Journal of Intelligent Manufacturing
https://doi.org/10.1007/s10845-020-01607-9

A multi-task and multi-scale convolutional neural network for automatic recognition of woven fabric pattern

Shuo Meng[1] · **Ruru Pan**[1] · **Weidong Gao**[1] · **Jian Zhou**[1] · **Jingan Wang**[1] · **Wentao He**[1]

Received: 20 August 2019 / Accepted: 14 June 2020

Abstract
The recognition of woven fabric pattern is a crucial task for mass manufacturing and quality control in the textile industry. Traditional methods based on image processing have some limitations on accuracy and stability. In this paper, an automatic method is proposed to jointly realize yarn location and weave pattern recognition. First, a new big fabric dataset is established by a portable wireless device. The dataset contains wide kinds of fabrics and detailed fabric structure parameters. Then, a novel multi-task and multi-scale convolutional neural network (MTMSnet) is proposed to predict the location maps of yarns and floats. By adopting the multi-task structure, the MTMSnet can better learn the related features between yarns and floats. Finally, the weave pattern and basic weave repeat are recognized by combining the yarn and float location maps. Extensive experimental results on various kinds of fabrics indicate that the proposed method achieves high accuracy and quality in weave pattern recognition.

Keywords Weave pattern recognition · Texture analysis · Computer vision · Multi-task learning · Convolutional neural network

Introduction

Woven fabrics are produced by interlacing two perpendicular sets of yarns: vertically passing warps and horizontally passing wefts. The cross states of warps and wefts are called floats. There are two types of floats: A warp float refers to a float with a warp passing above a weft, and a weft float denotes a float with a weft residing on top of a warp. The weave pattern is usually composed of the recurrence of the basic weave repeat (Schneider et al. 2015). The basic weave patterns are the plain, twill, and satin weave. Figure 1 shows a 2/2 left twill fabric image sample and its schematic diagram of the basic weave repeat. The weave pattern is an important parameter which makes the fabric not only strong and stable but also visually aesthetic. Conventional weave pattern recognition methods are mainly based on manual analysis, which is time-consuming and has high labor costs. Therefore, it is desirable to develop a high efficiency and robustness method for the automatic recognition of weave patterns.

The weave pattern is a relative abstract conception, which relies on the analysis of floats and the reasoning of the relationship between warps and wefts. However, the colors, diameters, and overriding relationships of yarns are quite diverse. Moreover, the automatic method should have high accuracy and generalization faced with various kinds of fabrics. All the above reasons make the problem a rather challenging task. In our previous work (Meng et al. 2019), we creatively explored a multi-scale convolutional neural network (MSnet) to measure the fabric density. The method adopted a portable device to capture fabric images and established a dataset. The MSnet can accurately locate yarns and outperforms the state-of-art methods. Despite the excellent

✉ Ruru Pan
prrsw@jiangnan.edu.cn

Shuo Meng
6180703014@stu.jiangnan.edu.cn

Weidong Gao
gaowd@jiangnan.edu.cn

Jian Zhou
jzhou@jiangnan.edu.cn

Jingan Wang
jingan.wang@vip.jiangnan.edu.cn

Wentao He
6180705044@stu.jiangnan.edu.cn

[1] Key Laboratory of Eco-Textiles, Ministry of Education, Jiangnan University, Wuxi 214122, Jiangsu, China

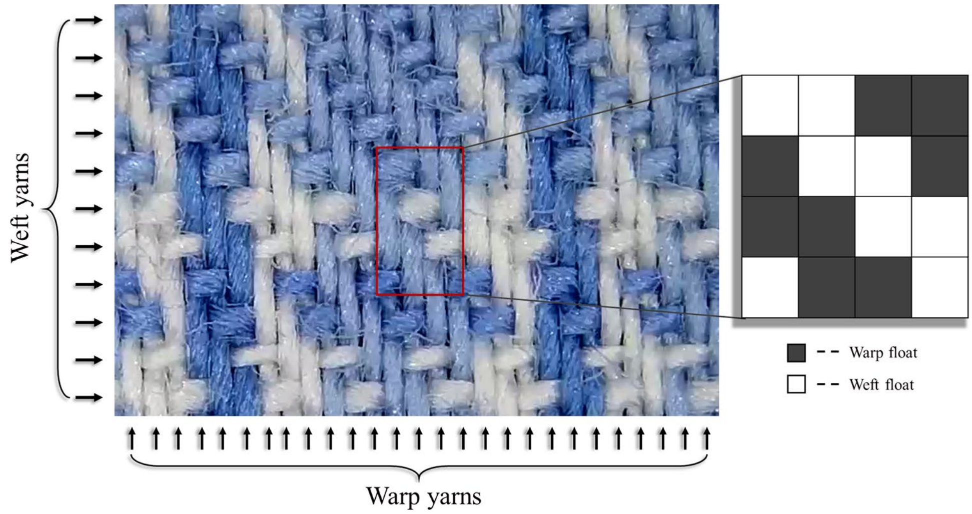

Fig. 1 A 2/2 left twill fabric image sample and its schematic diagram of the basic weave repeat

performance for fabric density measurement, using it alone does not enable to the recognition of the weave pattern.

In this paper, we develop a new multi-task and multi-scale convolutional neural network (MTMSnet) to learn the associativity of yarns and floats. The multi-task structure achieves a high performance that cannot be achieved by a single task. Figure 2 shows the flow chart of the proposed method. Besides, we further expand our fabric image dataset and label the information of floats. The establishment of this elaborated dataset allows us to train the network well and conduct a comprehensive evaluation.

The rest of the paper is organized as follows: First, some related works on the fabric weave pattern recognition are given in Sect. 2. In Sect. 3, the image acquisition system and the dataset establishment are briefly introduced. Section 4 provides a description of the MTMSnet and the following steps to recognize the weave pattern. In Sect. 5, the training and evaluation details are described. It is followed by the discussion of the model structure and comparisons of different methods in Sect. 6. Final Sect. 7 draws the conclusion of this paper.

Related works

With the development of computer vision and image processing algorithms, many scholars have committed to using image processing methods to automatically recognize the weave pattern (Xu 1996; Wang et al. 2010; Guo et al. 2019). In general, there are two ways to recognize the weave pattern: one is to classify the fabrics to basic weave patterns, which is considered as a classification problem; Another is to locate floats and then classify floats to realize weave pattern recognition.

The methods based on classification (Kinoshita et al. 1989; Li et al. 2013; Jing et al. 2014) extract the entire fabric image features and then use classification methods, like support vector machine (SVM) and probabilistic neural network (PNN), to classify the basic weave patterns. Their methods can classify the basic woven fabrics but cannot deal with unknown weave pattern fabrics which have limitations in generalization and varieties adaptability.

At present, the more widely used method is to locate the floats and then identify the float type. Due to the periodic nature of the weave pattern, some frequency domain methods (Lachkar et al. 2005; Shen et al. 2010) is adopted to locate the floats. However, the power spectrum is often noisy and hard to analysis due to the non-uniform distribution of yarns. The space domain methods mainly contain grey image projection (Wang et al. 2010) and gray line profile (Aldemir et al. 2018). Moreover, The Hough transform is utilized to detect the skew angle caused by the placement (Pan et al. 2010). Their methods can deal with some simple solid color fabrics but show low accuracy on complex color pattern fabrics.

Once accurately locating the floats, the next step can be seen as a binary classification problem. Kang et al. (1999) and Huang et al. (2000) utilized the geometrical shape of the floats to classify the floats, but the geometrical shapes are very similar in some fabrics which caused high misjudgments. Thus, some texture analysis methods are widely used to extract the features of floats such as GLCM (Wang et al. 2010), active grid model (AGM) (Xin et al. 2009), optical coherence tomography (OCT) (Sabuncu and Ã-zdemir 2015), and so on. Their methods require a high resolution of the fabric images, which causes the fabric image acquisition systems are expensive and inconvenient. After feature extraction of the floats, the floats are classified by the methods like Fuzzy C-means Clustering (FCM) (Wang et al. 2010; Schneider and Merhof 2015; Xiao et al. 2018), the BP neural network (Pan et al. 2011), and the pattern database (Pan et al. 2010). The classification methods rely heavily on

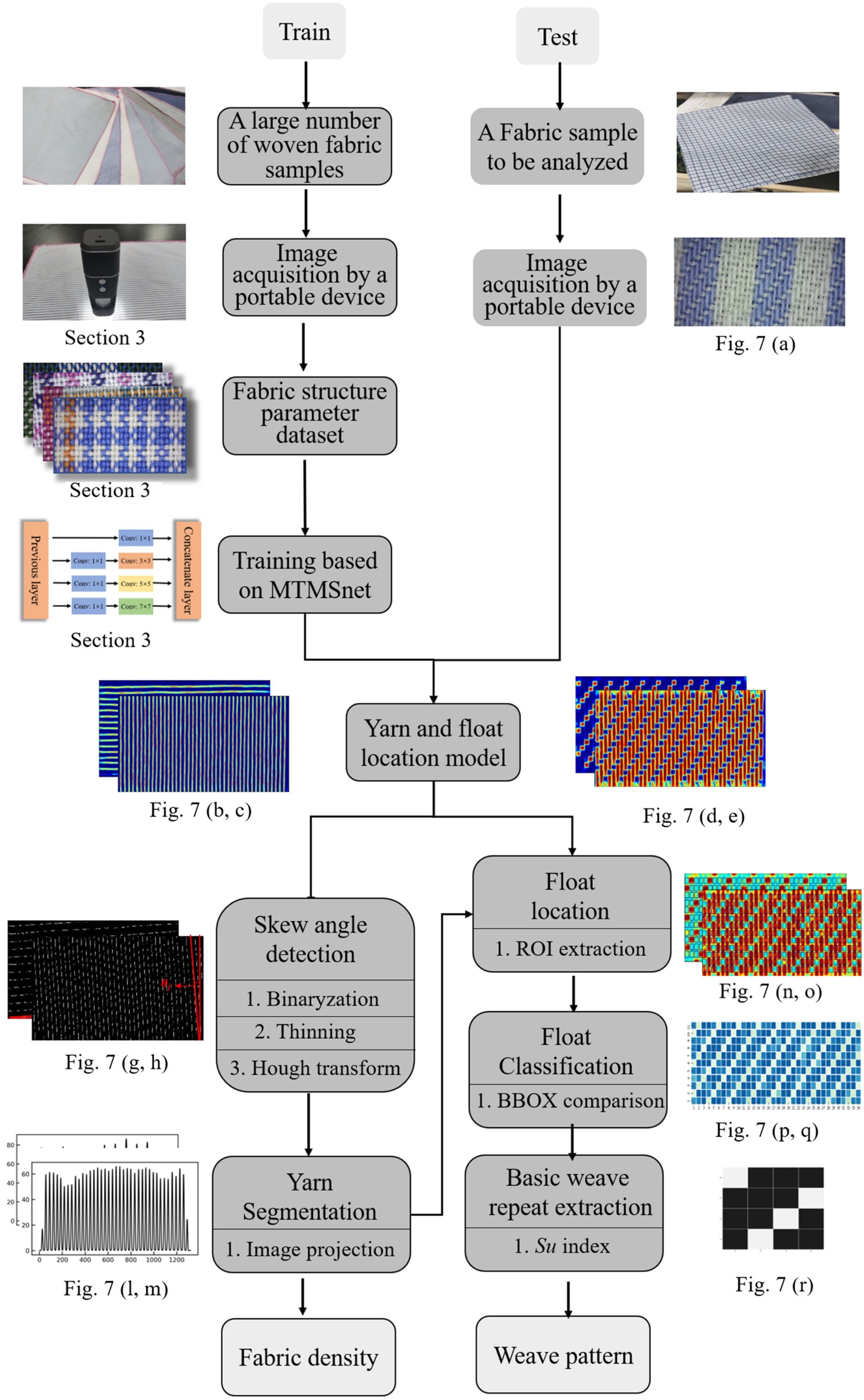

Fig. 2 The flow chart of the proposed method

feature extraction and some float types need the reasoning of the neighboring floats to decide, which cannot be easily classified by a single float.

In the past few years, owing to the excellent feature extraction capability, convolutional neural networks (CNN) are widely used to solve pattern recognition problems such as object classification (Malaca et al. 2019; Boonsirisumpun and Puarungroj 2018) and defect detection(Tabernik et al. 2020; Lin et al. 2019). Boonsirisumpun and Puarungroj (2018) and Xiao et al. (2018) used CNN to realize the automotive fabric pattern and knitted fabrics pattern classification respectively. They also converted the weave pattern recognition to a classification problem. For weave pattern recognition, it is more like a multi-task problem that needs to locate the floats and classify the float type. Sindagi and Patel (2017) designed a multi-task CNN to address crowd counting problems. Zhang et al. (2016) divided facial landmark problems into face classification, bounding box regression and facial landmark localization. The multi-task structure shows better performance when dealing with associative problems.

Although many traditional automatic methods based on image processing have been made some progress in the recognition of the weave pattern, they have some limitations on the efficiency and adaptability. In actual production, the portability and the application range of the acquisition equipment are highly required. To solve the problems in existing methods, we use a portable device and develop the MTMSnet to realize the recognition of the weave pattern.

As far as we know, there are still few studies based on CNN to address the weave pattern recognition problem. The paper has the following research contributions: (1) the fabric images are captured by a portable device, which extends the application range; (2) a more elaborated dataset with the information of yarns and floats is established; (3) Extensive experiments verify that the MTMSnet has high accuracy and robustness under various kinds of fabrics.

Image acquisition system and dataset establishment

Image acquisition system

A portable wireless device is used to conveniently acquire fabric images in sRGB mode with high-resolution. The device is equipped with a WIFI module and constant illumination. In each dealing process, it is only necessary to ensure that the fabric surface is flat and clean and the device is close to the surface of the fabric. The captured image is transferred wirelessly to a server and then the weave pattern can be recognized by the proposed method. The spatial resolution (PPI) of the image is fixed as 4680 pixel/inch with the size: 1280 pixels × 720 pixels.

Dataset establishment

In our initial exploration phase, we collected about 400 kinds of fabrics and established a fabric dataset of 600 images with only yarns information to measure the fabric density (Meng et al. 2019). In this paper, we expanded the dataset to 800 images and labeled the float information in addition. We only labeled a basic weave repeat to generate the whole float location to reduce the workload of labeling. The location of the basic weave repeat is randomly selected. If the fabric image does not have a periodic basic weave pattern, we labeled all floats in the image. Figure 3a shows a labeled image.

A combined strategy, which uses the smooth label to generate yarn location maps and the hard label to generate float location maps, is utilized to generate location maps as ground truth to train the model. For yarn location, we still use Gaussian distribution as a smooth label strategy to generate yarn location maps (Meng et al. 2019). The floats will be successfully located once the warps and wefts are segmented in the predicted yarn location maps. The next step is to classify the float into warp or weft float. Therefore, for float classification, we adopt 0 or 1 which is a hard label strategy to generate float location maps. Specifically, we generate two float location maps: warp float location maps and weft float location maps. If a pixel belongs to a float in

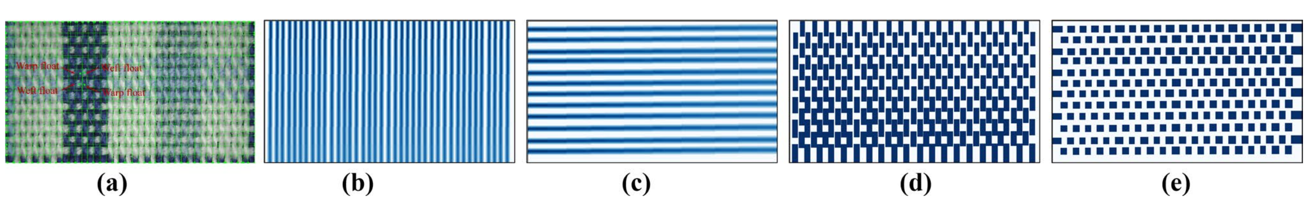

Fig. 3 A labeled image and its generated location maps. **a** a labeled image, **b** a warp yarn location map, **c** a weft yarn location map, **d** a warp float location map, **e** a weft float location map

the image, its value is 1 in the corresponding location map. As shown in Fig. 3b–e, four location maps are generated respectively.

The extended dataset covers a wide range of weave patterns and fabric types such as plain, twill, satin weave and some complex patterns. Figure 4 shows the distribution of the warps and wefts densities and fabric types. Meanwhile, we recorded the largest complete weave pattern in a fabric image and the basic weave pattern as ground truth for evaluation.

The proposed method

We improve the MSnet and design a novel multi-task and multi-scale convolution neural network (MTMSnet) to jointly realize yarn location and float classification. The flow chart of the proposed method is shown in Fig. 2. In this section, we introduce the structure of the MTMSnet and the processing steps based on the predicted location maps in detail.

The multi-task and multi-scale convolutional neural network

Inspired by the success of the CNN for resolving related multi-tasks (Sindagi and Patel 2017; Zhang et al. 2016; Dai et al. 2016), the recognition task is carried out as two related tasks: yarn location and float classification. Figure 5 illustrates the structure of the proposed MTMSnet. Two parallel stages are corresponding to the two tasks, with one stage learning warp and weft yarn location maps and the other stage generating warp and weft float location maps. Each stage contains two components: multi-scale feature encoder and location map decoder. The two stages share a set of convolutional features of the multi-scale feature encoder.

The structure of the encoder and decoder makes the network can address some general object location problems such as defect detection, face detection, and crowd counting. Owing to the multi-task structure, it is more suitable to extract related features like defect features or facial landmark.

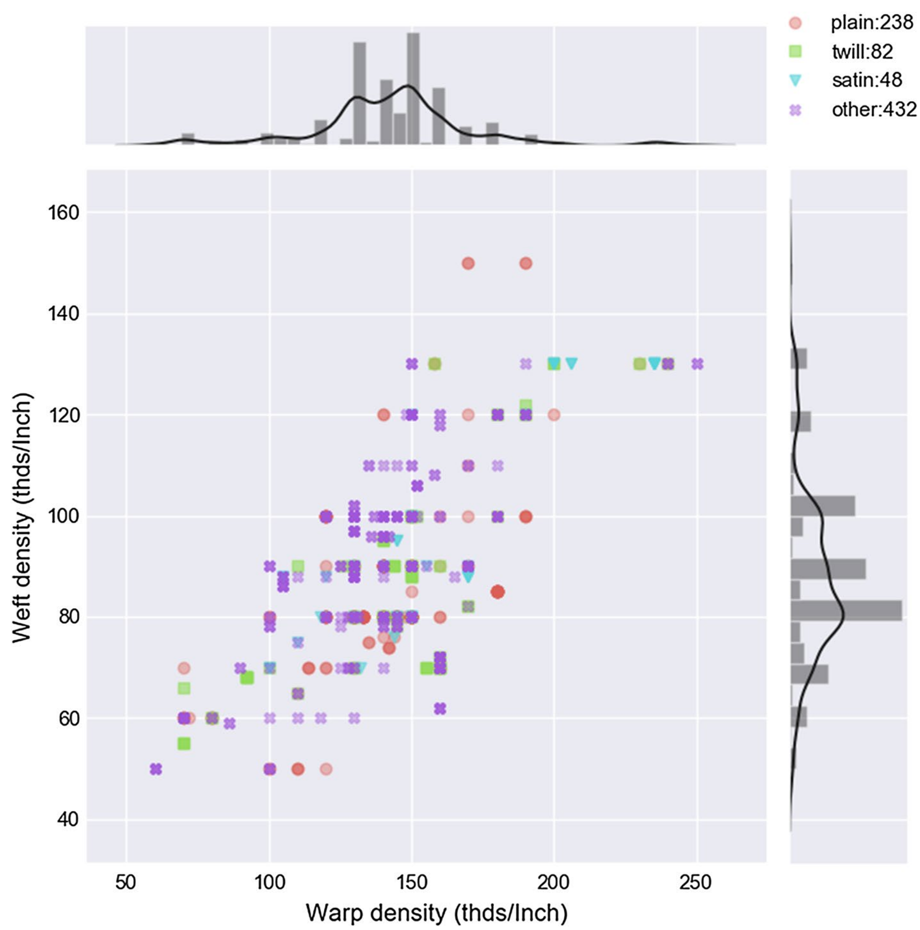

Fig. 4 The distribution of the warps and wefts densities and fabric types in the extended dataset

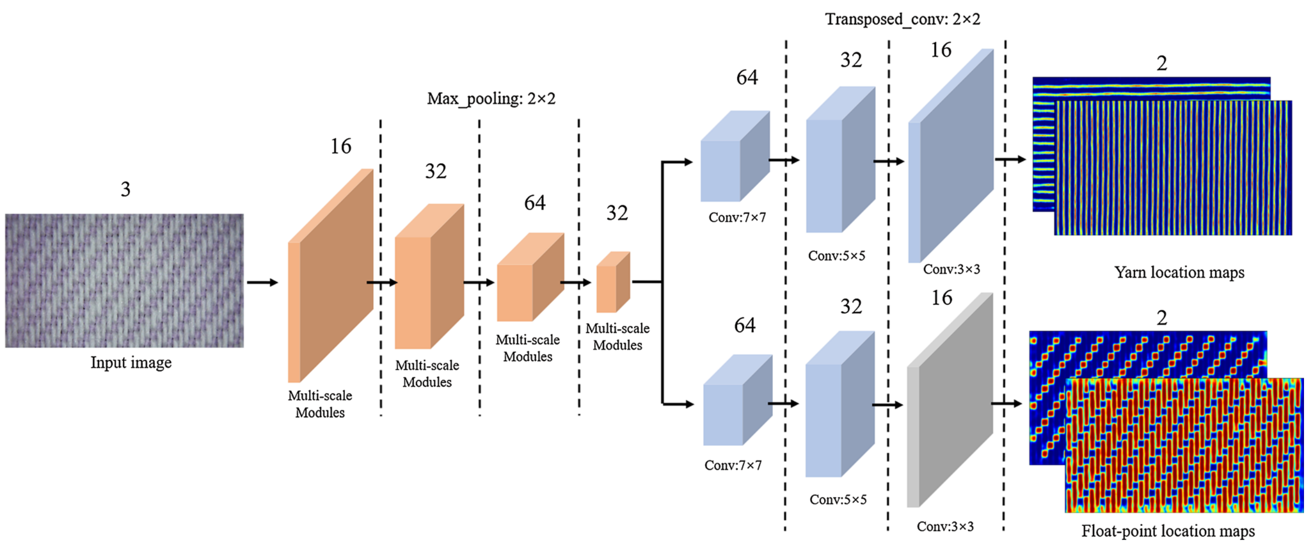

Fig. 5 The structure of the proposed MTMSnet

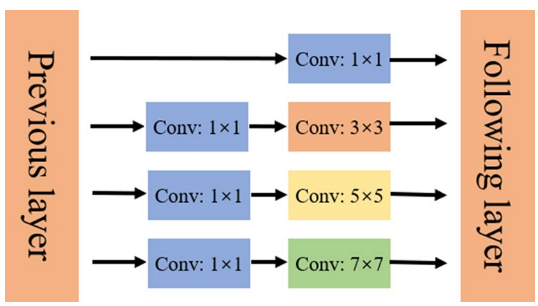

Fig. 6 The structure of the multi-scale module

Shared multi-scale feature encoder

The initial shared multi-scale feature encode*r contains four multi-scale modules to extract features of the raw input image. As shown in Fig. 6, the structure of the multi-scale module has four different sizes of filters which ensures the net has a more extensive local receptive field to address the problem of diverse yarn diameters and float sizes. To reduce the feature dimensions, a 1×1 filter is added before each filter. The feature maps of each module are 16, 32, 64, 32, respectively. Each interior layer has the same feature maps for concatenation. ReLU is adopted as the activation function after each layer. A 2×2 max-pooling layer is applied to eliminate most texture noises and retain robust features. The feature maps generated by this deep feature extract encoder are shared by the two decoders: yarn and float location map decoder. The presence of the shared convolutional features improves the attention filed of the net and reduces the number of parameters in some ways.

Yarn location map decoder

The feature maps obtained from the shared encoder are processed by yarn location map decoder to generate two location maps of warps and wefts. Four convolutional layers with the filter size of 7×7, 5×5, 3×3, 1×1 and the output channel of 64, 32, 16, 2 are adopted to refine the details of feature maps step by step. To ensure the outputs have the same size as the input image, three 2×2 transposed convolutional layers among the four convolutional layers are added. ReLU activation is applied in each layer to avoid gradient vanishing. Mean square error (MSE) and structural similarity index (SSIM) are used as loss function, which has been proved in our previous work (Meng et al. 2019). The definitions are as follows:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(G_i - Y_i\right)^2 \tag{1}$$

where N denotes total pixel number of an image, Y_i denotes the predicted value, the G_i denotes the ground truth.

$$\mathrm{SSIM} = \frac{(2\mu_Y\mu_G + C_1)(2\sigma_{YG} + C_2)}{(\mu_Y^2 + \mu_G^2 + C_1)(\sigma_Y^2 + \sigma_G^2 + C_2)} \tag{2}$$

$$\mu_Y = \frac{1}{N}\sum_{i=1}^{N} Y_i \tag{3}$$

$$\mu_G = \frac{1}{N}\sum_{i=1}^{N} G_i \tag{4}$$

$$\sigma_Y^2 = \frac{1}{N-1}\sum_{i=1}^{N} Y_i - \mu_Y \tag{5}$$

$$\sigma_G^2 = \frac{1}{N-1}\sum_{i=1}^{N} G_i - \mu_G \tag{6}$$

$$\sigma_{YG} = \frac{1}{N-1}\sum_{i=1}^{N} \left(Y_i - \mu_Y\right)\left(G_i - \mu_G\right) \tag{7}$$

where μ_Y and σ_Y^2 denote the local mean and variance of the predicted value, μ_G and σ_G^2 denote the local mean and variance of the ground truth, σ_{YG} denotes local covariance, C_1 and C_2 are small constants to avoid division by zero.

$$L = \mathrm{MSE} + \lambda(1 - \mathrm{SSIM}) \tag{8}$$

where L is the combined loss function, λ is a weight to balance the MSE and SSIM. In this paper, we set λ as 0.001 by experiments.

Float location map decoder

The structure of float location map decoder has minor differences with the yarn location map decoder. Because a hard label strategy is adopted to generate float location maps as the ground truth, the activation function of the output layer is set as Sigmoid and the loss function is set as binary cross entropy (BCE), which is defined as follows:

$$\mathrm{BCE} = -\frac{1}{N}\sum_{i=1}^{N}\left[G_i \ln Y_i + \left(1 - G_i\right)\ln(1 - Y_i)\right] \tag{9}$$

Using Sigmoid as activation function and BCE as loss function will accelerate the convergence and reduce the classification error for the binary classification problem (Golik et al. 2013).

Float location and classification

The network takes a fabric image of arbitrary size, and outputs two sets of the same size location maps: warp yarn and weft yarn location maps, warp float and weft float location maps. Figure 7a–f show a fabric image sample and its predicted location maps. As mentioned in our previous work (Meng et al. 2019), some of the predicted yarn location maps have a deficiency in the binding of warps and wefts due to warps and wefts masked by each other. We cannot easily locate the floats by summing the two warps and wefts location maps. Therefore, we introduce a new method of float location and classification based on the predicted location maps.

Skew angle detection

Due to the influence of the color pattern and the yarn texture, most of the existing methods need to tune parameters according to different fabric images and many outliers occur. However, the predicted yarn location maps have similar pattern features, which have already eliminated most of the noises and retained the information of the yarn location. Therefore, the skew angle of warps and wefts can be accurately detected based on the warps and wefts location maps. First, Otsu algorithm (Ohtsu 1979) is applied to convert the yarn location maps to binary images and Zhang–Suen thinning algorithm (Zhang and Suen 1984) is used to skeletonize the binary images for the purpose of saving computation time. Next, Hough transform (Duda and Hart 1972) is carried out to detect the skew angles of warps (θ_j) and wefts (θ_w) respectively. Figure 7g, h shows the skeletonized image and the detected skew angle. Finally, the predicted warp yarn location map and warp float location map are rotated $-\theta_j$, and the predicted weft yarn location map and weft float location map are rotated $-\theta_\mathrm{w}$. The image is cropped to retain a rectangle as the region of interest (ROI) which removes the incomplete yarn. Figure 7i–k show the ROI and the yarn location maps of the ROI. It should be noted that all the predicted location maps are rotated around the image center and the sizes of the final rotated images are the same. In this way, the relative location among the four predicted location maps has not been changed. More detailed comparisons between using the MTMSnet processing or not have been given in Sect. 6.6.

Yarn segmentation

The Image projection is adopted to locate the warp and weft yarns. The rotated warps location map is projected in the column and the rotated wefts location map is projected in the row to get the projection curves. To remove noises, a 5×1 minimum filter is used to smooth the curve. The final smoothed projection curves are shown in Fig. 7l, m The smoothed projection curves are almost noiseless, so by locating the local maxima (Pc) the fabric density (d) can be derived. Meantime, the two minima (Pl, Pr) on either side of a local maximum (Pc) represent the boundary of the yarn. The boundary box (BBOX) of each float can be finally obtained by combining the Pl and the Pr of each warp and weft. Figure 7n, o show the BBOX of each float in the warp float location map and weft float location map.

Float type classification

Once the BBOX of each float is got, the float type can be classified by calculating the sum values in the corresponding location of the float location maps. The float type is decided by follows:

$$F_\mathrm{p} = \frac{S_j}{S_j + S_w + C} \tag{10}$$

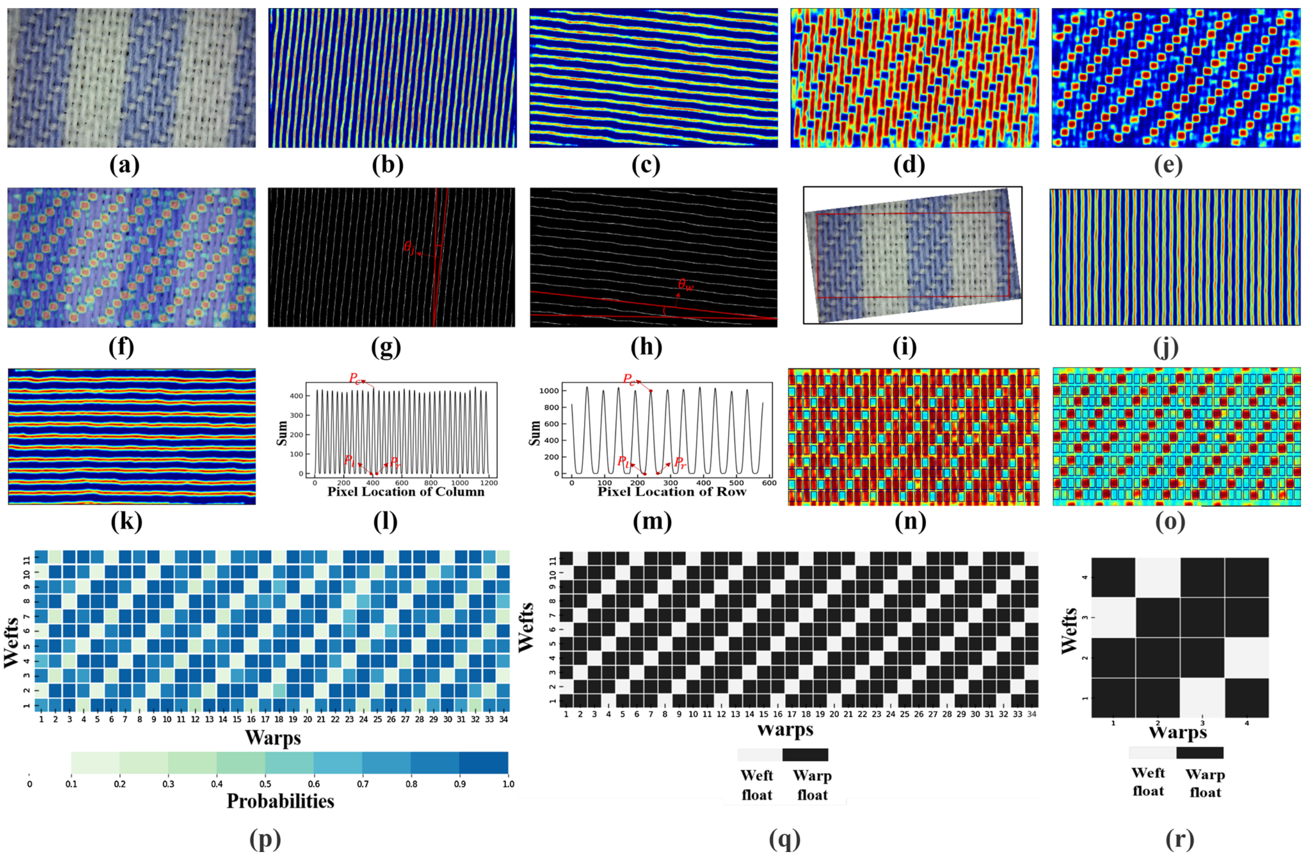

Fig. 7 Predicted location maps and processing images for the weave pattern recognition. **a** A test fabric image sample, **b** predicted warps location map, **c** the predicted wefts location map, **d** the predicted warp float location map, **e** the predicted weft float location map, **f** the mixed image of the weft float location map and the original image to show the predicted effect, **g** the skeletonized image of the warps location map, **h** the skeletonized image of the wefts location map, **i** the rotated original image and its ROI: the red rectangle region. **j** the rotated warps location map of ROI, **k** the rotated wefts location map of ROI, **l** the smoothed warp projection curve, **m** the smoothed weft projection curve **n** the rotated warp float location map of ROI and the BBOX of floats, **o** the rotated weft float location map of ROI and the BBOX of floats, **p** the heatmap for probabilities of the float as a warp float, **q** the diagram of the recognized weave pattern, **r** the diagram of the deduced basic weave repeat (Color figure online)

$$F = \begin{cases} \mathrm{WEFT}, F_{\mathrm{p}} < T \\ \mathrm{WARP}, F_{\mathrm{p}} \geq T \end{cases} \tag{11}$$

where F_{p} is the probability of the float is a warp float; S_j is the sum value within the float BBOX in the warp float location map; S_w is the sum value within the float BBOX in weft float location map; C is a small constant to avoid division by zero; F is the float type; T is a threshold and we set T as 0.5 in this paper. Figure 7p, q show the heatmap for probabilities of each float as a warp float and the schematic diagram of the recognized weave pattern.

Basic weave repeat recognition

Due to the number of yarns is relatively small, we calculate the size of basic weave repeat by the Su index which is proposed by Pan et al. (2010). The main principle is to calculate the probabilities of different repeat sizes from 2 to the number of yarns. Once we get the size of the basic weave repeat, the basic weave repeat can be extracted from randomly choosing the same size in the whole predicted weave pattern. Figure 7r shows the diagram of the deduced basic weave pattern.

Experiments

A series of experiments are conducted to evaluate the effectiveness of our proposed method. The hardware used in this study includes a sever with Intel(R) Core (TM) i9-7900x CPU, GTX 1080Ti GPU, and 32 GB RAM memory. The algorithm is implemented on the framework of Keras 2.2.4 with Tensorflow 1.13.0 as the backend.

Training details

Before training the MTMSnet, the dataset is randomly divided into three parts: 600 images are used for training, 100 images for validation, and the remaining 100 images for testing. The images in the training set and validation set are horizontal flipped, vertical flipped and shifted channel intensity for data augmentation. The final size of the train set and the validation set is 2400 and 400 images respectively. All the images are converted to [0, 1] for normalization. The parameters of the network are randomly initialized by Glorot uniform initializer (Glorot and Bengio 2010). Adam optimizer (Kingma and Ba 2014) with a small initial learning rate of 10^{-5} is used to train the model. At the same time, we monitor the validation loss and set patience of 20 epochs to reduce the learning rate. After 200 epochs of training, we obtain the final locating model.

Evaluation details

To evaluate the accuracy and robustness of the method, we calculate the fabric density and still utilize the mean-absolute-percentage error (MAPE) and the mean-squared-percentage error (MSPE) respectively for evaluation (Meng et al. 2019), which are defined as follows:

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|G_i - Y_i\right|}{G_i} \times 100\% \tag{12}$$

$$\mathrm{MSPE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{G_i - Y_i}{G_i}\right)^2} \times 100\% \tag{13}$$

where N is the number of test images, G_i is manual measurement value, Y_i is automatic measurement value.

For weave pattern recognition, the convention of existing works used recognition error (RE) (Guo et al. 2019; Kuo et al. 2016) and yarn location error (YLE) (Xin et al. 2009) to evaluate the model, which are calculated as follows:

$$\mathrm{RE} = \frac{N-R}{N} \times 100\% \tag{14}$$

$$\mathrm{YLE} = \frac{N-L}{N} \times 100\% \tag{15}$$

where N is the number of all test images, R is the number of correctly recognized images. L is the number of images for correctly locating the yarns. Because of the periodical nature of the weave pattern, the correctly recognized image means the recognized weave pattern is as same as the ground truth or the basic weave repeat is a subset of all possible basic weave repeats.

RE and YLE indicate the accuracy of the recognition based on the whole test set, which cannot illustrate the effect for every image. Therefore, we calculate the mean accuracy (MACC) based on an image, which is defined as follows:

$$\mathrm{MACC} = \frac{1}{N}\sum_{i=1}^{N}\frac{n_i - r_i}{n_i} \times 100\% \tag{16}$$

where N is the number of test images, n_i is the number of floats of the i image, r_i is the number of correctly recognized floats of the i image. If the model cannot locate all complete yarns in an image (i e, the size of predicted weave pattern is not same as the ground truth), we ignore the image to ensure the number of predicted floats is the same as the size of ground truth and use YLE to extra evaluate the model.

Because it is a binary classification problem for each float, the confusion matrix is introduced to further evaluate the robustness of the method. We plot the receiver operating characteristic curve (ROC) and calculate the area under the curve of ROC (AUC) to visually and comprehensively evaluate the model.

Results

Figure 8 shows some representative fabric samples and their predicted results. The results illustrate that the method can correctly recognize most of the weave patterns but with minor errors when deals with complex weave patterns such as jacquard fabrics. The estimated RE and YLE of the whole test set are 8% and is 6%. On the other hand, the proposed method can still realize accuracy fabric density measurement. The estimated MAPE of warps and wefts densities in the whole test set are 1.38% and 1.65% respectively, the MSPE are 2.28% and 2.43% respectively. All the above results approve that the proposed method can jointly realize weave pattern recognition and fabric density measurement with high accuracy and robustness.

In terms of computation time, the model takes about 4.62 s to deal with an image in the first loading. But it only takes about 0.59 s when we use the hot loading and the image of the weave pattern does not output. Moreover, because of using a portable device, the fabric density and weave pattern can be conveniently obtained within 10 s.

It should be noted that the MTMSnet do not rely heavily on the image acquisition device. As long as the yarns and floats in the image can be clearly distinguished by humans, the MTMSnet can achieve favorable results for the fabric density measurement and weave pattern recognition. In practice, the recommended range of the fabric density is 50 thds/inch to 250 thds/inch and the PPI of the fabric image is better between 3000 and 12000 pixel/inch. The fabric should not have large curved or overlapped yarns such as nap fabrics, two-layer fabrics and so on.

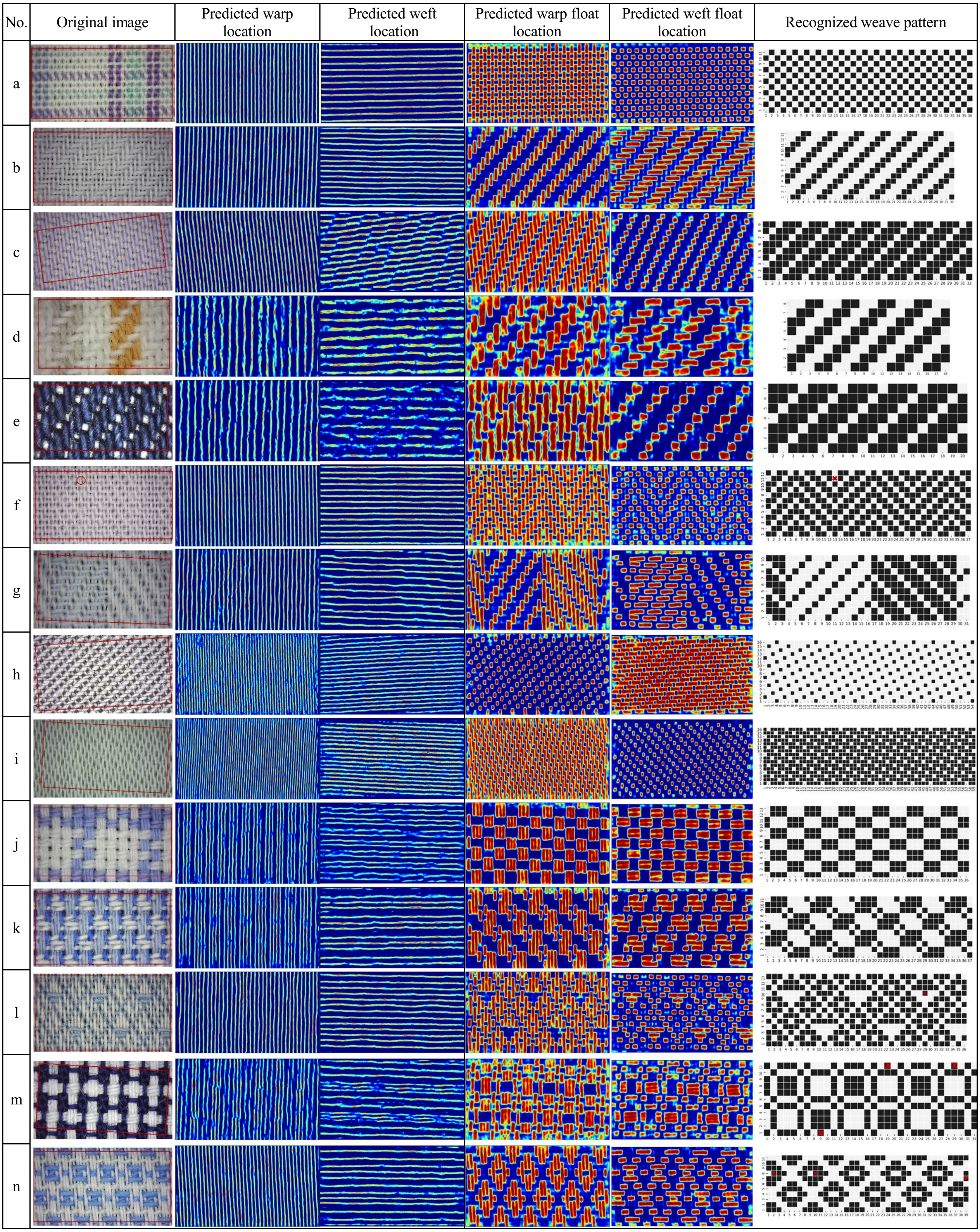

Fig. 8 Some representative fabric samples and their location maps in test set: the largest complete weave pattern is marked with a red rectangle and the misjudged floats are marked in red (Color figure online)

The model shows errors when deals with some more complex fabrics such as jacquard fabrics. Take Fig. 8n as an example. The float location and classification results are illustrated in Fig. 9. Most of the floats have been successfully located and classified but three floats are misjudged. Figure 9b, c show that the location of the misjudged floats has deviations, which affect the final classification results. The reason is that the Hough transform is used to calculate the average skew angle, but the skew angle of each yarn is not the same and some yarns are curved. In addition to these, even humans can hardly tell apart the float type in some images.

Discussion

Different loss functions and different strategies to transform labels to location maps

To improve the performance of the MTMSnet, different loss functions, activation functions, and strategies to transform labels to location maps were discussed. We used MSE + SSIM as loss function and ReLU as the activation, which has been verified in our previous work (Meng et al. 2019). Moreover, we also tried to use BCE as loss function and the Sigmoid as the activation of the output layer of the MTMSnet. Meanwhile, we adopted the hard and smooth label strategy respectively to generate float location maps for the aim of fitting with different loss functions. The hard label strategy has been described in Sect. 3.2. For the smooth label strategy, we assume that the pixel within the float obeys that the float center as the mean value and a scale parameter as the standard deviation. Therefore, the float can be represented as Gaussian kernels. We set as the center of the float and as 0.1 of the width or height of the float by experiments.

For qualitative analysis, Figs. 10a–d and 11a show the predicted location maps and the ROC of different loss functions and strategies respectively. Table 1 reports the results of the evaluation indexes. It can be seen that using the hard label to transform original labels to float location maps obtains lower error, and BCE is more suitable for this strategy. The reason is that using MSE + SSIM as loss function can accurately locate yarns, which is beneficial to measure fabric density. But for float classification, it is considered that using BCE as loss function and the Sigmoid as the activation is more suitable due to the effects of classification rely on the contrast between the two float location maps.

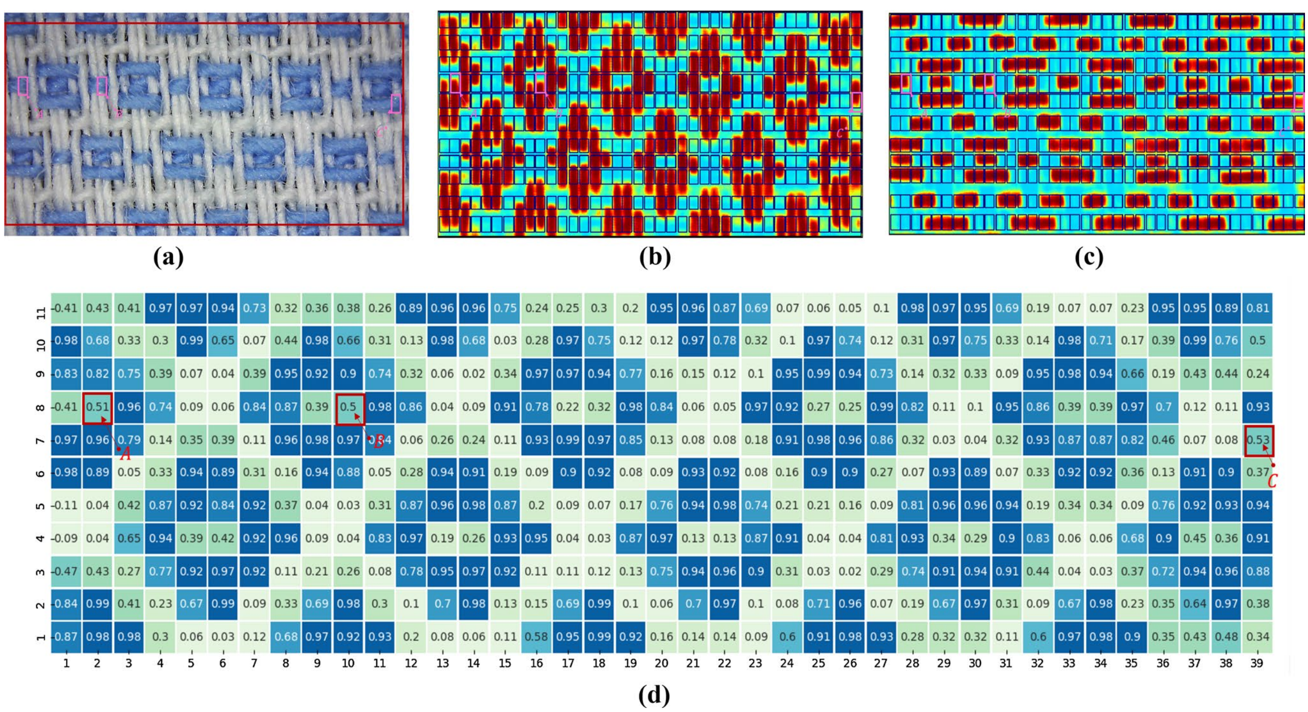

Fig. 9 The float location and classification results of Fig. 8n. **a** The original image marked with the ROI and the misjudged floats, **b** the rotated warp float location maps marked with float location and the misjudged floats, **c** the rotated weft float location maps marked with float location and the misjudged floats, and **d** the heatmap for probabilities of each float as a warp float

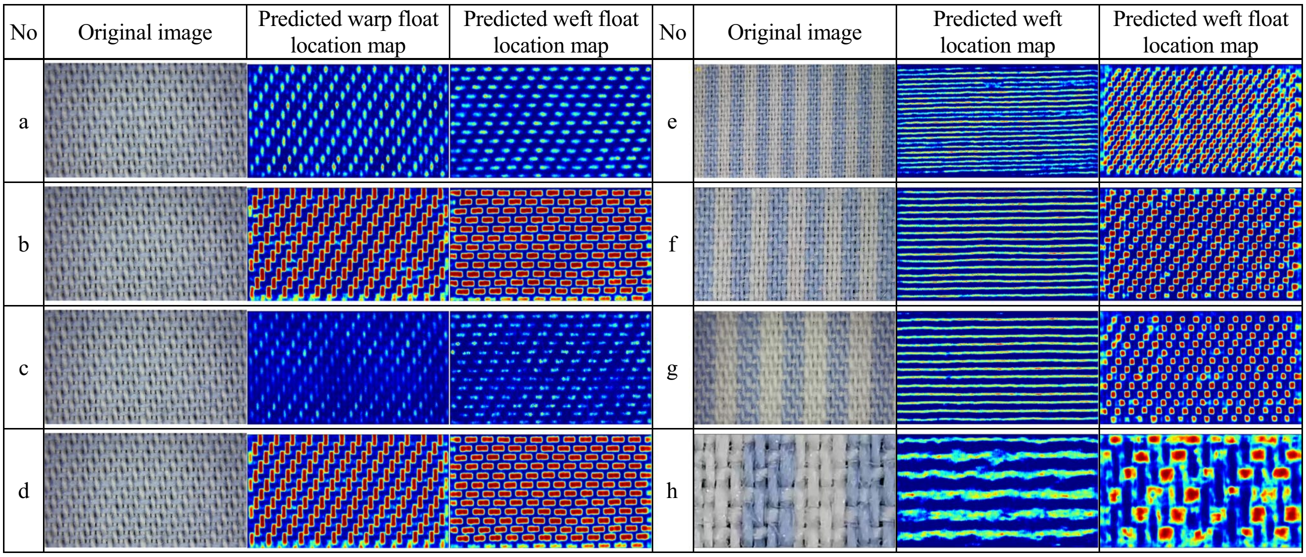

Fig. 10 Predicted yarn and float location maps with different loss functions, label strategies, and image resolutions: **a** the smooth label and MSE + SSIM as the loss function, **b** the hard label and MSE + SSIM as the loss function, **c** the smooth label and BCE as the loss function, **d** the hard label and BCE as the loss function, **e** 3078 PPI, **f** 4680 PPI, **g** 5925 PPI, **h** 12,312 PPI

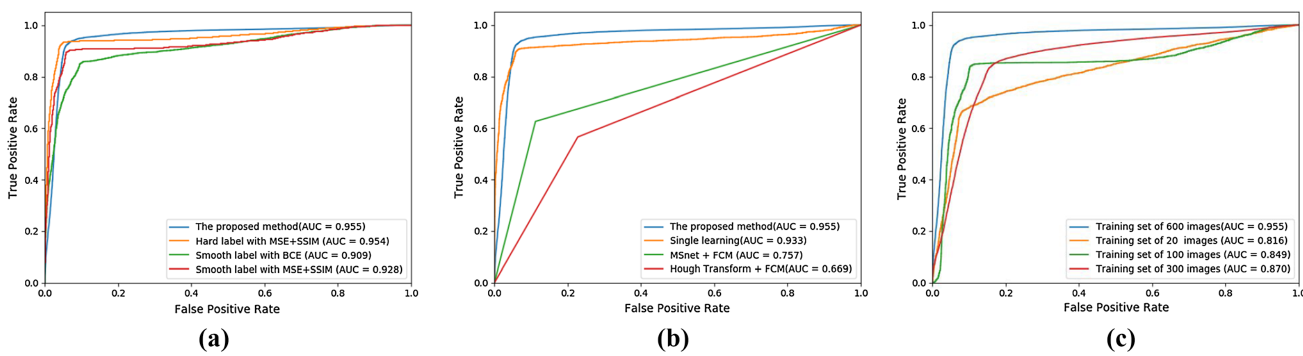

Fig. 11 The ROC of different methods for evaluation: **a** the ROC of different loss functions and strategies to transform labels to location maps, **b** the ROC of different methods, **c** the ROC of different dataset sizes

Table 1 Evaluation indexes of different loss functions and strategies to transform labels to location maps

Loss function	Label strategy	MACC (%)	RE (%)	YLE (%)
MSE + SSIM	Smooth label	90.16	11	**6**
MSE + SSIM	Hard label	92.88	**8**	**6**
BCE	Smooth label	89.50	12	**6**
BCE	Hard label	**92.97**	**8**	**6**

The boldface values refer to the best values

Multi-task structure

To evaluate the effectiveness of the multi-task structure, the MTMSnet was split into two deep neural networks which can be seen as two slightly different MSnet (Meng et al. 2019) (the difference is the loss function and the activation of the last layer). The training parameters of the MTMSnet are 2.1 million and the two divided nets are the same size of 1.6 million. We trained and evaluated the three networks based on the same dataset. The running of the two divided nets is separated due to the limitation of the finite computational ability of the hardware.

Table 2 reports the evaluation indexes. Figure 11b shows the ROC of joint learning (the proposed method) and the single learning (the combination of the two divided nets). The joint learning of yarn and float location reduces the error compared to the single-stage CNN, especially for locating yarns. The results for fabric density measurement is even better than the MSnet from the comparisons of the MAPE and MSPE. The main reason might be that the shared multi-scale encoder ensures the model can learn more general

Table 2 Evaluation indexes of the joint learning and the single learning

Structure	MAPE (%)	MSPE (%)	MACC (%)	RE (%)	YLE (%)	Processing time (s)
Single	1.56	2.60	91.84	8	7	6.23
Joint	**1.52**	**2.36**	**92.97**	**8**	**6**	**4.62**

The boldface values refer to the best values

The MAPE and MSPE are the mean value of the estimations of warps and wefts densities

Table 3 Evaluation indexes of different modifications of the structure

Structure	MAPE (%)	MSPE (%)	MACC (%)	RE (%)	YLE (%)	Parameters (million)
3×3 filter	1.96	2.72	90.53	12	8	**2.00**
9 multi-scale	1.61	2.363	**93.54**	**8**	**6**	3.39
MTMSnet	**1.52**	**2.355**	92.97	**8**	**6**	2.11

The boldface values refer to the best values

representations, which means the feature maps embedded with yarn and float information. Moreover, the cross states of warps and wefts are rightly the location of floats, which makes complements between the yarn location maps and float location maps in certain extent.

Multi-scale structure

We have explored some modifications of the structure of the MTMSnet to further enhance the recognition of the weave pattern. We modified the multi-scale module considering the diversity of the sizes of floats. In the multi-scale module, we added a parallel 1×1 and 9×9 filter to further extend the spatial awareness of the net. At the same time, we used the 3×3 filter to replace all of the filters of different sizes as a comparison.

Table 3 reports the evaluation indexes of different modifications of the multi-scale structure, which can be seen that the model obtains more promising results when using the multi-scale module. Although the extended multi-scale structure achieves higher MACC for weave pattern recognition, the overall results of the original multi-scale is better. The performance of the MTMSnet can be applied to the general weave pattern recognition.

Input image resolution

Image resolutions were discussed to evaluate the generalization and the application scope of the MTMSnet. In the experiments, we used different image resolutions of the same fabric for qualitative analysis. Figure 10e–h show the predicted weft yarn and float location maps under different PPI. The MTMSnet reveals certain adaptability under different PPI, but too small or large of PPI of the input image will result in the deficiency of the predicted location maps, which reduces the accuracy. The main reason is that the yarns are relatively dense or few beyond the receptive field of the net.

Table 4 Evaluation indexes of the MTMSnet trained with different dataset sizes

Dataset size	MAPE (%)	MSPE (%)	MACC (%)	RE (%)	YLE (%)
20	12.34	18.43	81.73	28	20
100	5.80	9.51	84.50	22	14
300	3.68	4.70	89.11	14	9
600	**1.52**	**2.36**	**92.97**	**8**	**6**

The boldface values refer to the best values

Dataset size

Conventionally, training a deep neural network needs the dataset large enough. Table 4 and Fig. 11c illustrate the evaluation indexes and the ROC with different sizes of the train set. The train sets are selected to cover most common fabric types but there is not a same image between the train set and the test set. Although the error on the test set is still high when trained with a small set, the model shows a certain ability to locate yarns and classify floats, owing to the multiple transposed convolution layers in the decoder stage and fabric images have some similar patterns. The results also illustrate that the larger the training set, the better the performance. So, it is necessary to establish a relatively large dataset to improve the robustness of the

Table 5 Evaluation indexes and processing times of different methods

Method	MACC (%)	RE (%)	YLE (%)	Processing time (s)
Hough Transform +Image projection + GLMC + PCA + FCM (Wang et al. 2010)	68.57	75	62	**1.75**
MSnet (Meng et al. 2019) + GLMC + PCA + FCM (Wang et al. 2010)	76.14	38	8	3.14
MTMSnet	**92.97**	**8**	**6**	4.62

The boldface values refer to the best values

The main parameter settings are as follows: (1) the step angle, min angle, and min distance of the Hough transform are 1°, 10°, 9 respectively; (2) the fuzzy weighting exponent and cluster number of FCM are set as 2, 2

model. Part of the dataset will be on publication to encourage the study of this problem.

Comparisons with different methods

We made comparisons between existing weave pattern recognition methods and the proposed method to further demonstrate the performance. Some methods have not been realized due to the requirement of the image acquisition system such as transmitting images (Li et al. 2019), dual-side captured images (Xin et al. 2009). In this paper, we realized the image projection to locate floats and the FCM to classify the float types which is widely used in the weave pattern recognition (Wang et al. 2010; Schneider and Merhof 2015; Xiao et al. 2018). The detailed steps can be seen in Table 5. Meantime, we used the MSnet to locate floats and the FCM to classify float types. Some parameters are set according to their study and our experiments to ensure a fair comparison. In the experiment, some fabrics are failed to be recognized and many outliers occurred in the test set, the reasons are that our dataset is more complex and the fine textures in the image are not clear enough. Figure 11b shows the ROC of different methods. The evaluation indexes and parameter settings can be observed in Table 5. The proposed method reaches higher accuracy compared with other methods. Moreover, using the MSnet to predict the location maps can highly improve the accuracy of the FCM to classify the floats.

Conclusion

Aiming at the recognition of the weave patterns of various kinds of fabrics, a novel MTMSnet is presented in this paper. The multi-task structure improves the ability to extract features from related tasks. The multi-scale structure enhances the adaptability of the network, which ensures a more extensive local receptive field to deal with different sizes of the interesting objects. Extensive experiments prove that: (1) the proposed method can jointly realize yarn location and weave pattern recognition with high accuracy, (2) the proposed method shows a better adaptability under a wide range of fabric densities and weave patterns, and (3) the potable image acquisition device makes it efficient and convenient for the recognition of the weave pattern.

Although the proposed method reaches high performance, it still has some limitations: (1) the establishment of the dataset is time-consuming, (2) the method cannot deal well with fabrics which contain large curved or overlapped yarns. In the future, the detailed directions of our research are as follows: (1) the further enhancement of the structure of the MTMSnet and new end-to-end methods will be discussed for the better recognition of the woven fabric pattern, (2) some new unsupervised models will be researched, (3) the automatic analysis of some other fabric structure parameters will be studied such as the layout of color yarns in yarn-dyed fabrics, and (4) an online web service will be established for automatic detection of fabric structure parameters.

Acknowledgements The authors are thankful to the National Natural Science Foundation of China under Grant 61976105, for providing financial support for this research work.

References

Aldemir, E., Özdemir, H., & Sarı, Z. (2018). An improved gray line profile method to inspect the warp–weft density of fabrics. *The Journal of The Textile Institute, 104,* 1–12.

Boonsirisumpun, N., & Puarungroj, W. (2018). Loei fabric weaving pattern recognition using deep neural network. In *2018 15th International joint conference on computer science and software engineering (JCSSE), 2018* (pp. 1–6). IEEE.

Dai, J., He, K., & Sun, J. (2016). Instance-aware semantic segmentation via multi-task network cascades. In *Proceedings of the IEEE conference on computer vision and pattern recognition, 2016* (pp. 3150–3158).

Duda, R. O., & Hart, P. E. (1972). Use of the hough transform to detect lines and curves in pictures. *Communications of the ACM, 15*(1), 11–15.
Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, 9,* 249–256.
Golik, P., Doetsch, P., & Ney, H. (2013). Cross-entropy vs. squared error training: A theoretical and experimental comparison. In: *Interspeech, 2013* (pp. 1756–1760).
Guo, Y., Ge, X., Yu, M., Yan, G., & Liu, Y. (2019). Automatic recognition method for the repeat size of a weave pattern on a woven fabric image. *Textile Research Journal, 89*(14), 2754–2775.
Huang, C.-C., Liu, S.-C., & Yu, W.-H. (2000). Woven fabric analysis by image processing: Part I: Identification of weave patterns. *Textile Research Journal, 70*(6), 481–485.
Jing, J., Xu, M., Li, P., Qi, L., & Liu, S. (2014). Automatic classification of woven fabric structure based on texture feature and PNN. *Fibers and Polymers, 15*(5), 1092–1098.
Kang, T. J., Kim, C. H., & Oh, K. W. (1999). Automatic recognition of fabric weave patterns by digital image analysis. *Textile Research Journal, 69*(2), 77–83.
Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:14091556.
Kinoshita, M., Hashimoto, Y., Akiyama, R., & Uchiyama, S. (1989). Determination of weave type in woven fabric by digital image processing. *Journal of the Textile Machinery Society of Japan, 35*(2), 1–4.
Kuo, C.-F. J., Shih, C.-Y., Huang, C.-C., Su, T.-L., & Liao, I.-C. (2016). A novel image processing technology for recognizing the weave of fabrics. *Textile Research Journal, 86*(3), 288–301.
Lachkar, A., Benslimane, R., D'orazio, L., & Martuscelli, E. (2005). Textile woven fabric recognition using Fourier image analysis techniques: Part II–texture analysis for crossed-states detection. *Journal of the Textile Institute, 96*(3), 179–183.
Li, Z., Meng, S., Wang, L., Zhang, N., & Gao, W. (2019). Intelligent recognition of the patterns of yarn-dyed fabric based on LSRT images. *Journal of Engineered Fibers and Fabrics, 14,* 1558925019840659.
Li, P. F., Wang, J., Zhang, H. H., & Jing, J. F. (2013). Automatic woven fabric classification based on support vector machine. In *International conference on automatic control and artificial intelligence, 2013* (pp. 581–584).
Lin, H., Li, B., Wang, X., Shu, Y., & Niu, S. (2019). Automated defect inspection of LED chip using deep convolutional neural network. *Journal of Intelligent Manufacturing, 30*(6), 2525–2534.
Malaca, P., Rocha, L. F., Gomes, D., Silva, J., & Veiga, G. (2019). Online inspection system based on machine learning techniques: real case study of fabric textures classification for the automotive industry. *Journal of Intelligent Manufacturing, 30*(1), 351–361.
Meng, S., Pan, R., Gao, W., Zhou, J., Wang, J., & He, W. (2019). Woven fabric density measurement by using multi-scale convolutional neural networks. *IEEE Access, 7,* 75810–75821.
Ohtsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems Man and Cybernetics, 9*(1), 62–66.
Pan, R., Gao, W., Liu, J., & Wang, H. (2010a). Automatic recognition of woven fabric patterns based on pattern database. *Fibers and Polymers, 11*(2), 303–308.
Pan, R., Gao, W., Liu, J., & Wang, H. (2010b). Automatic detection of the layout of color yarns for yarn-dyed fabric via a FCM algorithm. *Textile Research Journal, 80*(12), 1222–1231.
Pan, R., Gao, W., Liu, J., & Wang, H. (2011). Automatic recognition of woven fabric pattern based on image processing and BP neural network. *Journal of the Textile Institute Proceedings and Abstracts, 102*(1), 19–30.
Pan, R., Gao, W., Liu, J., Wang, H., & Zhang, X. (2010c). Automatic detection of structure parameters of yarn-dyed fabric. *Textile Research Journal, 80*(17), 1819–1832.
Sabuncu, M., & Ã-zdemir, H. (2015). Recognition of fabric weave patterns using optical coherence tomography. *Journal of the Textile Institute Proceedings and Abstracts, 107*(11), 1406–1411.
Schneider, D., Gloy, Y. S., & Merhof, D. (2015). Vision-based on-loom measurement of yarn densities in woven fabrics. *IEEE Transactions on Instrumentation and Measurement, 64*(4), 1063–1074.
Schneider, D., & Merhof, D. (2015). Blind weave detection for woven fabrics. *Pattern Analysis and Applications, 18*(3), 725–737.
Shen, J., Zou, X., Xu, F., & Xian, Z. (2010). Intelligent recognition of fabric weave patterns using texture orientation features. In *International conference on information computing and applications, 2010* (pp. 8–15). Berlin: Springer.
Sindagi, V. A., & Patel, V. M. (2017). CNN-based cascaded multi-task learning of high-level prior and density estimation for crowd counting. In *2017 14th IEEE international conference on advanced video and signal based surveillance (AVSS), 2017*.
Tabernik, D., Šela, S., Skvarč, J., & Skočaj, D. (2020). Segmentation-based deep-learning approach for surface-defect detection. *Journal of Intelligent Manufacturing, 31*(3), 759–776. https://doi.org/10.1007/s10845-019-01476-x.
Wang, X., Georganas, N. D., & Petriu, E. M. (2010a). Fabric texture analysis using computer vision techniques. *IEEE Transactions on Instrumentation and Measurement, 60*(1), 44–56.
Wang, X., Georganas, N. D., & Petriu, E. M. (2010). Automatic woven fabric structure identification by using principal component analysis and fuzzy clustering. In *2010 IEEE instrumentation and measurement technology conference proceedings, 2010* (pp. 590–595). IEEE.
Xiao, Z., Guo, Y., Geng, L., Wu, J., Zhang, F., Wang, W., et al. (2018a). Automatic recognition of woven fabric pattern based on TILT. *Mathematical Problems in Engineering, 2018,* 1–12.
Xiao, Z., Liu, X., Wu, J., Geng, L., Sun, Y., Zhang, F., et al. (2018b). Knitted fabric structure recognition based on deep learning. *The Journal of The Textile Institute, 109*(9), 1217–1223.
Xin, B., Hu, J., Baciu, G., & Yu, X. (2009). Investigation on the classification of weave pattern based on an active grid model. *Textile Research Journal, 79*(12), 1123–1134.
Xu, B. (1996). Identifying fabric structures with fast Fourier transform techniques. *Textile Research Journal, 66*(8), 496–506.
Zhang, T. Y., & Suen, C. Y. (1984). A fast parallel algorithm for thinning digital patterns. *Communications of the ACM, 27*(3), 236–239.
Zhang, K., Zhang, Z., Li, Z., & Yu, Q. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters, 23*(10), 1499–1503.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.