Fabric defect detection plays an important role in controlling the quality of textile production. In this article, a novel fabric defect detection algorithm is proposed based on a multi-scale convolutional neural network and low-rank decomposition model. First, multi-scale convolutional neural network, which can extract the multi-scale deep feature of the image using multiple nonlinear transformations, is adopted to improve the characterization ability of fabric images with complex textures. The effective feature extraction makes the background lie in a low-rank subspace, and a sparse defect deviates from the low-rank subspace. Then, the low-rank decomposition model is constructed to decompose the feature matrix into the low-rank part (background) and the sparse part (salient defect). Finally, the saliency maps generated by the sparse matrix are segmented based on an improved optimal threshold to locate the fabric defect regions. Experimental results indicate that the feature extracted by the multi-scale convolutional neural network is more suitable for characterizing the fabric texture than the traditional hand-crafted feature extraction methods, such as histogram of oriented gradient, local binary pattern, and Gabor. The adopted low-rank decomposition model can effectively separate the defects from the background. Moreover, the proposed method is superior to state-of-the-art methods in terms of its adaptability and detection efficiency.

**Keywords**
Multi-scale convolutional neural network, fabric image, low-rank decomposition, defect detection

Date received: 26 February 2019; accepted: 26 November 2019

**Introduction**

Fabric defect detection is the key phase of textile quality control. Traditional fabric defect detection is mainly performed through a visual inspection of skilled workers. However, its reliability is restricted by vision fatigue and human errors. An automatic visual textile detection system based on machine learning can provide a promising solution that not only reduces high labor costs but also improves accuracy and efficiency of fabric defect detection.\(^1\)

Traditional machine-vision-based fabric defect detection methods can be categorized into two categories,\(^2,3\) namely, non-feature extraction methods and feature extraction methods. Among the non-feature extraction methods, Gabor filtering is the most effective method.\(^4\) However, it has strict requirements on filter parameters in order to achieve good performance. Li and Zhang\(^1\) proposed an embedded machine vision system using Gabor filters and pulse-coupled neural network to automatically identify defects of warp-knitted fabrics, which consisted of image
enhancement implemented by Gabor filtering with optimal parameters to make the defects more obvious, and image segmentation achieved by a parameter adaptive pulse-coupled neural network layer by layer. Bo et al. proposed an unsupervised learning method of the training model without the image label, which uses k-singular value decomposition (k-SVD) method to learn the sparse dictionary from an image block, and then space pyramid pooling and orthogonal matching pursuit method to build hierarchical characteristics from a dictionary. However, this method would bring higher feature dimension and cannot be used in the visual task of large-scale database. Li et al. proposed an effective fabric detection method based on biological vision modeling (BVM), which simulated the mechanism of biological visual perception and applied a robust feature descriptor from the biological modeling of P ganglion cells to characterize fabric texture, and then low-rank representation adopted to model visual saliency. In addition, image segmentation techniques, according to the image’s color, texture, shape features, and so on, can partition an image into homogeneous regions. An image segmentation method based on Hough transform is used to detect the target contour directly by the global characteristics of image space and parameter space before and after image transformation. Mathematical morphology based on erosion, dilation, opening, and closing provides an effective approach to analyzing digital images. Morphological filters exploit geometric rather than analytic features of signals. The advantages of the morphological over linear filters are direct geometric interpretations, simplicity, and efficiency in hardware implementation.

Feature extraction methods are often widely used, mainly to extract fabric images’ texture, color, shape, and spatial relationship characteristics. Based on the extracted feature, template matching, neighborhood information, Fourier transformation, and wavelet decomposition are adopted to localize the defect region. Due to their diversity, fabric texture and defect are difficult to efficiently characterize using one kind of feature; this reduces the adaptivity of the defect detection methods. Mak et al. detected fabric defects using previously trained Gabor wavelet networks and morphological elements with a linear structural element. However, this method fails to detect the defect in the fabric images with black and white patterns. Tsai et al. proposed a regularity measurement for defect detection in non-textured and homogeneously textured images using principal component analysis (PCA), which is an orthogonal transformation used to transform linearly and non-linearly the correlation of the source variables into a subspace in which the variables are not correlated. It is widely used in feature extraction and data compression, as well as typically utilized for data pre-processing in defect detection. Notably, the features of the scale invariant feature transform (SIFT) method were originally extracted at scale-space extrema and used for feature point matching. The SIFT method is invariant to scale, rotation, and shift. The features of the histogram of oriented gradient (HOG) method and later SIFT method were densely computed by entire image pyramids. They all described the patch of an image in terms of image gradient histogram. On the basis of feature extraction, some methods are operated at the patch level instead of the pixel level, and each pixel is simply assigned the saliency value of its enclosing patch. Furthermore, all image patches are treated as independent data samples for classification or regression even when they are overlapping.

With the advances of artificial intelligence, the feature extraction is gradually integrated by a deep-learning algorithm in the training process. Deep neural networks (DNNs), especially convolutional neural networks (CNNs), can automatically extract and learn in-depth features of an image, which have been proved to be better than hand-crafted features extracted by carefully designed algorithms. CNN’s convolution operation significantly reduces the number of parameters in a trained model and improves the model’s efficiency, thereby avoiding complex feature selection and manual extraction. In addition, CNNs can directly process the original test images and generate the multi-layer feature of complex fabric texture images by multiple nonlinear transformations. Deep learning has achieved very good results in some tasks, mainly boosted by the feature learning performed, which allows the method to extract specific and adaptable visual features depending on the data. Hinton et al. applied an unsupervised representation learning algorithm to help learning internal representations by providing a local training signal at each level of a hierarchy of features. Unsupervised representation learning algorithms can be adopted several times to learn different layers of a deep model. Several unsupervised representation learning algorithms retain many properties of artificial multi-layer neural networks, relying on the back-propagation algorithm to estimate stochastic gradients. Abid adopted a polynomial interpolation and multilayer perceptron method to train a neural network to detect and locate regions of defects. Ren et al. presented a generic deep-learning method that used a pre-trained network and transferred features to build classifier, and then convolved the trained classifier over input image to make pixel-wise prediction. Li and Yu trained a DNN for deriving a saliency map from multi-scale features extracted using deep CNNs. Wang et al. adopted a DNN to learn local patch features for each centered pixel. Weimer et al. proposed a novel deep CNN architecture to detect defects, which took all types of defect-free and defective samples together as the input. Li et al. proposed a discriminative representation for patterned fabric defect detection using Fisher criterion-based stacked denoising auto-encoders (FCSDA). Fabric images were divided into patches of the same size. Then, these fabric patches with defect-free and
defective classes were used to train FCSDA. Finally, test patches were classified using FCSDA in defective and defect-free classes. Experimental results indicated that the FCSDA method could obtain the superior results on a complex Jacquard warp-knitted fabric.

Low-rank decomposition (LRD) techniques\(^{30}\) can divide an image matrix into two parts: low-rank matrix and sparse matrix, where the low-rank matrix indicates a smooth background and the sparse matrix indicates the salient regions. It has been successfully used in a variety of applications, such as subspace segmentation,\(^{31}\) visual tracking,\(^{32,33}\) image clustering,\(^{34}\) and video background–foreground separation.\(^{35,36}\) Shen and Wu\(^{37}\) provided a unified framework for integrating high-level knowledge and low-level features, which is based on the assumption that an image could be represented as the sum of the background being low rank and the salient regions being sparse. Peng et al.\(^{38}\) introduced a tree-structured sparsity-inducing norm regularization to provide a hierarchical characterization for saliency detection through low-rank and structured sparse matrix decomposition. Yang et al.\(^{39}\) proposed a saliency detection method of constructing an affinity matrix and scoring each node with its similarity to background and foreground cues through graph-based manifold ranking. Wang and Huang\(^{40}\) proposed a salient object detection method using low-rank approximation and \(l_{2,1}\) -norm minimization, which is based on an underlying assumption that an image is a combination of background regions being low rank and salient objects being sparse. The normal fabric images with complex textures have large visual redundancy, and the defects are outstanding from the background. The efficiently deep feature can make the background lie in a low-rank subspace, the defect region deviate from the background. Therefore, a novel fabric defect detection algorithm is proposed based on deep feature and LRD. Deep learning is used to extract a CNN feature, and the LRD model is used to separate the background and the defect region.

Proposed algorithm

Although fabric images have numerous kinds of defects and a complex texture, the defects are more salient in the complex texture background. Therefore, it is of great value to study the defect detection by combining a deep-learning method and LRD technology. The existing deep-learning method typically uses the features of the last convolution layer to carry out the target detection, which causes the loss of detailed texture information. Fabric image texture is relatively abundant, and the low-texture features for defect detection and recognition are crucial; so, a novel fabric defect detection algorithm is proposed based on MCNN and LRD, and its overall structure is illustrated in Figure 1. The multi-layer features are extracted and integrated by MCNN, preserving the semantic information of the high-level characteristics and the detailed texture information of the low-level features. LRD technique was then adopted to divide the generated feature matrix into the low-rank matrix that indicates the background and the sparse matrix that indicates the salient defect. In the end, the iterative optimal threshold segmentation algorithm was utilized to segment the saliency maps generated by the sparse matrix to locate the fabric defect area.

Unsupervised pre-training

When the labeled data are insufficient, an auxiliary supervision training can be adopted, and fine-tuning the particular area can also improve its effectiveness. The current study adopts the sparse auto-encoder (SAE) to pre-train the filters of CNN, and its purpose is in line with the statistical characteristics of datasets and achieving a better initial value of filter sets. For each extracted patch, dictionary was generated by SAE, which can describe their characteristics. And then, sparse matrix was obtained through the linear combination. In the network, the overall cost function is defined as

\[
J_{\text{sparse}}(W, b) = J(W, b) + \beta \sum_{j=1}^{s} KL(\rho \| \hat{\rho}_j) \quad (1)
\]

where the first term is the traditional basic neural network, as shown in formula (2); the second is the sparse penalty factor, as shown in formula (3); and \(\beta\) is the weight parameter balancing the sparse penalty factor and \(J(W, b)\) as well as \(s_2\) which is the number of neurons in the hidden layer

\[
J(W, b) = \frac{1}{m} \sum_{i=1}^{m} \left[ \frac{1}{2} h_{w,b}(x^{(i)}) - \hat{y}^{(i)} \right]^2 + \frac{\alpha}{2} \sum_{j=1}^{n} \sum_{s=1}^{s_1} (W_{\beta}^{(s)})^2 \quad (2)
\]

where the first term is a mean-square-error term; \(h_{w,b}\) denotes an activation function, which is actually a non-linear transformation function with parameters \(w\) and \(b\); and the second term represents a weight attenuation term to
prevent overfitting, and the weight attenuation parameter $\theta$ is adopted to balance the above-mentioned two terms. Relative entropy is generally employed to measure the dissimilarity between two probability distributions. The definition of relative entropy between two mean Bernoulli random variables of mean $\rho$ and mean $\hat{\rho}_j$ is given by

$$
KL(\rho \| \hat{\rho}_j) = \rho \log \frac{\rho}{\hat{\rho}_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \hat{\rho}_j}
$$

(3)

where $\rho$ denotes the sparse parameter and is typically a smaller value close to 0 and $\hat{\rho}_j$ represents the average activity of the hidden neurons $j$. It can be seen from formula (3) that the average activity $\hat{\rho}_j$ approximates to the sparse parameters $b^{(l)} = b^{(l)} - \alpha \frac{\partial}{\partial b^{(l)}} J_{\text{sparse}}(W, b)$ by sparsity constraint of the SAE neural network. In other words, most of the neurons are inhibited, and only a few of them are activated.

After the overall cost function is obtained, the parameters are updated according to the following formulas

$$
W^{(l)} = W^{(l)} - \alpha \frac{\partial}{\partial W^{(l)}} J_{\text{sparse}}(W, b)
$$

(4)

$$
b^{(l)} = b^{(l)} - \alpha \frac{\partial}{\partial b^{(l)}} J_{\text{sparse}}(W, b)
$$

(5)

where $\alpha$ is the learning rate. The latter two derivatives of formulas (4) and (5) are calculated by the back-propagation algorithm. The updating of the whole coding network is completed until the parameters are converged and the characteristic parameters $W$ and $b$ are obtained. It is known that $m$ is the number of hidden layer nodes of layer $l$, and $W^{(l)}$ is then decomposed into a parameter set of the number $m$. Each parameter set is a filter, which leads to a pre-trained filter set.

**MCNN feature extraction and training**

Considering that the fabric image has complex texture and diversities, MCNN model can learn a hierarchy of features from the raw image input by automatically updating the filters during training on massive amounts of training data. So, the MCNN is adopted to extract a deep feature of fabric images. The architecture of the proposed MCNN is presented in Figure 2.

A typical MCNN architecture consists of several nested convolutional layers and pooling layers followed by fully connected layers at the end. In the convolutional layer, the feature maps of the previous layer are convolved with the learned convolution kernels, and an activation function then acts on that value to form the output feature maps. For each block of fabric images, its output size after convolution can be expressed as

$$
\alpha_m = \left[ \frac{\alpha_m^{l-1} - W_m^l}{P_m + 1} \right] + 1
$$

(6)

$$
\alpha_n = \left[ \frac{\alpha_n^{l-1} - W_n^l}{P_n + 1} \right] + 1
$$

(7)
where \( l \) denotes the image layer, \( (\alpha_m, \alpha_n) \) is the input image size of each layer, and \((w_m, w_n)\) is the convolution kernel, where each convolution kernel operates on the effective area of input image. Ignoring parameter \((P_m, P_n)\) specifies the image pixel ignored by convolution kernel. The output node of a convolutional layer is expressed as

\[
\alpha^l_{mn} = \sum_{k} \left( \alpha^{l-1}_{mn} \times w^l_{mn} \right) + b^l_n
\]

where \( \alpha^l_{mn} \) and \( \alpha^{l-1}_{mn} \) are the feature maps of the current layer and the previous layer, respectively; \( w^l_{mn} \) is the convolution kernel that is from the \( m \)th feature map of the previous layer to the \( n \)th feature map of the current layer; \( f(x) = \max(0, x) \) is the neuron activation function; and \( b^l_n \) represents the neuronal bias. The output node of the pooling layer can be expressed as

\[
\alpha^l_{n} = \frac{1}{s^2} \sum_{s \times s} \alpha^{l-1}_{n} + b^l_n
\]

where \( s \times s \) is the scale of the down-sampling template and \( w^l_n \) is the weight of the template.

The fully connected layers are the last part of the neural networks. All of the neurons in the fully connected layers are connected to all of the units of the last layer. Therefore, the final output of the CNN fully connected layer can be expressed as

\[
\alpha^{\text{out}}_{n} = f \left( \sum_{m} \left( \alpha^{\text{out}-1}_{m} \times w^\text{out}_{mn} \right) + b^\text{out}_n \right)
\]

After combining all convolutional layers and pooling layers, the pre-training of SAE and MCNN’s supervisory training can achieve the optimal weight along with improved training speed. We present the parameter optimization for MCNN in Algorithm 1.

Given an input image or a patch, it is input to each channel of the MCNN for training. After training, the corresponding features are also represented by the feature vector, successively. The use of two-layer full connection reduces the characteristic dimension of training. Although the feature dimension is lower, the more texture information of fabric images is retained for further detection and identification.

**LRD model and optimal solution**

Given a test fabric image, it can be divided into image blocks \( \{B_i\}_{i=1, \ldots, N} \) with different sizes, where \( N \) is the number of image blocks. The characteristic representation \( f_i \in IR^D \) of each block is generated based on the MCNN. Then, we combine all the feature vectors \( f_i \) into a feature matrix to represent the image, which can be expressed as \( X = [f_1, f_2, \ldots, f_N] \in IR^{D \times N} \), where \( D \) is the dimension of the feature matrix and \( f_i \) represents the feature vector of the \( i \)th block. The feature matrix \( X \) can be decomposed into a low-rank matrix (corresponding to the background) and a sparse matrix (corresponding to the defective regions) by
utilizing the LRD model. In general, we require the rank of the decomposed low-rank matrix and the sparse part to be small. It can be described as follows

$$\min \left( \text{rank}(L) + \lambda \|S\|_0 \right)$$

subject to

$$X = XL + S$$

where $\text{rank}(L)$ is the rank of matrix $L \in \mathbb{R}^{D \times N}$ and $\|S\|_0$ is the $l_0$-norm (number of non-zero elements) of matrix $S \in \mathbb{R}^{D \times N}$. Here, the first term is to guarantee the low-rank matrix; the second term corresponds to sparsity; and the parameter $\lambda > 0$ is used to balance the effects of the low rank and sparsity, which could be chosen according to the properties of the two norms or tuned empirically. Due to non-smoothness and non-convexity, the above objective optimization problem is a tough task and comes with no closed-form solutions. To alleviate the difficulty, we can replace $\text{rank}(L)$ with $\|L\|_*$ and $\|S\|_0$ with $\|S\|_{2,1}$, resulting in the following convex optimization problem

$$\min \left( \|L\|_* + \lambda \|S\|_{2,1} \right)$$

subject to

$$X = XL + S$$

where $\|L\|_*$ is the nuclear norm of the low-rank matrix $L$ and $\|S\|_{2,1} = \sum_{j=1}^{N} \left( \sum_{i=1}^{D} (S_{ij})^2 \right)$ is the $l_{2,1}$-norm of the sparse matrix $S$. In this article, augment Lagrange multipliers (ALMs) are used to solve formula (12). Correspondingly, formula (12) is equivalently converted into the following Augmented Lagrangian function form

$$\min \left( \|L\|_* + \lambda \|S\|_{2,1} \right)$$

subject to

$$X = XL + S$$

where $Y_1$ and $Y_2$ are Lagrange multipliers and $\mu > 0$ is a penalty parameter. For the above problem, we choose the inexact ALMs, which we outline in Algorithm 2.

### Generation and segmentation of the saliency map

The generated feature matrix $X$ can be decomposed into low-rank matrix $L$ corresponding to the background and sparse matrix $S$ corresponding to the defect by formula (11). Each column $S_i$ of the sparse matrix $S$ corresponds to the possibility of an image block as a defect. In this article, the $l_{2,1}$-norm of $S_i$ is used to denote the prominence of image block $i$

$$\text{Sal}(B_i) = \|S_i\|_{2,1} = \sum_{j=1}^{N} \left( S_{ij} \right)^2$$

If $\|S_i\|_{2,1}$ is larger, the prominence of the image block $B_i$ becomes larger, which indicates the likelihood of the defect is greater. The prominence of all image blocks constitutes a saliency map $\text{SM}$. Noise reduction processing obtains $\hat{S}$ for the visual saliency map $\text{SM}$, that is

$$\hat{S} = g \times (\text{SM} \circ \text{SM})$$

where $g$ is a circular smoothing filter, “$\circ$” is the Hadamard inner product operator, and “$\ast$” represents the convolution operation. The saliency map $\hat{S}$ is converted into a gray-scale image $G$

$$G = \frac{\hat{s} - \min(s)}{\max(s) - \min(s)} \times 255$$

Finally, the iterative optimal threshold segmentation algorithm is adopted to segment the saliency map generated by the sparse matrix. Thus, the defective region is located.
Algorithm 2. Solving problem (12) by inexact ALMs

**Input:** Feature matrix $X$, parameter $\lambda$

**Initialize:** $L = J = 0$, $S = 0$, $Y_1 = 0$, $Y_2 = 0$, $\nu = 10^{-6}$, $\max_{\nu} = 10^{10}$, $\rho = 1.1$, $\epsilon = 10^{-9}$

**While** the formula is not converged, do

1. Fix the others and update $J$ by
   $$J = \arg \min \frac{1}{\nu} \| X - Y_1 \|_F^2 + \frac{1}{2} \| J - (L + Y_2) / \mu \|_F^2$$

2. Fix the others and update $L$ by
   $$L = (I + X^T X)^{-1} (X^T X - X^T S + J + (X^T Y_1 - Y_2) / \nu)$$

3. Fix the others and update $S$ by
   $$S = \arg \min \frac{\lambda}{\mu} \| S \|_{S, 1} + \frac{1}{2} \| S - (X - XL + Y_1 / \mu) \|_F^2$$

4. Update the multipliers
   $$Y_1 = Y_1 + \mu (X - XL - S)$$
   $$Y_2 = Y_2 + \mu (L - J)$$

5. Update the parameter $\nu$ by $\mu = \min (\nu, \max_{\nu})$

6. Check the convergence conditions
   $$\| X - XL - S \|_F < \epsilon \text{ and } \| L - J \|_F < \epsilon$$

**end while**

**Experimental results and analysis**

In the current study, numerous simulation experiments were conducted on the fabric images to evaluate the performance of the proposed algorithm. We selected several kinds of fabric defect images from two fabric image databases (including broken end, netting multiple, hole, thick bar, thin bar, dot-patterned, and star-patterned, etc.). The first is the fabric database of the University of Hong Kong, and the second is the TILDA fabric image database of the University of Freiburg. All of the algorithms were coded and executed in MATLAB R2016a, and all of the experiments were implemented on a 3.30 GHz, Intel (R) Core (TM) i3-2120 CPU PC.

To reduce the computation cost of the proposed algorithm, a whole fabric image was first down-sampled to 256 × 256 pixels and then fed into the network as the input in the experiments. The training was carried out, and features were extracted for the image block by MCNN. The multi-layer feature was fused to realize the representation of the fabric image, and the feature matrix was then composed. The low-rank matrix and the sparse matrix were decomposed by the ALM. Finally, the iterative optimal threshold segmentation algorithm was used to segment the saliency map generated by the sparse matrix. Thus, the defective area of the fabric is located.

To demonstrate the effectiveness of the adopted feature extraction method, we first compare the other four feature extraction methods with our adopted MCNN method: (1) Shen and Wu used intensity, texture, and orientation features plus low-rank matrix recovery (LRR) to detect an object; (2) Li et al. adopted an HOG feature and an LRD algorithm to detect fabric defects; (3) Li et al. applied the hierarchical information processing of the biological visual system and established BVM to characterize all types of fabric textures, which can quickly locate the salient object; and (4) Zhang et al. proposed a Gabor feature and LRD to detect defects. Figures 3 and 4 indicate the detection results of these different feature extraction methods in a plain-texture fabric and patterned fabric defect detection. Figure 3 demonstrates the detection result of the plain-texture fabric, and Figure 4 illustrates the detection result of a patterned fabric. Figures 3(a) and 4(a) are the original fabric images; (b) to (f) are some saliency maps generated by Shen and Wu, HOG + LRD, BVM + LRD, Gabor + LRD, and our method, respectively; (g) shows our segmentation maps. It can be clearly seen from the experimental results that the five feature extraction algorithms have good detection performance for plain and twill fabric, except that the defect area of the third column and the fourth column are not continuous. For the fabric defect images with a plain weave, the fabric defect area in Shen and Wu and Zhang et al. can be highlighted in the visual effect. For the patterned fabric, the background texture is very complex, and the effect of saliency maps generated by the method of Shen and Wu and Zhang et al. is not ideal. However, the detection performance of our method is significantly enhanced. The reason is that MCNN can be more effective in extracting the texture information of fabric images compared with other feature extraction algorithms, and the LRD model was integrated into the proposed algorithm. The fabric defects were accurately segmented, and the defect area was more prominent and was located effectively.

Figures 3 and 4 show the saliency maps generated by different feature extraction methods in plain-texture fabric and patterned fabric defect detection. The first column indicates the original images; the second column indicates the saliency maps generated by Shen and Wu; the third column indicates the saliency maps generated by the HOG feature with the LRD model; the fourth column is the saliency maps generated by the BVM with the LRD; the fifth column is the saliency maps generated by the Gabor feature with the LRD model; and the last two columns indicate the saliency maps and segmentation maps generated by our algorithm, respectively.

In the MCNN network, we added batch normalization before each activation layer. The batch normalization stores the running average of the mean and standard
deviation from its inputs. The stored mean is subtracted from each input of the layer, and division by the standard deviation is also performed. It has been demonstrated that by applying batch normalization, overfitting is decreased, and higher learning rates for training are achieved. In addition, we applied the padding strategy that pads zeros around the borders of the feature maps after each convolution leading to unchanged output shape in each channel. In addition, the pooling strategy adopted in all of the pooling layers is max-pooling, which is robust to small distortions. The output of the max-pooling creates a new set of image channels that are then fed through another layer of convolutions and nonlinearities. Finally, the last fully connected layer generates the output vectors, which are then stacked to a feature matrix. The parameters of the network, such as weights of the convolutional filters, are optimized through back propagation.

The selection of a stride and convolutional kernel size are critical for training network parameters and improving MCNNs’ performance. A larger value of the stride parameter reduces accuracy, although the training speed is improved. If the convolutional kernel is small, the local feature of the image cannot be extracted effectively. Conversely, if the convolutional kernel is too large, the calculation complexity will be far higher. The current study adopted multi-scale input images to increase local invariant information and collocation of different sampling intervals with different sizes of the convolutional kernel to obtain feature invariance, thus ensuring that the detailed information of the fabric image texture is not

Figure 3. Visual saliency maps of plain-weave fabric defect images: (a) original images, (b) Shen and Wu,37 (c) Li et al.,19 (d) Li et al.,4 (e) Zhang et al.,42 (f) ours, and (g) segmentation.
lost. We selected the cross-entropy function as the loss function of our proposed algorithm, and during the MCNN training process, the stochastic gradient descent (SGD) with mini-batches of 50 samples was applied to update the weight parameters. We also incorporated the momentum and learning rate decay into the SGD optimizer, and the updating rules of the weights in each iteration are as follows

\[
v_{i+1} = \mu \cdot v_i - a \cdot \nabla g
\]  

\[
w_{i+1} = w_i + v_{i+1}
\]  

\[
a \leftarrow a \cdot \frac{1}{1 + d \cdot i}
\]  

where \(i\) is the iteration index, \(w\) is the weight hyperparameter, \(\mu\) is the momentum coefficient, \(v\) is the current velocity vector, \(a\) is the learning rate, \(d\) is the decay parameter of the learning rate, and \(\nabla g\) is the average value of gradients with respect to \(w\) over the mini-batch at each iteration. In our experiments, the momentum parameter and the weight attenuation coefficient were set to 0.9 and \(10^{-5}\), respectively. The learning rate was initialized to 0.1 and then reduced by half the increase of the number of iterations. Apart from that, the dropout strategy with a probability of 0.5 was applied to the last fully connected layers that also help to avoid overfitting. In addition, we initialized the bias value to 0 for every layer and chose the ReLU activation function according to the work of He et al.\(^{44}\)

An important parameter in neural network training is the iteration number. If it is too small, the network prediction error increased. Furthermore, the input defect images cannot be fully learned, resulting in low detection accuracy of

Figure 4. Visual saliency maps of patterned fabric defect images: (a) original images, (b) Shen et al.,\(^{37}\) (c) Li et al.,\(^{19}\) (d) Li et al.,\(^{6}\) (e) Zhang et al.,\(^{42}\) (f) ours, and (g) segmentation.
the final multi-layer deep-learning algorithm; on the contrary, the long calculation time and the accuracy rate gradually increase with a higher number of iterations. Figure 5 illustrates the relationship between the number of iterations and both the loss function and the accuracy rate.

As can be seen in Figure 5, the accuracy rate is improved as the increase of iteration number and stability are achieved. In the experiments, the number of iterations was set to 3600 for a compromise of iteration times and accuracy. When training from scratch, the performance did not seem to improve beyond 3600 iterations, thus paused the back propagation there. Although the performance was much lower than for the fine-tuned networks, the network learned to predict the fabric defect.

However, the accuracy rate was not a good measure for imbalanced datasets as performance on the defect detection. Therefore, in the current study, we adopted means and standard deviations of average precisions, recalls, F-measure, and mean absolute error (MAE). The precision (also known as the positive predictive value or false positive rate, defined as TP/(TP + FP)) is the ratio of the correct significant region and the total detection area, which reflects the testing accuracy and measures the refused ability of the false detection rate in the test system error and test method. The recall (also known as sensitivity or true positive rate, defined as TP/(TP + FN)) is the probability of the correct significant area out of the total significant areas, which reflects the comprehensiveness of detection. Here, TP is the number of true positives, FP the number of true negatives, and FN the number of false negatives. The higher the precision and recall values, the better the detection effect, but there is a certain contradiction between the two. For example, when the detection area is very large, it can be used to ensure that the recall ratio is high, but the accuracy rate is not enough. To keep the balance between the two, it is necessary to adopt the F-measure value to synthesize the evaluation criteria of the two indexes. It is defined as

\[
F = \frac{(1 + \beta^2)P \times R}{\beta^2 P + R}
\]

where \( \beta^2 = 0.3 \), and according to Yang et al., it is a constant. Furthermore, MAE was also applied to evaluate the effectiveness of saliency detection of the proposed algorithm. The result is illustrated in Figure 6.

In Figure 6, the detection performance of our proposed algorithm is better than other existing methods, and the multi-layer features of MCNN for fabric defect description are more remarkable than some hand-crafted features.

**Conclusion**

Fabric defect detection is a key component of quality control in the textile industry. In the current study, we propose a novel fabric defect detection algorithm based on MCNNs and LRD. The proposed algorithm has two contributions: (1) the traditional methods of extracting features are replaced by MCNN to characterize a fabric defect image’s texture and (2) LRD is adopted to decompose the feature matrices in the low-rank and sparse parts. The detection results are obtained by segmenting the saliency map generated by the sparse part. The experimental results demonstrate that the proposed method can accurately detect the defect regions in various fabric defect images, even for images with complex textures, and it is superior to the state of the art.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China (Grant Nos 61772576 and 61379113), the Key Natural Science Foundation of Henan Province (Grant No. 162300410338), Science and Technology Innovation Talent Project of the Education Department of Henan Province (Grant No. 17HASTIT019), Intelligent Image Analysis Processing and Machine Vision Innovation Team in Henan Province (Grant No. 2018091), and the Henan Science Fund for Distinguished Young Scholars (Grant No. 184100510002).

ORCID iD
Baorui Wang https://orcid.org/0000-0002-1735-031X

References
1. Li Y and Zhang C. Automated vision system for fabric defect inspection using Gabor filters and PCNN. J Springerplus 2016; 5(1): 765.
2. Zhou J, Semenovich D, Sowmya A, et al. Dictionary learning framework for fabric defect detection. J Text Inst 2014; 105(105): 223–234.
3. Zhu Q, Wu M, Li J, et al. Fabric defect detection via small scale over-complete basis set. J Text Res J 2014; 84(15): 1634–1649.
4. Jung C and Kim C. A unified spectral-domain approach for saliency detection and its application to automatic object segmentation. IEEE T Image Process 2012; 21(3): 1272–1283.
5. Bo L, Ren X and Fox D. Unsupervised feature learning for RGB-D based object recognition. New York: Springer, 2013, pp. 387–402.
6. Li C, Gao G, Liu Z, et al. Fabric defect detection based on biological vision modeling. IEEE Access 2018; 6: 27659–27670.
7. Cheng D, Tian F, Liu L, et al. Image segmentation based on multi-region multi-scale local binary fitting and Kullback–Leibler divergence. J Signal Image Video Proc 2018(2): 1–9.
8. Xia D, Jiang G, Li Y, et al. Warp-knitted fabric defect segmentation based on non-subsampled Contourlet transform. J Text Inst 2017; 108(2): 239–245.
9. Jiang ZY and Hao FZ. Injured ticket detection and location based on improved template matching method. J Electr Design Eng 2018; 26: 175.
10. Wang G. Woven fabric defects based on singular value decomposition. Shanghai, China: Donghua University, 2014.
11. Zhu D, Pan R and Gao W. Fabric defect detection using characteristic spectrum of Fourier transform and correlation coefficient. J Comput Eng Appl 2014; 50(10): 866–873.
12. Guan S, Gao Z, Wu N, et al. Defect detection of plain weave based on visual saliency mechanism. J Text Res 2014; 35(4): 56–61.
13. Ngan HYT, Pang GKH, Yung SP, et al. Wavelet based methods on patterned fabric defect detection. J Pattern Recogn 2005; 38(4): 559–576.
14. Arivazhagan S, Ganesan L and Bama S. Fault segmentation in fabric images using Gabor wavelet transform. J Mach Vis Appl 2006; 16(6): 356–363.
15. Chunlei LI, Zhang Z, Liu Z, et al. A novel fabric defect detection algorithm based on textural differential visual saliency model. J Shandong Univ 2014; 44(4): 1–8.
16. Mak KL, Peng P and Yu KFC. Fabric defect detection using morphological filters. J Image Vis Comput 2009; 27(10): 1585–1592.
17. Tsai DM, Chen MC, Li WC, et al. A fast regularity measure for surface defect detection. J Mach Vis Appl 2012; 23(5): 869–886.
18. Lowe DG. Distinctive Image Features from Scale-Invariant Keypoints. J Int J Comput Vis 2004; 60(2): 91–110.
19. Li C, Gao G, Liu Z, et al. Fabric defect detection algorithm based on histogram of oriented gradient and low-rank decomposition. J Text Res 2017; 38: 153–158.
20. Long HX, Zhuo L, Jia-Feng LI, et al. A low-redundancy dense SIFT feature extraction algorithm. J Meas Control Technol 2017; 36(3): 20–23,27.
21. Li G and Yu Y. Deep contrast learning for salient object detection. Comput Vis Pattern Recogn 2016: 478–487.
22. Lu H and Zhang Q. Application of deep convolutional neural network in computer vision. J Data Acquisit Process 2016; 31(1): 1–17.
23. Hinton GE, Osindero S and Teh YW. A fast learning algorithm for deep belief nets. J Neural Comput 2006; 18: 1527–1554.
24. Abid S. Texture defect detection by using polynomial interpolation and multilayer perceptron. J Eng Fiber Fab 2019; 14: 1–12.
25. Ren R, Hung T and Tan KC. A generic deep-learning-based approach for automated surface inspection. IEEE Trans Cybernetics 2018; 48(3): 929–940.
26. Li G and Yu Y. Visual saliency based on multiscale deep features. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), Boston, MA, 7–12 June 2015, pp. 5455–5463. New York: IEEE.
27. Wang L, Lu H, Ruan X, et al. Deep networks for saliency detection via local estimation and global search. In: Proceedings of the IEEE conference on computer vision and pattern recognition, Boston, MA, 7–12 June 2015, pp. 3183–3192. New York: IEEE.
28. Weimer D, Scholz-Reiter B and Shpitalni M. Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection. CIRP Ann Manuf Technol 2016; 65(1): 417–420.
29. Li Y, Zhao W and Pan J. Deformable patterned fabric defect detection with fisher criterion-based deep learning. IEEE T Autom Sci Eng 2017; 14(2): 1256–1264.
30. Candès E, Li X and Ma J. Robust principal component analysis. J ACM 2011; 58(3): 11.
31. Yin M, Cai S and Gao J. Robust face recognition via double low-rank matrix recovery for feature extraction. In: Proceedings of the 2013 IEEE international conference on image processing, Melbourne, VIC, Australia, 15–18 September 2013, pp. 3770–3774. New York: IEEE.
32. Wan M, Qian W, Gu G, et al. Total variation regularization term-based low-rank and sparse matrix representation model for infrared moving target tracking. Remote Sens 2018; 10: 510.
33. Vaswani N, Bouwmans T, Javed S, et al. Robust subspace learning: robust PCA, robust subspace tracking and robust subspace recovery. *IEEE Signal Pr Mag* 2018; 35(4): 32–55.

34. Zhang Z and Zhao K. Low-rank matrix approximation with manifold regularization. *IEEE T Pattern Anal Mach Intel* 2013; 35(7): 1717–1729.

35. Bouwmans T, Sobral A, Javed S, et al. Decomposition into low-rank plus additive matrices for background/foreground separation: a review for a comparative evaluation with a large-scale dataset. *Comput Sci Rev* 2017; 23: 1–71.

36. Xue Y, Gu X and Cao X. Motion saliency detection using low-rank and sparse decomposition. In: *Proceedings of the IEEE international conference on acoustics, speech and signal processing (ICASSP)*, Kyoto, Japan, 25–30 March 2012.

37. Shen X and Wu Y. A unified approach to salient object detection via low rank matrix recovery. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, Providence, RI, 16–21 June 2012, pp. 853–860. New York: IEEE.

38. Peng H, Li B, Ji R, et al. Salient object detection via low-rank and structured sparse matrix decomposition. In: *Proceedings of the twenty-seventh AAAI conference on artificial intelligence, AAAI’13*, Bellevue, Washington, DC, pp. 796–802. Reston, VA: AAAI Press.

39. Yang C, Zhang L, Lu H, et al. Saliency detection via graph-based manifold ranking. *Comput Vis Found* 2013: 3166–3173.

40. Wang S and Huang A. Salient object detection with low-rank approximation and l2,1-norm minimization. *J Image Vision Comput* 2017; 57: 67–77.

41. Li C, Gao G, Liu Z, et al. Defect detection for patterned fabric images based on GHOG and low-rank decomposition. *IEEE Access* 2017; 7: 83962–83973.

42. Zhang D, Gao G and Li C. Fabric defect detection algorithm based on Gabor filter and low-rank decomposition. In: *Proceedings of the eighth international conference on digital image processing*, Chengdu, China, 20–23 May 2016, pp. 1–6. Bellingham: International Society for Optics and Photonics.

43. Ioffe S and Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv Preprint, arXiv:1502.03167v3.

44. He K, Zhang X, Ren S, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In: *Proceedings of the IEEE international conference on computer vision*, Santiago, Chile, 7–13 December 2015, pp. 1026–1034.

45. Yapi D, Allili MS and Baaziz N. Automatic fabric defect detection using learning-based local textural distributions in the contourlet domain. *IEEE T Autom Sci Eng* 2018; 15(3): 1014–1026.