] presented an approach to estimate weave pattern, yarn paths, widths, and their variations for the visual prototyping of woven cloth. This method, to a certain extent, was still dependent on manual selection of model parameters. Trunz et al. [17] put forward an approach for the identification and localization of the occurring stitch types. The underlying grid-like structure of knitted fabric was inferred from these localized stitches. Their experiments were designed only for knit and purl stitch types. The above approaches are reverse engineering-based prototyping of fabrics using a single image only. Yildiz proposed a principal component analysis-based dimensionality reduction method for feature extraction and classification in defective fleece fabric. He compared Naive Bayes and K-nearest neighbor classifier and reported that the latter had better accuracy. This approach only focused on binary patterns rather than complex texture patterns [18]. Li et al. [19] used the local binary pattern (LBP) and gray-level co-occurrence matrix (GLCM) to extract the fabric features and support vector machine (SVM) was used to classify the fabric textures. They merged LBP and GLCM, which improved the recognition performance as compared with the feature extraction algorithm based upon any one of them. This method only relied on handcrafted features and its application on nonwoven and knitted fabric was unexplored. Guo et al. [20] presented a local feature similarity (LFS) method based on seven characteristics of weave pattern for the recognition of weave pattern repeat size. Xiao et al. [21] used transform invariant low-rank textures (TILT) and histogram of oriented gradients (HOG) for the identification of woven fabric pattern, and fuzzy c-means clustering (FCM) was used to recognize the warp and weft cross points. However, the limitations of these studies was the poor recognition of double-yarn weave patterns, as well as large rotational variations that were present in the fabric. Khan [22] introduced a biologically inspired model to recognize the fabric weave and color, and it was robust when the yarn color and fabric appearance were changed. Liu et al. [23] proposed a fabric defect recognition method using optimized convolutional neural network (CNN) with a visualization technique for complicated textures. They preferred only a limited number of images for experiments but in a very constrained environment.

There are several limitations to the recognition methods of woven fabric. The existing studies relied heavily on handcrafted features engineering. The availability of woven fabric images database was limited. The rotational variations in fabric directly influenced the texture features extraction during image acquisition. The improper lighting effect could result in unclear texture images.

Inspired by these previous works and to address the shortcomings, we propose a deep learning model based on data augmentation and transfer learning technique for the recognition and classification of woven fabric texture. A pretrained convolutional neural network (CNN) model, based on residual network architecture (ResNet-50) have been used as a deep model [24]. The proposed model performed end-to-end fabric feature extraction and classification using the woven fabric images that we had developed. In addition, we evaluated different performance metrics such as accuracy, balanced accuracy, precision, recall, and F1-score on the proposed model. We further compared the proposed

model with the VGG-16 pretrained CNN model. The result demonstrates that the proposed method achieves better classification accuracy as compared with the other traditional approaches. Hence, our approach should assist textile manufacturers to save time, cost, and extensive labor through the reliable, efficient, and consistent evaluation of weave patterns to produce high-quality fabrics.

The remainder of this paper is organized as follows: In Section 2, we describe the framework of our proposed model, details of the dataset, data augmentation, along with the DCNN models used; in Section 3, we present our experimental results and performance metrics used; in Section 4 we present a discussion and comparison with other works; and finally, the conclusions are drawn in Section 5.

2. Materials and Methods

2.1. Convolutional Neural Networks

In recent years, convolutional neural networks (CNNs), which are capable of recognizing patterns in images, have achieved remarkable performance in the fields of object recognition [25,26], tracking [27], and, especially, image processing [28]. CNN architectures automatically learn the high-level descriptive features instead of relying on handcrafted features, as is done in traditional machine learning algorithms. A typical CNN consists of several building blocks, namely convolutional, pooling, and fully connected layers. The shallow CNN networks such as AlexNet [29] and VGGNet [30] are formed by stacking several blocks together. Alternatively, deeper CNN architectures are more complex as they use complex alternating connections among layers, such as ResNet.

2.2. VGGNet

The VGGNet architecture contains 144 million parameters with a stack of small-sized convolutional kernels [30]. It is comprised of 16 convolutional layers with small-sized kernels (3 × 3), five max-pooling layers, three fully connected layers, and an output classifier layer with Softmax nonlinear activation. Since the architecture contains a large number of parameters as compared with AlexNet, it is more expensive computationally because it requires an extensive amount of memory.

2.3. ResNet

He et al., in 2016 [24], were the first to introduce the concept of the residual network (ResNet). The main advantage of ResNet is that it solves the problem of vanishing gradient and degrading accuracy by introducing a concept of shortcut connections making it flexible, task-dependent, and capable of training extremely deep neural networks. These shortcut connections are allowed to skip one or more subsequent layers. The pretrained ResNet-50 model is shown in Figure 1.
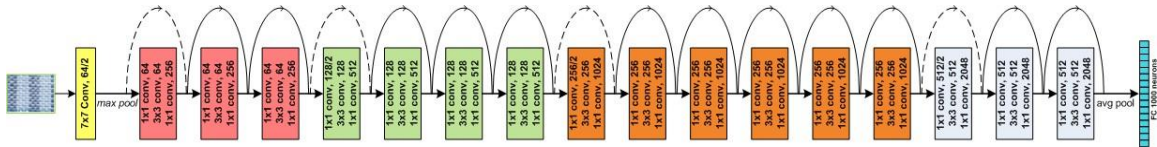

Figure 1. ResNet-50 pretrained network with residual connections.

The basic concept is to use "identity shortcut connections" by skipping the blocks of convolutional layers. These basic blocks are called "bottleneck" blocks which follow two heuristics as follows: (i) for the feature map having the same output size, the depth of filter map remains the same; (ii) if the size of the output feature map is halved, then, the depth of the filter map is doubled. The convolutional layers perform downsampling directly with a stride of 2; between each convolution and ReLU activation function, a batch normalization is processed. On the one hand, for the same dimensions of input and output, identity shortcut is used. On the other hand, for unequal dimensions, the linear projection shortcut is used to insert 1 × 1 convolutions and match the dimensions. In both these cases, when

the identity shortcuts exceed the size of two feature maps, they are implemented by a stride of 2. The architecture ends at the fully connected layer having 1000 neurons activated by Softmax function. These 1000 neurons indicate the 1000 classes of ImageNet challenge [29].

The deep architectures usually contain millions of parameters and training them with random weights initialization can take weeks. A huge amount of data and heavy computations are involved (GPUs) to train these CNNs from scratch. To suppress these problems, transfer learning is commonly used, allowing a CNN model to train on a large dataset, and the features learned while training are transferred to the new model. The last fully connected layer is removed, and then the remaining layers are treated as a feature extractor to adapt to the new task. Consequently, the dense layers of the proposed model are trained.

CNN performs convolutions to extract patterns from the images. The proposed CNN model is trained on a very large ImageNet dataset. In the first few layers, it learns basic patterns such as dots, lines, edges, and diagonals, and later layers combining these basic patterns to complex patterns. Hence, the last layer is capable of learning significant objects such as ships, cars, dogs, fruits, etc. Using transfer learning, we transfer the features (weights) learned from these layers to the new model. The important features learned from the pretrained model for the identification of ImageNet objects are used to classify woven fabric images. As a result, the transfer learning method speeds up the training process and the new CNN model is constructed easily.

2.4. Proposed Model

We proposed a pipeline-based approach for our deep learning model. The pipeline consisted of several stages; the first stage received the fabric images, which ended with the classification of the model. The output of each stage in the pipeline acted as the input to the next stage. The proposed pipeline approach is shown in Figure 2 and the details are described as follows:
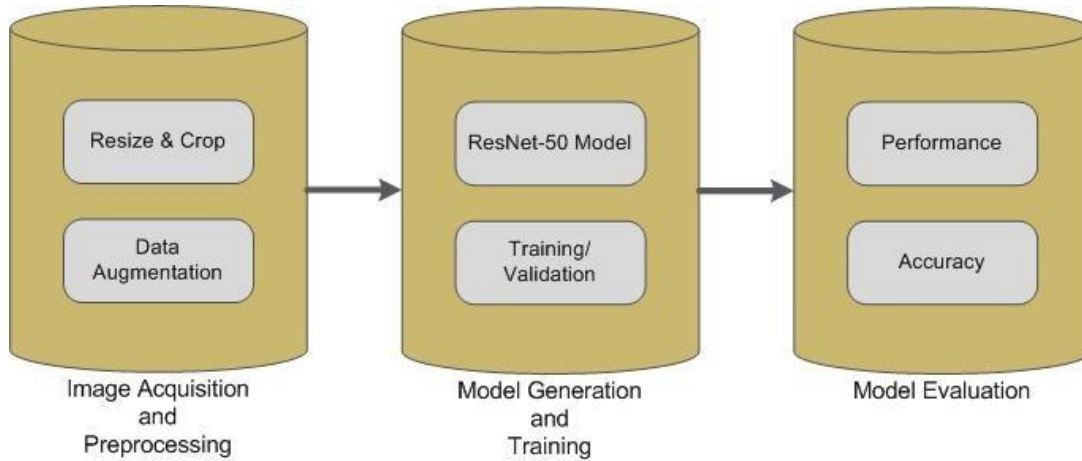

Figure 2. Proposed pipeline approach.

- **Image Acquisition and Preprocessing**

The woven fabric images were collected to form a dataset. The dataset required suitable conversions, resizing, and preprocessing of the images. The number of images was smaller in size, so we used various augmentation techniques to increase the dataset, which helped the model have a good generalization and accomplish better recognition.

- **Model Generation and Training**

A learning algorithm that receives input data "X" (map into attributes to the target) and predicts the output "Y" is called a model. For our model, we employed residual network (ResNet-50) architecture. During training, the algorithm performed optimization on the parameters (update weights and biases) which was used for the recognition of the model.

- **Model Evaluation**

The performance of our model was evaluated using various evaluation metrics such as accuracy, balanced accuracy, precision, recall, and F1-score.

2.5. Dataset

The image acquisition of the woven fabric texture images was done using a digital camera surrounded by a light source to control the lighting illumination conditions, as shown in Figure 3. The Nikon D5600 digital camera attached with a micro Nikorr lens of focal length 45 mm was used. The ISO speed value was set to 100 with a f-stop of f/2.8. The woven fabric samples were collected from various warehouses and textile factories. We captured 3540 images from different locations of 880 pieces of fabric. Out of these 3540 images, we kept 2832 images for our testing dataset, while the remaining 708 images were applied through various techniques of data augmentation to generate a total of 11,328 training samples. A few sample images in each class are shown in Figure 4. These images were subdivided into three classes, namely plain, satin, and twill weave fabrics.
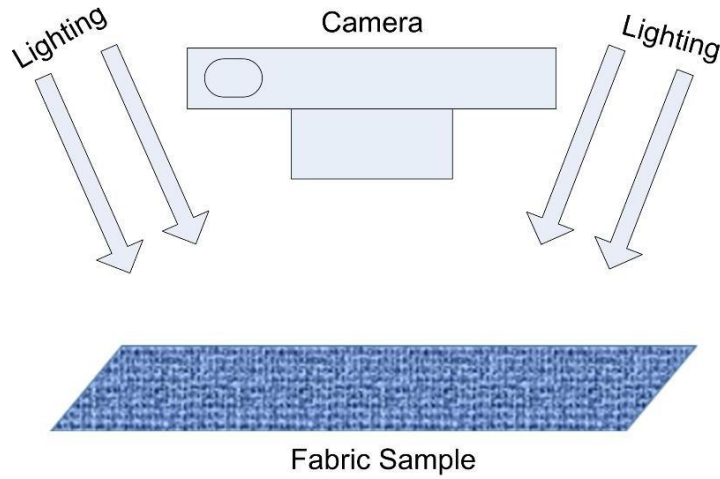

Figure 3. Woven fabric image acquisition system.

2.6. Data Augmentation

The problem of insufficient size of the training dataset has been solved using techniques of data augmentation [31]. Data augmentation performs several manipulations such as scaling, skewing, flipping, and lighting on the entire dataset to form a set of different images as a result expanding the dataset. For larger datasets, deep learning models perform very well. By using augmentation, the total number of images in the dataset is increased, allowing the model to train effectively. It is known that data augmentation is a kind of regularization implemented on the overall dataset, consequently, it reduces overfitting problem and the generalization capability is increased by expanding the dataset, which is the major issue, without performing any alterations that affect the structure of the model. Woven fabric image datasets are not easily available and they are difficult to collect.

In this study, we applied several augmentation techniques on the images such as horizontal and vertical flips, shifting, rotation (images are rotated at fixed angles of 30° starting from 0°, 30°, 60°, 90°, and so on), zooming, shearing, and brightness manipulation. An illustration of these augmented images is shown in Figure 5. It was important to rotate the images in order to identify the warp and weft yarns that were oriented in different directions due to the variations that occurred during the image acquisition. Zooming clearly identified the interlacing pattern of woven fabrics. Shearing created the local deformation in the images. All these augmentation techniques related to the situations occurring in the real scenario.

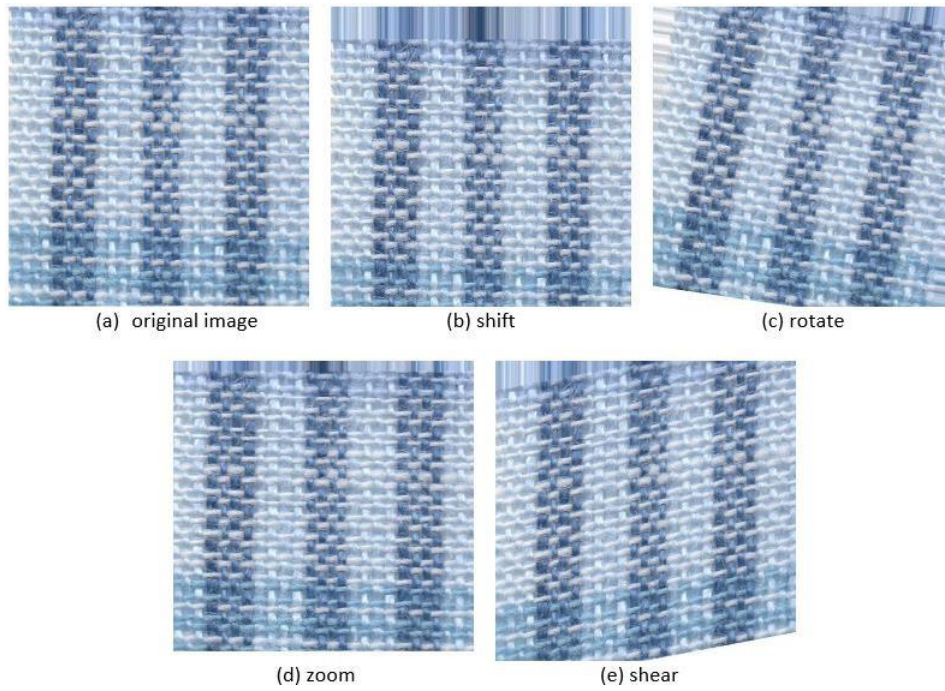**Figure 4.** Few samples of woven fabric images from the dataset.

Figure 5. Various augmentation techniques applied to an original image.

3. Experimental Results

The ResNet-50 pretrained CNN architecture is used for the classification of woven fabric images divided into 3 classes. The number of training and testing images in the woven fabric dataset are 11,328 and 2832, respectively. In this work, 80% of the training images are used to train the model and the remaining 20% is assigned to form a validation subset for validating the model. The performance of the

proposed deep learning model was evaluated on the test set. The structural properties of the woven fabrics used for the experiments are as follows: Note that the values are mentioned in ranges, yarn linear density (Ne: 6–40), yarn count per cm (25–58 ends/cm), and fabric areal density (125–485 gsm).

3.1. Experimental Framework

The woven fabric images were reshaped to 224 × 224 dimensions and the images were also preprocessed by subtracting the mean red-green-blue (RGB) value from each pixel so as to feed it into the model. In this work, we employed transfer learning method to ResNet-50 pretrained CNN architecture, which used the weights of the network learned from ImageNet. The pretrained weights were used to avoid the poor initialization of the model as compared with its counterpart "random initialization of weights". We removed the fully connected layer, which was the last layer that classified the images into ImageNet classes, and the early convolutional layers of the pretrained model acted as a base network for the new customized architecture. Afterward, a global average pooling layer followed by two pairs of batch normalization, fully connected, and dropout layers were, respectively, stacked to the base network. The two fully connected layers encompassed 512 and 256 neurons, respectively. Each fully connected layer was followed by a ReLU activation layer. The batch normalization layers helped to improve the training time of the pretrained model. The inclusion of global average pooling and dropout layers inherently reduced the problem of overfitting. By adding dropout layers, it randomly deleted the redundant neurons; hence, the performance of the model was enhanced. In deep architectures, the problem of overfitting usually fails to have a good generalization on the data which has never been seen before (test data). Finally, the last layer of the proposed model used the Softmax activation function to classify the woven fabric images into three classes. An overall outline of the customized deep learning model is shown in Figure 6.
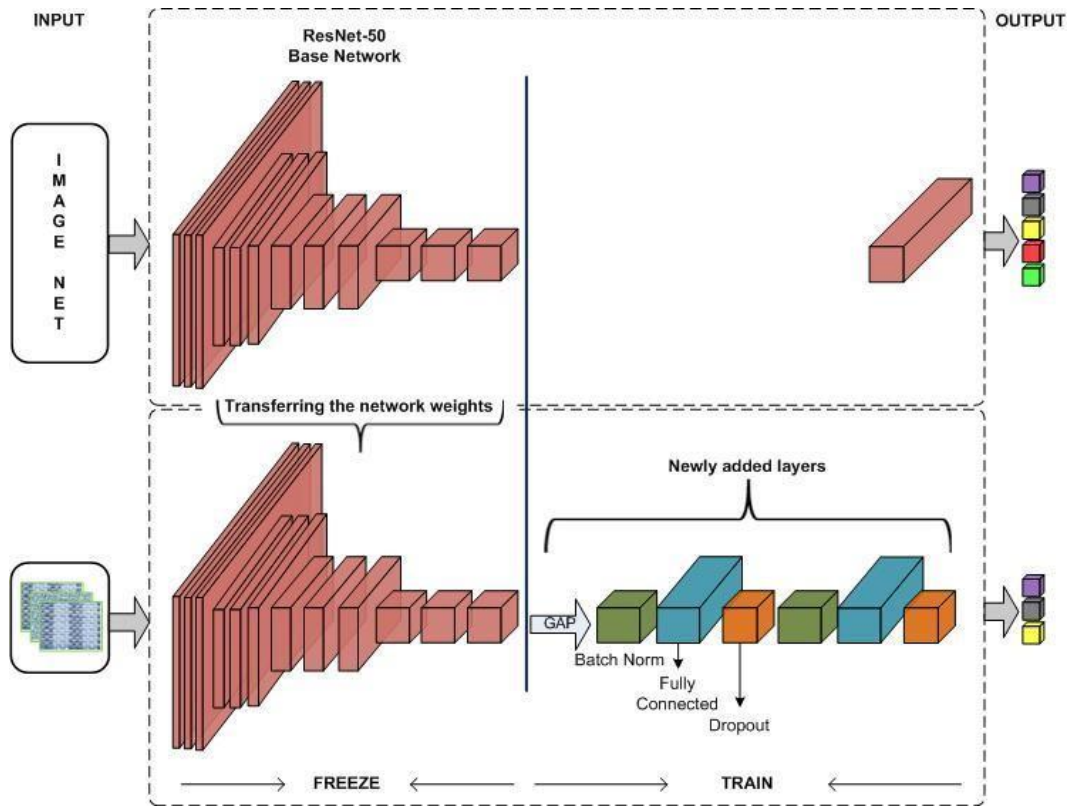

Figure 6. Overall framework of the proposed model.

The pretrained ResNet-50 model was trained on the woven fabric dataset generated in this work. In this architecture, only the customized newly added layers attached to the base network were trained, keeping the initial convolutional layers as frozen. The main idea of freezing these layers was to improve the convergence rate, as well as avoid the gradient explosion during the training process. After texture features were extracted, then, classification was carried out to compare the predicted class with the actual class. During the training process, the computation cost of the network was decreased since the total trainable parameters of the customized CNN model were also reduced.

The proposed model used a method for stochastic optimization, namely Adam optimizer for the optimization of the parameters. The learning rate was set to 0.0001. The dropout ratios for both the dropout layers were chosen as 0.50. The batch size was set to 32.

3.2. Results

The proposed model was trained and tested using NVIDIA GeForce GTX 1060 MQ using 6 GB graphical processing unit (GPU) with Intel i7-8750H @ 2.2 GHz and 16 GB RAM. We used Python 3.6 to implement the model, using Keras library as frontend and Tensorflow as backend.

Evaluation Metrics

The most commonly used evaluation metric for classification is accuracy. It is the ratio between the number of correct predictions to the total number of predictions. Usually, when the dataset is imbalanced, we encounter high accuracy showing that it is inclined towards the class, having more samples of images. In an exceptional situation, each test case can be assigned to the large class by the classifier; as a result, accuracy is achieved equal to the fraction of the more frequent labels in the test set. Thus, accuracy can be a confusing evaluation metric. In such cases, the appropriate performance evaluation metric used is the balanced accuracy, as shown in Equation (1). The class count is represented by "l". We achieved an accuracy of 99.3% and a balanced accuracy of 99.1%.

$$Balanced\ Accuracy = \frac{\sum_i^l (TP_i + TN_i)(TP_i + FP_i + TN_i + FN_i)}{l} \tag{1}$$

The confusion matrix is obtained to show the predictions made by the proposed model on the test dataset and to understand the number of images incorrectly classified. The true class and predicted class are represented by the rows and columns, respectively. Confusion matrix results are shown in Figure 7.
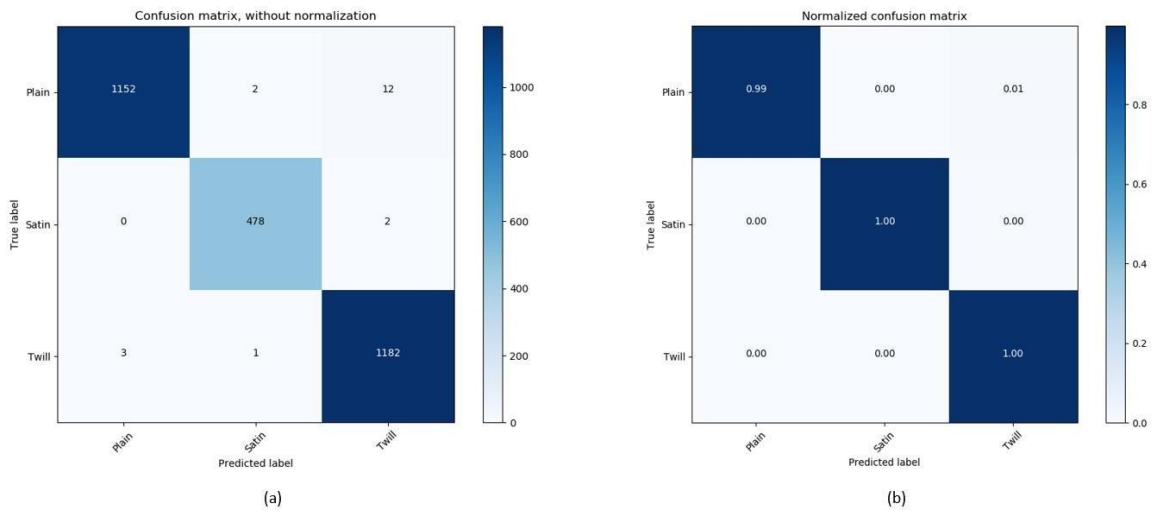

Figure 7. Confusion matrix for plain, satin, and twill weave fabrics. (**a**) Confusion matrix, without normalization; (**b**) Normalized confusion matrix.

The averaged precision, recall, and F1-score values were calculated for the proposed model on the test dataset. The detailed performance analysis of these values based on each class are given in Table 1. The average precision, recall, and F1-score values were 99%.

Table 1. The classification report generated on the test dataset.

Fabric Type	Precision	Recall	F1-Score	Number of Test Samples
Plain	1.00	0.99	0.99	1166
Twill	0.99	1.00	0.99	480
Satin	0.99	1.00	0.99	1186

In this work, we also compared the results with the VGG-16 pretrained CNN model. A significant difference in performance is observed in terms of accuracy and other evaluation metrics. As shown in Table 2, the model outperformed the VGG-16 pretrained model. The VGGNet architecture only obtained a classification accuracy of 92.4% which was about 7 percentage points less than the accuracy obtained by our proposed model. The most likely reason for this was that the VGGNet contained more trainable parameters (134M) and it also did not contain the skip connections to make the computations easier. On the basis of the results, we can say that the proposed model based on residual network ResNet-50 is a robust woven fabric classification CNN that is able to extract robust texture features for recognition and classification.

Table 2. Performance evaluation on our proposed method and VGG-16 pretrained model.

Model	Precision	Recall	F1-Score	Accuracy	Balanced Accuracy
VGG-16	0.918 ± 0.032	0.923 ± 0.016	0.920 ± 0.010	0.924 ± 0.009	0.921 ± 0.013
ResNet-50	0.983 ± 0.015	0.991 ± 0.006	0.986 ± 0.013	0.993 ± 0.003	0.991 ± 0.002

4. Discussion

In previous studies, researchers have proposed methods based on traditional machine learning techniques involving handcrafted features, making it a more tedious and time-consuming job. Table 3 presents a comparison of this work with other baseline approaches.

Table 3. Performance comparison against baseline approaches.

Authors	Method		Accuracy
	Feature Extraction	Classification	
Li et al. [19]	LBP + GLCM	SVM	87.77
Kuo et al. [14]	CIE + Co-occurrence matrix	SOM	92.63
Xiao et al. [21]	TILT + HOG	FCM	94.57
This work	ResNet-50		99.30

LBP, local binary pattern; GLCM, gray level co-occurrence matrix; SVM, support vector machine; SOM, self-organizing map; TILT, transform invariant low-rank textures; HOG, histogram of oriented gradients; FCM, fuzzy c-means.

Li et al. [19] proposed LBP and GLCM for feature extraction and later used SVM as a classifier to classify the woven fabric types. Their proposed method reported an accuracy of 87.77%. Kuo et al. [14] used CIE-Lab color model and co-occurrence matrix to extract features, and then applied a SOM network for classification. The highest classification accuracy obtained was 92.63%. Xiao et al. [21] proposed a method based on TILT and HOG to identify the texture features and later FCM clustering was used for classification. Their method achieved an accuracy of 94.57%. The problem with these approaches was that the datasets were limited to only a small number of images. The authors also excluded the consideration of uneven light during image acquisition, as well as physical properties such as yarn thickness, diameter changes, and rotational variation of fabric.

As shown in Table 3, the proposed ResNet-50-based CNN model outperformed other methods. In our opinion, the proposed deep learning model achieved better accuracy for recognizing and classifying the woven fabric images. Texture-based classification of fabrics is a challenging task because the availability of datasets is limited, and therefore in order to make it more robust, variations in the fabric color, yarn diameter, orientation, and uneven light are considered while performing image acquisition. Hence, making the datasets more detailed and representative. A neural network which has depth, such as the ResNet model, is capable of managing these variations and learning the high-level descriptive features. Hence, the model shows that it is diverse and can easily handle the complexity present in the woven fabric images by outperforming the other approaches. Furthermore, the data augmentation techniques increased the diversity of the available data, and as a result, improved the overall performance of the model.

The higher recognition and classification accuracy reflect the following three main features: (i) The model is still robust when variations in physical properties such as fabric color, yarn thickness, diameter, orientation, and uneven light are considered. (ii) The transfer learning technique allows us to train the model with a lower number of parameters making it computationally cost-effective. (iii) The proposed model does not rely on handcrafted features, i.e., the feature extraction and classification are performed in a fully automated end-to-end architecture. In the future, our work can be extended for recognition and classification of other types of fabric, such as nonwoven and knitted fabrics.

5. Conclusions

In this paper, we proposed a customized deep learning model for the recognition and classification of woven fabrics. The proposed deep learning model is based on residual network (ResNet-50) architecture. First, the image acquisition and preprocessing of fabric images are completed and the data augmentation techniques are applied to increase the size of dataset. Second, a pretrained CNN model is used where only the newly attached layers are trained keeping the other layers frozen. The high-level texture features are extracted, and then finally classified based on the types of woven fabric (plain, twill, and satin). We evaluated the performance of our model using various performance metrics such as accuracy, balanced accuracy, precision, recall, and F1-score. A comparative analysis was carried out against other baseline approaches and we also compared the results of the VGGNet pretrained model. The experimental results showed that the proposed method performed better than the other existing studies. The model is robust when variations such as fabric color, yarn thickness, rotational orientation, and uneven light are considered. The proposed model uses fewer parameters while training making it computationally cost-effective, and thus possess potential for the textile and fashion industry. In the future, we intend to extend our work to other woven and nonwoven fabric types.