Neurocomputing 411 (2020) 112–127

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Detection of cervical cancer cells based on strong feature CNN-SVM network

A. Dongyao Jia, B. Zhengyi Li, C. Chuanwang Zhang *

Beijing Jiaotong University, School of Electronics and Information Engineering, No.3 Shangyuancun Haidian District, Beijing 100044, China

ARTICLE INFO

Article history:
Received 20 February 2019
Revised 27 May 2020
Accepted 2 June 2020
Available online 13 June 2020
Communicated by Dr. Nianyin Zeng

Keywords:
Cancer cell detection
Strong features
CNN-SVM
Sample amplification

ABSTRACT

Traditional screening of cervical cells largely depends on the experience of pathologists, which also has the problem of low accuracy and poor efficiency. Medical image processing combining deep learning and machine learning shows its superiority in the field of cell classification. A new framework based on strong feature Convolutional Neural Networks (CNN)-Support Vector Machine (SVM) model was proposed to accurately classify the cervical cells. A method fusing the strong features extracted by Gray-Level Co-occurrence Matrix (GLCM) and Gabor with abstract features from the hidden layers of CNN was conducted, meanwhile the fused ones were input into the SVM for classification. An effective dataset amplification method was designed to improve the robustness of the model. The proposed method was evaluated on two independent datasets with the metrics of accuracy (Acc), sensitivity (Sn), and specificity (Sp). Our approach outperformed than the state-of-the-art models with the Acc, Sn, and Sp of 99.3, 98.9, 99.4 for 2-class detection in the mass, respectively. The results indicated that the strong feature CNN-SVM model could be applied in cell classification for the early screening of cervical cancer.

1. Introduction

Cervical cancer is the second leading cause of death among malignancies, which has brought a fatal impact on women around the world [1]. Due to the harsh medical and living environment, 85 percent of the cervical diseases occurred in underdeveloped or developing countries [2]. Cervical cancer is usually early detected before deterioration because precancerous lesions can be found and treated by the Papanicolaou (Pap) test [3]. Therefore, regular early screening and treatment can effectively reduce the mortality of cervical cancer. Traditional cytology test, such as pap smear test requires cell-level examination using a microscope by pathologists for detecting nucleoplasm and karyoplasm variation, which brings burden to their work also causes errors to the test [4]. Computer-based systems considering various features of cells similarly to manual diagnosis are becoming supplementary approaches to effectively detect and analyze cell images. The task of cell classification for cervical cancer screening has become a key researching content in auxiliary diagnostic systems.

At present, many methods for cell classification have been proposed in the literature [5–7]. The mainstream cell classification methods in machine learning are generally based on feature extraction, which improved the accuracy of classification meanwhile reducing the computational complexity. Some commonly used features for cervical cells are nuclei and cytoplasm brightness, chromatin patterns and texture, combined morphology and color features, as applied in literature [8–10]. SVM is one of the most cited techniques for illness treating like cervical cancer [11]. For instance, Cheng et al. [12] applied SVMs with several filters and parameter sets for feature selection to detect cervical cells. Hyeon et al. [13] also introduced least square support vector machine (LSSVM) for softmax regression using the features extracted from CNN and produced good results, which is also one of the work for reference by the proposed model. An Radial Basis Function (RBF)-SVM also achieved the satisfying result and outperformed than methods of logistic regression and random forest by Bora et al [14]. Gupta et al. [15] proposed Artificial neural networks (ANN) on cervical cell classification and evaluated the model with an accuracy of 78.0%. GençTav A et al. [16] aimed at the problem of unbalanced medical dataset and applied an unsupervised approach for classification of cervical cells without any parameter adjustments. Y Marinakis et al. [17] combined particle swarm optimization with k-nearest neighbors' (k-NN) membership values and outperformed than other basic classifiers. Bora et al. [5] used an ensemble model with a weighted majority voting of other classifiers and obtained good performance. It is summarized that single machine learning model could achieve classification results while

* Corresponding author.
E-mail address: cwzhang1995@qq.com (C. Chuanwang Zhang).

https://doi.org/10.1016/j.neucom.2020.06.006
0925-2312/

as for more complicated datasets ensemble approaches usually behave better.

Deep learning based cell image analysis on cervical cancer has been generalized due to the development in image processing methods and computing power [18]. More deep learning technologies are also being applied to various fields related to medical treatment [19–21]. CNN as one of the deep learning approaches was widely applied for detection and classification of cervical cancer [22]. Such as Zhang et al. [18] applied a fine-tuning AlexNet to classify nuclei-centered image patches on two Pap smear datasets and obtained good results. Data augmentation was also employed due to the lack and unbalanced medical datasets [23]. Jith et al. [24] proposed a smaller model with only three of the AlexNet's convolution layers and obtained a more viable method for cell classification. An AlexNet approach was applied by Gautam et al. [25] and achieved a better accuracy in 2-class and 7-class problem. Some approaches [26–28] also introduced CNNs in feature extraction process to obtain the abstract characters of cells, then machine learning models like SVM are used for classification. These combinations of different models behaved better in classification task than single approaches. Although the outcomes of mentioned models achieved promising results, they still have some space for improvement on the existing accuracy, which actually motivated our work to establish the strong feature CNN-SVM model to efficiently detect cervical cells for classification.

A new cervical cell classification approach was proposed in this paper, which was based on strong feature extraction and CNN-SVM model. The workflow of the method is shown in Fig. 1. In our approach, a novel feature extraction method named GLCM+Gabor was first applied to extract the strong features, relevant experiments were conducted considering the accuracy to propose this method. Besides, the hidden layer of LeNet-5 model was used as the CNN feature extraction module to get the other abstract features, which are more scientific and not easy to cause the overfitting of the mode l [29]. Later, these two kinds of features are fused according to the optimal proportion from experiment after dimension reduction. Finally, the fused features were input into the SVM classifier for cell detection and classification. The proposed ensem-

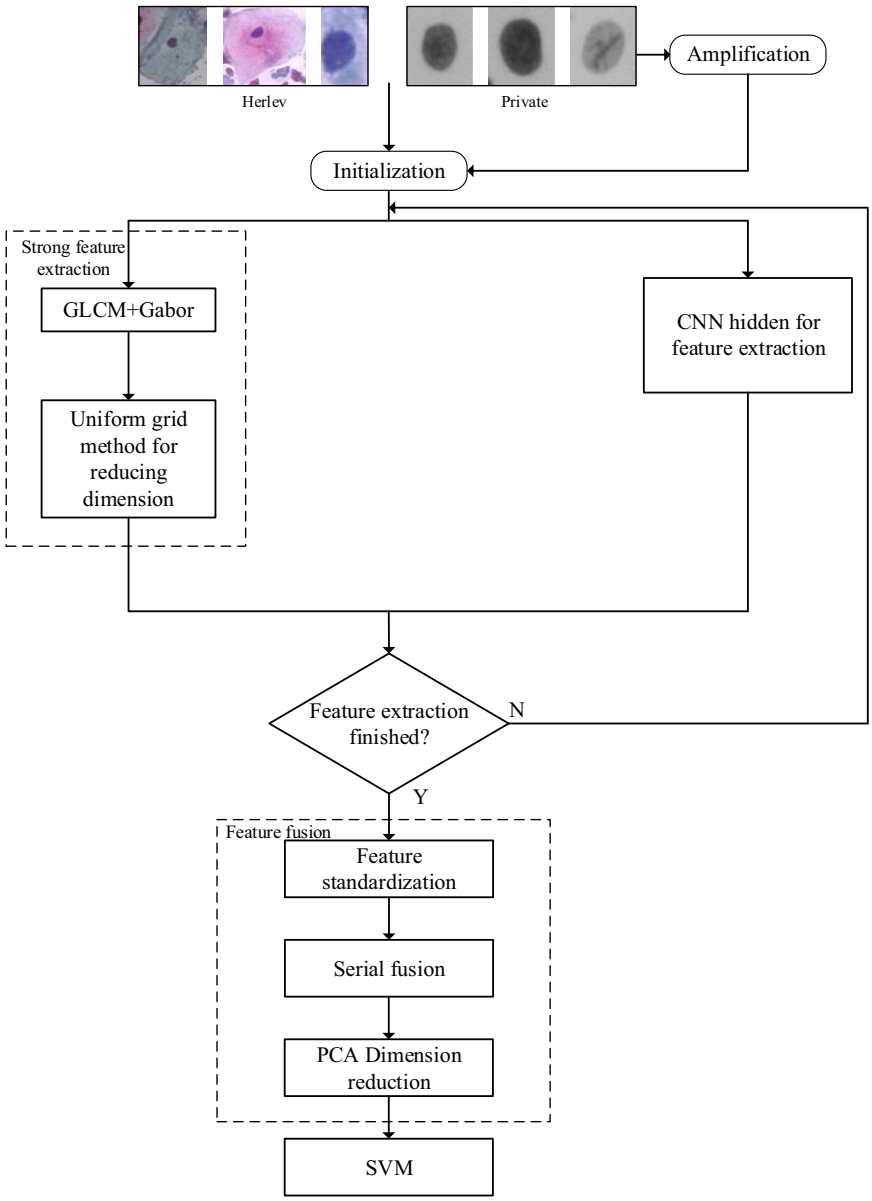

Fig. 1. Flow chart of the proposed classification algorithm based on strong feature CNN-SVM model.

ble model was more stable and behaved better in robustness [27,28]. Our approach was evaluated with the Herlev public dataset and a private dataset using data augmentation, and compared with baselines as well as some state-of-the-art models. Results showed that the proposed method outperformed than other related approaches in term of accuracy, sensitivity, and specificity. Therefore, the method in this paper will greatly reduce the work intensity and detection cost of cervical cancer, meanwhile improving the detection efficiency and the accuracy of screening, so that indirectly reduce the mortality of the patients with cervical cancer.

The rest of the paper is organized as follows. Experimental design is described in Section 2. Section 3 introduces the details of the proposed method. Section 4 talks about the experimental results and discussions. Conclusions are given in Section 5.

2. Experimental design

2.1. Data description

The proposed method in this paper was tested on two different datasets, including a publicly available dataset Herlev and one private dataset. The private dataset contained 2000 cell images from 200 patients who have conducted cervical cancer screening from August 2016 to December 2017 in Guangdong Province People's Hospital, which is our cooperative medical unit. All the data have been examined by the hospital ethics committee, whose sensitive information have been removed, including name, age, etc. And they will only be used for medical purpose with the requirement for informed consent waived.

50% of 200 cervical liquid-based cell slides in private dataset mentioned above are cervical cancer positive samples, the remaining 50% were negative samples. The final criterion for cervical cancer samples come from FIGO (International Federation of Gynecology and Obstetrics) and IGCS (International Gynecologic Cancer Society). Clinically, cervical cancer is classified into stages 0, Ⅰ, Ⅱ, III and IV according to the biopsy results. 40% of our cervical cancer positive samples are stage 0 and the rest of samples are stageⅠ. The main way of preventing cervical cancer is early screening, which is the reason to set the dataset like this. Instruments for private dataset establishment were color industrial camera (DFK33G274, Imaging Source) and microscope (DM3000, Leica), which made the total magnification of our image are 200 times. Through the self-assembled scanning system, 100 cell images can be obtained from each liquid-based cell slide, 10 typical cells were selected from each sample. The resolution of each cell image is 70 × 70. Pathological diagnosis has been confirmed by biopsy and case screening, cell images have been marked by four experienced doctors.

In addition to the private dataset, we also conducted model testing on the public dataset Herlev, which have 917 cells in total. Herlev images will be adjusted to the size of 70 × 70 in favor of being input into CNN before feature fusion. Model evaluation of 2 and 7 class was conducted respectively. Detailed descriptions of these two datasets are as shown in Table 1. Typical cells in two datasets are shown in Fig. 2, pictures of single position in slides are shown in Fig. 3. The upper parts in Fig. 3 are images of negative slides, and the lower parts are positive ones.

Table 1
Detailed description of the proposed 2 datasets.

Dataset	Private	Herlev	
(a) Stage 0 (positive)	331	(a) Superficial Squamous Epithelial	74
(b) Stage 1 (positive)	269	(b) Intermediate Squamous Epithelial	70
(c) Negative	1400	(c) Columnar Epithelial	98
		(d) Mild Dysplasia	182
		(e) Moderate Dysplasia	146
		(f) Severe Dysplasia	197
		(g) Carcinoma In Situ	150

2.2. Data preprocessing

In order to increase the robustness and the generalization ability of the model, it is often necessary to cover the sample space as much as possible [30]. As for the strong feature CNN-SVM network proposed in this paper, more effective data will bring more valid information for the model. However, it is difficult to ignore the problem of insufficient medical image data and unbalanced sample types. Aiming at this problem, an effective way is to create slightly changed copies using the existing data, which needs to follow the natural law in the process of image acquisition. For example, medical image data remains effective when the lens rotates, while we only get one sample at a time in the same position. The knowledge was summarized by human beings, while the model knows none about it. From this point of view, dataset expansion is also a method of manual supervision. The proposed data expansion method is only used for private dataset in order to facilitate the comparison of accuracy. Herlev dataset remained the same.

Dataset was mainly expanded by methods of rotation, scaling, random crop, and shear-warp transformation in this paper. It should be noted that the newly expanded data should keep their attribute labels unchanged. Technical details of the dataset expansion methods are as follows. Each specific way of dataset expansion is called as mode. RO (Rotation), SA (scaling), RA (random crop) and SH (shear-warp transformation) are listed according to the naming style. The optimal expansion method will be selected by conducting different dataset amplification methods and their combinations in experiment.

(1) Rotation

Rotating the original image in a clockwise direction of θ, θ∈(0°, 360°). As is shown in Fig. 4, the coordinates of point (x_0, y_0) turned to $(x_1 y_1)$ after the clockwise rotation of θ, θ = 90°, 180°, 270° are conducted to turn the original images. The dataset with an original capacity of N can be expanded to the number of $4 \times N$ after the image is rotated as is depicted in Fig. 5.

(2) Scaling

Image scaling simulates the imaging status of different multiple microscopes in operation, which is essentially to zoom the vector of each pixel in X and Y directions. Each cell image will be doubled and halved compared to its primary size, so that the private dataset can be expanded to three times of its original size, as shown in Fig. 6.

(3) Random crop

Random crop only improved the accuracy of the model, but also enhanced its stability [31], especially when the transformation range is appropriate [32]. The cropping ratio is scaled relatively to the horizontal and vertical direction, whose implementation is shown in Fig. 7.

Fig. 7 a) shows the original image that needs to be cropped. The small area in the dotted line frame in Fig. 7 b) is the value range of the upper left image after cropping. The point in this region is randomly selected so that the ratio of random crop is random, which can be achieved as in Fig. 7 c). We clipped the 1/3 position of the whole image to achieve the expansion of dataset in this way. Modes with random crop operations were recorded as RA1, those without were recorded as RA0.

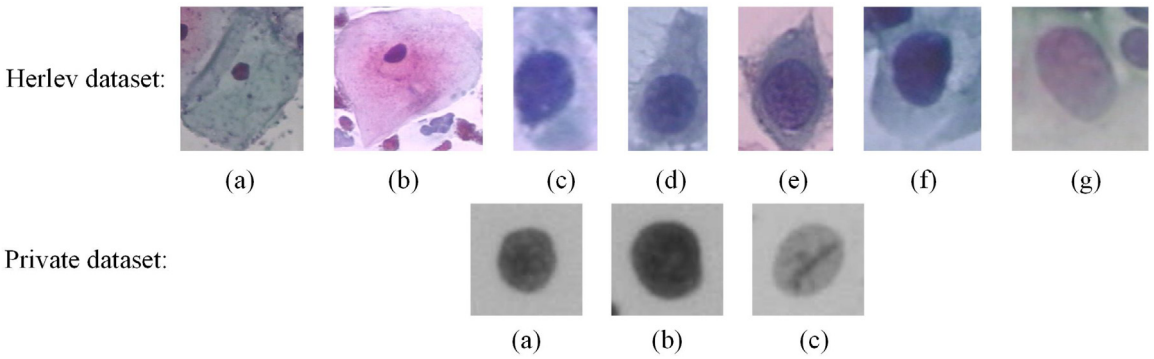

Fig. 2. Samples of Herlev & private dataset.

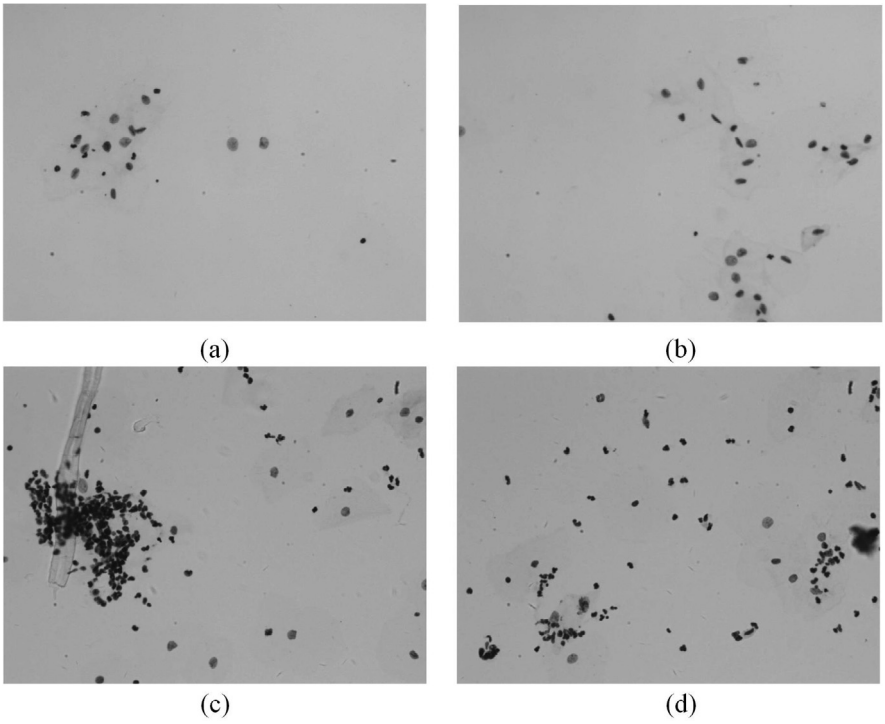

Fig. 3. Sample images of the single location in private dataset.

(4) Shear-warp transformation

Samples may be uneven or inclined due to the problem of production flow. For instance, the angle of the camera lens is not completely perpendicular to the slides. Shear-warp transformation is applied to simulate the similar distortion of images in X and Y direction meanwhile expanding datasets. Shear-warp transformation is calculated according to the mathematical matrices in Fig. 8.

The target pixel is calculated by the nearest-neighbor interpolation considering the influence on decimal part of pixels when coordinates change [33]. Angles of shear-warp transformation are horizontal (SH1), vertical (SH2) and 60° (SH3), the mode without shear-warp transformation was recorded as SH0. The diagram of shear-warp is shown as follows in Fig. 9.

2.3. Evaluation metrics

Medical image classification does not only require high accuracy, the sensitivity and specificity are also closely connected with the evaluation of models. Following evaluation criterions are given in order to calculate the metrics of different approaches.

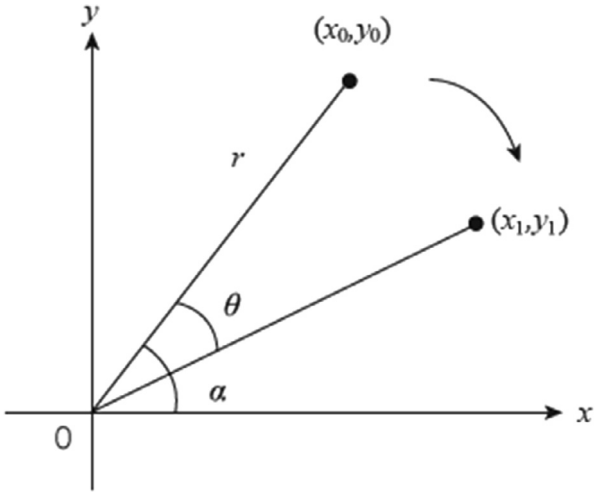

Fig. 4. Schematic diagram of image rotation.

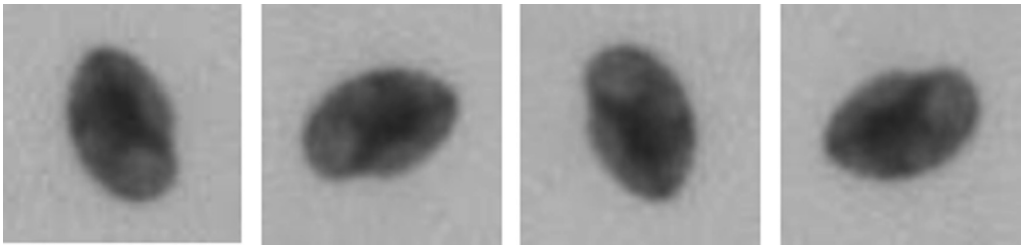

Fig. 5. Original (RO0) and θ = 90°, 180°, 270° (RO1, RO2, RO3) rotating mode of the image.

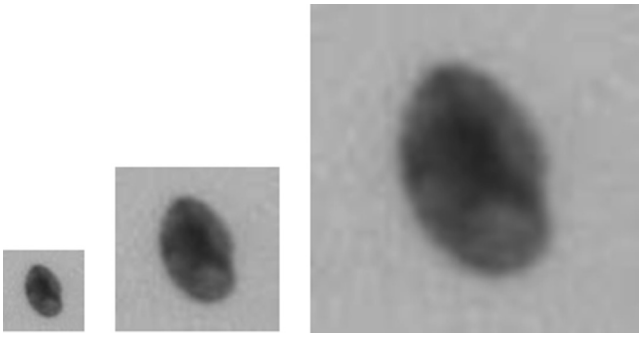

Fig. 6. Original cell image (SC0) and its half-doubled (SC1, SC2) conversion diagram.

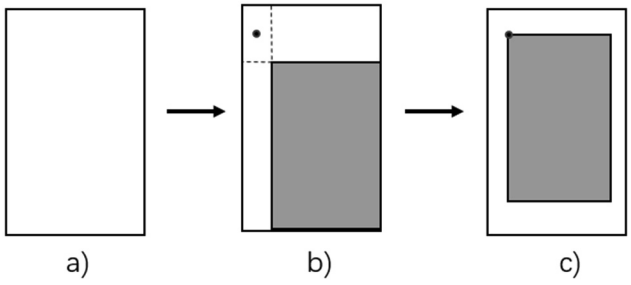

Fig. 7. Sketch map of random crop.

$$\begin{bmatrix} 1 & -tan\alpha \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ -tan\alpha & 1 \end{bmatrix}$$

Fig. 8. Matrix of X-direction and Y-direction shear-warp transformation.

The equations of accuracy, sensitivity and specificity is defined as follows.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

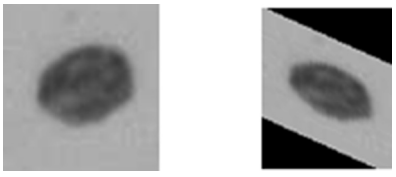

Fig. 9. Sketch map of shear-warp transformation.

$$Sn = \frac{TP}{TP + FN} \tag{2}$$

$$Sp = \frac{TN}{TN + FP} \tag{3}$$

The accuracy *Acc* refers to the proportion of the samples correctly classified in the total sample size, which represents the correctness of whole model; S_n is the percentage of true positive cases from all the positive ones, the higher sensitivity is, the lower probability of missed diagnosis will be; S_p is the percentage of samples correctly classified as negative from the total actual negative cases, the higher specificity is, the less chance of misdiagnosis there will be. In order to demonstrate the superiority of the proposed system, three indicators mentioned above are applied to evaluate the system performance.

3. Methods

The proposed approach consisted of three main parts, which are feature extraction, feature fusion and cell classification. The abstract features extracted from the hidden layer of LeNet-5 [34] are fused with the strong features after dimension reduction, then fused features are input into the SVM classifier for classification.

3.1. Feature extraction

3.1.1. Strong feature extraction

The changes of cell features in texture, morphology and chroma are very important for the detection of cancerization, which are called strong features in this paper. In order to extract the strong features of cell image and fuse them with the abstract features extracted by convolutional neural network, several common feature extraction methods were introduced to select the best strong feature extraction approach.

(1) Gray-Level Co-occurrence Matrix (GLCM)

GLCM is a statistical analysis method, which is achieved by calculating the gray level image, then some eigenvalues of the matrix are obtained by applying the co-occurrence matrix to represent the texture features of image respectively. GLCM can reflect the comprehensive information of image gray level about direction, adjacent interval, and variation range, which are the basis of analyzing the local patterns of images. Contrast, energy, entropy, correlation, and other features were applied to represent texture. The advantage of GLCM is that it is simple and easy to be implemented. And it specially suits images without obvious regularity, like the random patterns in cell images.

(2) Fourier transformation

Fourier spectrum contains abundant image information and can roughly describe texture patterns. The energy distribution of the

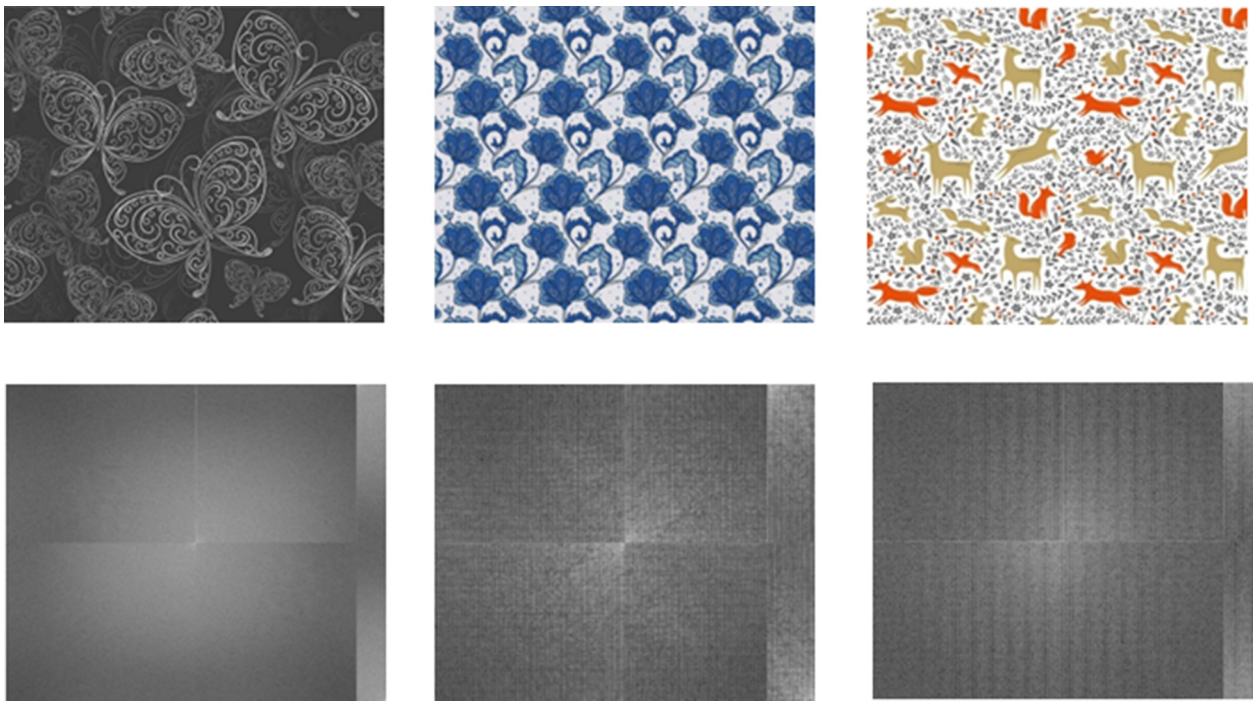

Fig. 10. Sample images and their Fourier spectrum.

image can be observed by the conduction on 2-dimensional Fourier transformation. If there are more dark points in the spectrum, it shows that the actual image is softer; while if there are more bright points, the actual image is sharper, which means the contrast is stronger and the boundary is clearer. The first line of Fig. 10 showed four texture images with obvious directivity, periodicity and randomness. The second line of Fig. 10 is the corresponding Fourier spectrum.

(3) Gabor transformation

Gabor transformation is very sensitive to the local space. When the standard Gauss kernel function is used, the rotation angle of the Gauss kernel function is the same as the angle of the sinusoidal plane wave, the Gabor function is simply described as follows.

$$\varphi(x,y,w,\theta)=\frac{1}{2\pi\sigma^2}\exp\left(-\left(\frac{x'^2y'^2}{2\sigma^2}\right)\right)\left[\exp\left(iwx'-\exp\left(-\frac{w^2\sigma^2}{2}\right)\right)\right] \tag{4}$$

$$\begin{cases} x'=x\cos\theta+y\sin\theta \\ y'=y\cos\theta+x\sin\theta \\ w=2\pi f \end{cases} \tag{5}$$

The feature extraction of Gabor is mainly reflected in the frequency scale f and the angle parameter θ, in which f represents the central frequency of the filter, and θ is its angle. Various filters can be achieved by combining different f and θ. θ is selected in eight directions 0, $\pi/8$, $\pi/4$, $3\pi/8$, $\pi/2$, $5\pi/8$, $3\pi/4$, $7\pi/8$ in this paper. f is selected in five scales such as $f=0,1,2,3,4$ to filter the image of cells. A total of 40 texture features can be obtained in eight directions and five scales, which also generated 40 filters. 40 Gabor feature maps in Fig. 11 are obtained after 40 kinds of filtering using the sample image. It is concluded that Gabor kernel can capture the spatial frequency of different directions in the local area as well as some structural features. The Gabor nucleus has strong ability of feature extraction and good spatial locality as well as direction selectivity.

(4) Markov random field

The feature extraction method based on Markov random field can extract feature points adaptively without manual parameter adjustment [35]. This algorithm is simple in calculation, relatively fast in speed and adaptive. The label set is selected by calculating the random field corresponding to each label adjustment. The optimal label set is found to extract the feature points until the energy of the whole system is stable.

3.1.2. Feature extraction using CNN

The hidden layer of LeNet-5 model is used to extract the CNN features in this paper. Fig. 12 is the schematic diagram of the network for extraction.

As is shown in Fig. 12, the first layer of the neural network is the input layer, the input image are epithelial cells with the size of 70×70. C1 layer of the network consisted of five convolution cores, the size of each group's visual field window was set to 5×5, the sliding window step was 1. C1 layer was obtained by convoluting visual window with the input image. Therefore, the C1 layer included five feature maps, the size of each feature map is $(70-5+1)\times(70-5+1)=66\times66$. S2 is the downsampling layer, which used the max-pooling for downsampling. It also applied the pooling window of the 2×2 size with a 2 steps length. Areas of 2×2 sizes in the original feature graph were selected without overlapping, the maximum value in the area is taken as the unique value for new feature map. Due to the mapping relation between convolution layer and subsampling layer, there are still five feature maps in the layer S2. And the dimension of each feature map is $(66/2)\times(66/2)=33\times33$. The feature map of layer S2 was convoluted by the visual field window with the size of 5×5, the step size of sliding window was also set to 1, the number

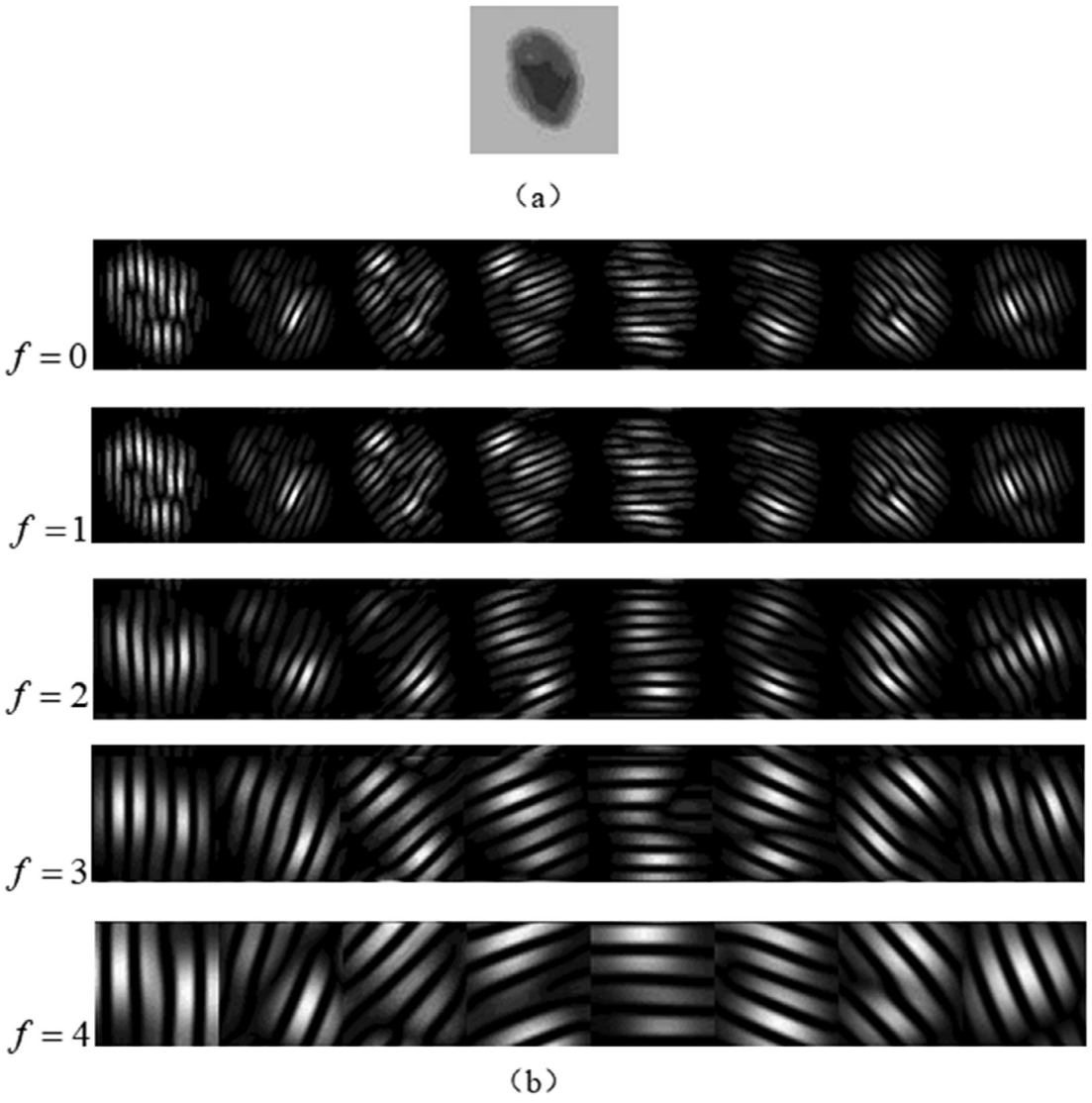

Fig. 11. Gabor feature map with different filtering scale f and frequency angle θ.

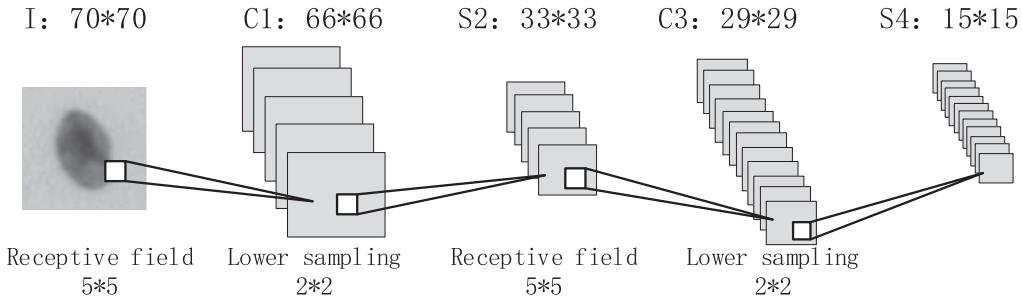

Fig. 12. Proposed LeNet-5 model for feature extraction.

of feature maps was 12, in this way the layer C3 was obtained. As for layer C3, the dimension of each feature graph was $(33-5+1)\times(33-5+1)=29\times29$. S4 is a pooling layer with the max-pooling method. The pooling window size was also set to 2×2, step size was 2. So S4 layer has the same 12 feature maps as C3 layer after mapping, and the dimension of each feature map is $(29+1)\div2=15$, so the final feature dimension is $15\times15\times12=2700$. In addition, sigmoid function was used for all activation functions in CNN. Parameters of CNN network is shown in Table 3.

3.2. Multi feature fusion

The features extracted by CNN are first normalized before being fused with strong features, after that the serial fusion is carried out. Finally, the Principal components analysis (PCA) method was

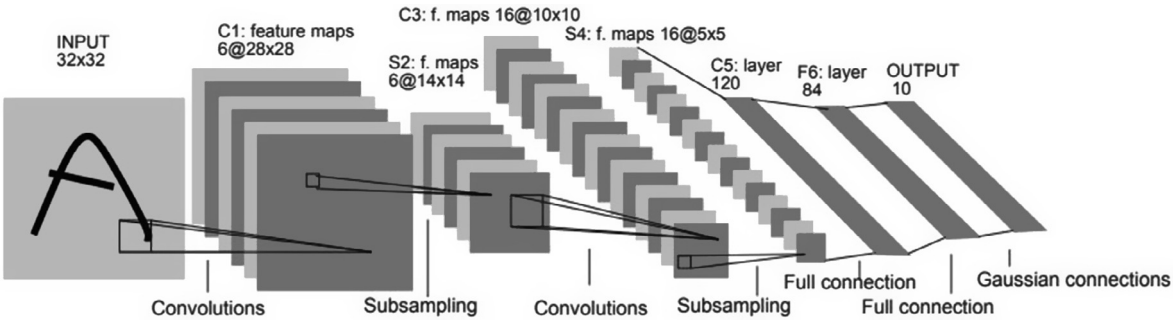

Fig. 13. Structure diagram of classical LeNet-5 neural network.

applied to reduce the dimension of fused features. Feature fusion is conducted as follows.

(1) Standardization of parameters

$$x_i' = \frac{x_i - \mu}{\sigma} \tag{6}$$

μ, σ are the mean and standard deviation of the eigenvector parameters respectively in formula 6,

(2) Multi feature serial fusion

After standardization, the Gabor + GLCM feature vector is $X_1 = [x_{11}, x_{12}, \cdots x_{1,m}]$, and the vector of abstract features extracted by CNN is $X_2 = [x_{21}, x_{22}, \cdots x_{2,n}]$. After the serial combination, fused feature $W = (w_1, w_2, \cdots, w_{m+n}) = (\alpha X_1, \beta X_2)$ was obtained. α and β are weighted coefficients for the features of strong feature and CNN, representing the weight of different features after fusion. The range of values is $(0, 1)$ and $\alpha + \beta = 1$. A group of experiments showed that when $\alpha = 0.4, \beta = 0.6$ were set, the classification effect was the best.

(3) Feature dimensionality reduction

After the multi feature serial fusion, the dimension of the feature vector W is up to 2700 + 2560 = 5260. If the fused feature is directly input into the SVM classifier, the computational complexity is too high that the classification performance is also affected. Therefore, the PCA method is used to reduce the dimension of the eigenvector W.

The overall scatter matrix P for the set of training samples is shown as follows:

$$P = \frac{1}{N}\sum_{i=1}^{N}(a_i - m)(a_i - m)^T = AA^T \tag{7}$$

In the formula (7), a_i represents the training samples, $m = \frac{1}{N}\sum_{i=1}^{N} a_i$ represents the mean vector of sample A, and N is the number of samples in A. A set of eigenvalues and the corresponding eigenvector can be obtained by solving the matrix P, and the vectors are sorted according to the size of their eigenvalues. A $m \times n$ dimension transformation matrix S can be created by applying the former n column vector. Finally, PCA is used to reduce the dimension of the fusion features and get the first 799 dimensional features, so that the characteristic matrix $W_{799\times N}$ can be trained. Similarly, assuming the number of test samples is M, there is a testing characteristic matrix $W_{799\times M}$.

3.3. CNN-SVM network model

3.3.1. Dual channel network structure

CNN is a network structure which has been developed in recent years. The neurons between its adjacent layers are not fully connected, but partly linked, which means that the perceptual region of a neuron comes from part of the neurons in the upper layer. The convolution layer and the max-pooling layer constitute the hidden layer of CNN, which can be connected in series [36]. The structure of the proposed LeNet-5 neural network is shown in Fig. 13 Table 2.

The core idea of SVM is to find an optimal hypersurface to maximize the distance between the two sides, meanwhile achieving the 2 classification. Linear classifier is mostly used to find the hyperplane in multidimensional space to divide two types of separable eigenvectors, as shown in Fig. 14.

In general, the classical convolution neural network can be applied as the feature extraction module and classifier [37], layers

Table 2
Evaluation criterion confusion matrix.

	Prediction example (P′)	Negative prediction (N′)
Actual positive example (*P*)	True example (*TP*)	False negative example (*FN*)
Actual negative example (*N*)	False positive example (*FP*)	True negative example (*TN*)

Table 3
CNN network parameter setting.

Layer category	number of feature graph	feature graph size
Input layer I	1	70 × 70
Convolution layerC1	5	66 × 66
Lower sampling layerS2	5	33 × 33
Convolution layer C3	12	29 × 29
Lower sampling layer S4	12	15 × 15

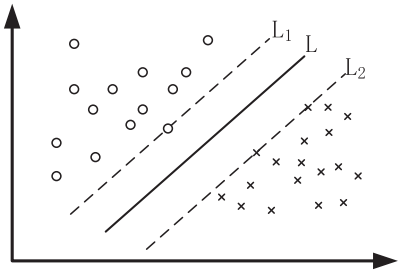

Fig. 14. Schematic diagram of linear SVM.

of CNN can be used to extract the features of the image. Features extracted from convolutional neural networks are more abundant and applicable than those extracted manually by adjusting the parameters. CNN itself can accomplish the task of classification independently, while the recognition effect of traditional model is limited. SVM classifier is usually introduced to improve the recognition effect of CNN, whose full connection layer is replaced by SVM, the remaining convolution layer and sub-sampling layer automatically extract features from the input data. These fused features are processed by SVM algorithm to achieve the cell classification, which is the CNN-SVM classification network.

Strong features have important reference value for the judgment of cervical cancer. However, features extracted by convolutional neural network are abstract that the physical meaning is not clear. Therefore, in view of strong features, we proposed the method of fusing strong features extracted by prior knowledge with the abstract features extracted by convolutional neural network after standardization as well as dimensionality reduction and input them into the trained SVM model. The topology structure of the model is shown in Fig. 15.

In this model, feature extraction is divided into two parallel paths, which are strong feature extraction and convolution neural network path. The network integrates two features in serial and input them into the SVM classifier.

3.3.2. Cancer epithelial cell recognition

The training steps of strong feature CNN-SVM classifier for cell classification are as follows.

Algorithm 1	The training algorithm for strong CNN-SVM model
Step 1	Judge the training sample set: if the sample size is insufficient, sample set will be amplified; otherwise, the algorithm will directly enter the next step
Step 2	Create CNN network, randomly initialize CNN and SVM parameter values
Step3	Generate the Gabor filter and combine with GLCM, extract the total 40 feature maps of $\theta = 0, \pi/8, \pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4, 7\pi/8$ direction and $f = 0, 1, 2, 3, 4$ scale from the sample image I_i.9 × 9 grid is used to reduce the dimension of 70 × 70 feature to 8 × 8, the first link of the feature graph form the feature vector $X_{i1} = [x_{11}, x_{12}, \cdots x_{1,m}]$
Step 4	Sort the same sample image and input into the CNN to calculate the output of hidden layer, meanwhile extracting the characteristic part $X_{i2} = [x_{21}, x_{22}, \cdots x_{2,n}]$
Step 5	If all samples have been extracted, continue; otherwise, go to Step 3
Step 6	Strong features of all samples are $X_1 = [x_{11}, x_{12}, \cdots x_{1,M}]$, features extracted automatically by CNN are $X_2 = [x_{21}, x_{22}, \cdots x_{2,N}]$. X_1, X_2 are normalized and serially fused to obtain the fusion feature $W = (w_1, w_2, \cdots, w_{M+N}) = (\alpha X_1, \beta X_2)$. Finally, the PCA method is applied to reduce the dimension of W and get the final fusion feature vector W^*.
Step 7	Input the feature vector W^* to SVM and train model until reaching the allowable range of error or the maximum number of iterating generations

4. Experimental results and discussions

The experiment was implemented under Microsoft Visual Studio 2013 and Python 2.7 platform, which is executed on a computer with 2.8 GHz Intel Core i7 processor and 16 GB memory. Experimental data were derived from the private dataset and the publicly available Herlev dataset introduced in Table 1. The results obtained in the experiment were the optimal ones.

4.1. Comparisons of dataset expansion

The private dataset is divided into training set and testing set according to the ratios of 4:1, fusion ratio was 2:8. In this paper, the size of the image is set to 70 × 70 pixels, and the weighted vote is carried out by using the Gaussian distribution weight with the standard deviation of 35 pixels. System accuracy is the average value of all data blocks using the 5-fold cross-validation.

The parameter settings of model A-I are listed in Table 4. RO (Rotation), SA (scaling), RA (random crop) and SH (shear-warp transformation) were conducted on original pictures. Data of mode A is not expanded as a baseline for experimental comparison.

From the comparison of mode A&I, mode RO can improve the accuracy, indicating that the rotation operation is effective and should be considered. As for mode A&B&C, mode SC need to be used with caution, scaling mode SC0-SC1 can improve the accuracy while mode SC0-SC2 decreased the accuracy rate in contrast. Experimental results show that a certain degree of scaling can improve the accuracy and enhance the robustness of the model, such as model B; but excessive scaling can lead to a decline in accuracy, such as model C. From the comparison of mode A&D and mode C&E, random crop can effectively improve the accuracy of the model, which also conformed to the general rules of model training. As for the mode A&F&G&H, shear-warp transformation was not stable for improving the performance, comparison between the models showed that the optimization effect is not obvious, and the accuracy of the modified model was lower than original. Subsequently, no shear-warp transformation will be applied in this paper. In summary, considering the advantages of model A&B&D&E&I, model J was chosen as the best model, and the accuracy of model J was 91.1% by comparing and analyzing different dataset expansion technologies.

Besides, an experiment on exploring the relationship between dataset size and accuracy was conducted. Training conditions remained the same. Considering the results in Table 4, datasets were amplified using the approaches of rotation, scaling and random crop. Therefore, the largest magnification times reached $4 \times 3 \times 2 = 24$, while the minimum kept 1. Changes of three evaluation parameters under datasets of different sizes are shown in Fig. 16.

It is concluded from the Fig. 16 that three types of parameters gradually increased upward as the size of datasets grew. The proposed model behaved the worst when the dataset magnification was three. The Model effect performed best when its size maximized. Besides, the sensitivity and specificity curves were relatively stable, and the accuracy kept above 85%. This experiment showed that the strong feature CNN-SVM model proposed in this paper had good robustness with the increase of dataset, its model performance may be further improved in larger and more balanced datasets.

4.2. Comparison of strong feature extraction methods

In the analogy experiment of strong feature extraction methods, we investigated different feature extraction methods and their combinations based on the accuracy and efficiency of classification using our private dataset. All the images have been

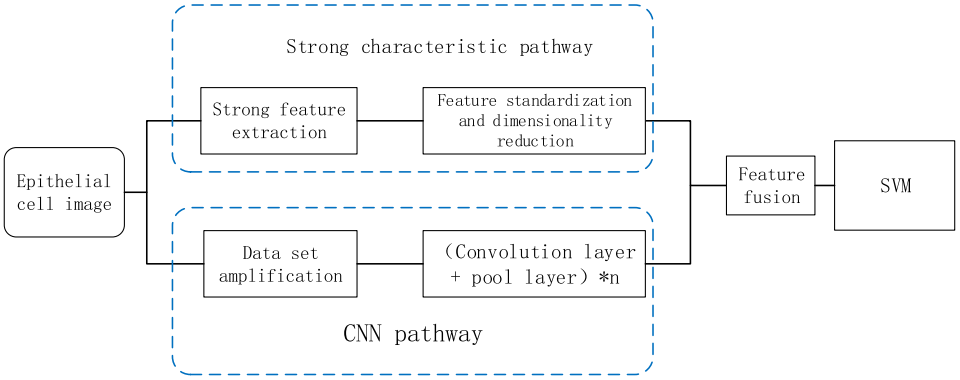

Fig. 15. Topology diagram of CNN-SVM model based on strong feature.

Table 4
Parameters of dataset expansion settings.

Strong characteristic CNN-SVM network	Rotation mode	Scaling mode	Random crop mode	Shear-warp transformation mode	Dataset expansion multiplier	Accuracy
A	RO0	SC0	RA0	SH0	1	88.6
B	RO0	SC0-SC1	RA0	SH0	2	89.0
C	RO0	SC0-SC2	RA0	SH0	3	85.4
D	RO0	SC0	RA1	SH0	1	89.2
E	RO0	SC0-SC2	RA1	SH0	3	86.7
F	RO0	SC0	RA0	SH0-SH1	2	84.6
G	RO0	SC0	RA0	SH0-SH2	3	86.8
H	RO0	SC0	RA0	SH0-SH3	4	85.4
I	RO0-RO3	SC0	RA0	SH0	4	90.3
J	RO0-RO3	SC0-SC1	RA1	SH0	**8**	**91.1**

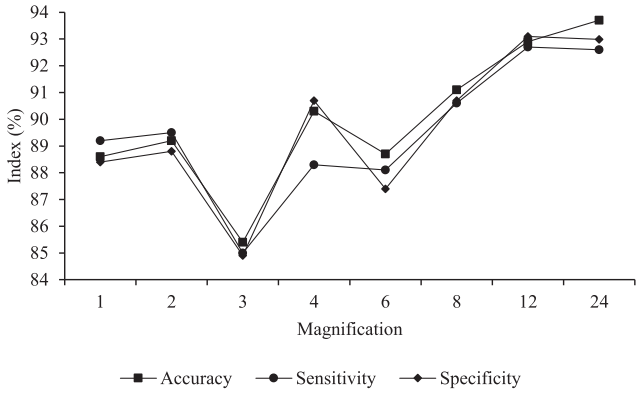

Fig. 16. Comparison of model accuracy under different scale datasets.

cut and normalized to the size of 70 × 70. In order to further expand the experiment, the gray images of 48 × 48 were also normalized.

In this section, strong features of samples from private dataset were extracted with GLCM, Fourier transformation, Gabor transformation, Markov random field and their combinations. Later the dimensions of feature vectors extracted from feature maps were reduced to get similar dimensions with those extracted by CNN. The proportion of feature fusion is set to 1:1 consistently to control variables. Feature vectors were input into the SVM classifier for classification after feature fusion. Ratio between the training set and the test set was divided into 4:1, accuracy is calculated according to 5-fold cross-validation. Different feature extraction methods and their corresponding accuracy are shown in Table 5. Besides, total time of each feature extraction method has been recorded and used for comparative experiments.

Table 5 introduced the accuracy and running time of each feature extraction method on different image sizes. It is concluded that the performance of GLCM + Gabor was better than that of other methods, both in accuracy and processing time. In addition, Gabor transformation and Fourier + Gabor also gave a similar result to GLCM + Gabor, whose performances were better than other methods. While MRF obtained a relatively low accuracy. We speculate it is because MRF only considers the features of a point that relates to a small area nearby, rather than the other regions. Private cell images have continuity on the pixels, and the gray value of the cell has some relationship with the background, so the classification effect of MRF is not ideal. The running time of GLCM was shorter than that of the others, while its accuracy was not the highest. Moreover, the operation speed of small pictures was better than that of larger ones, and the running time of the same method on different data can be compared horizontally. It is concluded that GLCM based on statistical analysis and Gabor filter based on signal analysis have better classification accuracy and response speed when used jointly. Therefore, GLCM + Gabor method is applied as the strong feature extraction method in this paper.

4.3. Experiment on weight selection of multi feature fusion

Strong features and CNN self-extracted features are fused to get a new feature $W = (w_1, w_2, \cdots, w_{m+n}) = (\alpha X_1, \beta X_2)$, $\alpha + \beta = 1$. Weights α, β are selected in order to get the best classification performance.

The private dataset weas applied for 5-fold cross-validation, ratio of training to testing set was 4:1. Different weights and experimental groups were set up and sent to strong feature CNN-SVM classifier with Gauss kernel function. Table 6 shows the accuracy of 10 datasets under different combinations of α, β.

It is concluded from the data in Table 6 that the accuracy of classifier increased with the ratio of α, β gradually balanced. The

Table 5
Accuracy and running time of feature extraction methods on private dataset.

Method	70 × 70		48 × 48	
	Accuracy	Time (s)	Accuracy	Time (s)
GLCM	88.6	31	89.1	27
Fourier	89.8	46	89.5	41
Gabor	91.3	53	90.1	48
MRF	75.7	38	76.5	35
GLCM + Fourier	87.0	48	84.4	45
GLCM + Gabor	**92.1**	**46**	**90.9**	**49**
GLCM + MRF	89.5	58	87.3	42
Fourier + Gabor	91.4	45	90.7	43
Fourier + MRF	85.9	44	82.8	49
Gabor + MRF	82.6	50	79.3	45

classifier obtained the highest accuracy when $\alpha/\beta = 4/6$, which was chosen as the optimum proportion.

4.4. Feature extraction experiment of CNN

Aiming at evaluating the performance of CNN model as the feature extraction method and its effects in the classification of cell images from different datasets. Features from various layers of CNN were picked out for feature fusion and then input into SVM for classification. 2-class classification task are implemented using the private dataset and Herlev, accuracy, sensitivity, and specificity were considered as evaluation indexes. Training set and testing set is still divided as 4:1, the average of classification metrics was calculated by the method of 5-fold cross-validation, fusion ratio is set as 4:6. Tables 7 and 8 summarized the performances of each layer. It is concluded that low sampling layer S4 achieved the best classification performance. The results of other layers like C1, S2 and C3 were lower than that of S4. The deeper network structure, like convolution and lower sampling layer will enhance the accuracy of CNN in extracting features. As for the 2-class classification task of private dataset and Herlev, the deepest layer of the CNN feature extraction module, lower sampling layer S4, obtained the best classification accuracy and sensitivity, which also shows the rationality of the CNN structure proposed in this paper.

Table 7
Accuracies (%) and sensitivity (%) of 2-class classification using private dataset.

CNN layer	Accuracy	Sensitivity	Specificity
Convolution layer C1	90.4	90.0	90.2
Lower sampling layer S2	94.2	92.0	92.3
Convolution layer C3	93.2	91.8	91.7
Lower sampling layer S4	**94.5**	**93.0**	**93.5**

Table 8
Accuracies (%) and sensitivity (%) of 2-class classification using Herlev.

CNN layer	Accuracy	Sensitivity	Specificity
Convolution layer C1	97.5	97.7	98.0
Lower sampling layer S2	98.5	97.3	98.3
Convolution layer C3	98.9	98.6	98.8
Lower sampling layer S4	**99.3**	**98.9**	**99.4**

4.5. Evaluation of strong feature CNN-SVM and its basic models

Strong feature CNN-SVM model proposed in this paper originated from classical CNN and SVM classifier. Comparative experiments were conducted to illustrate the superiority of the proposed model. Private dataset mentioned before was applied for comparison. Strong features and features extracted from the hidden layer of CNN were fused according to ratio of 4:6, training set to the testing set was set as 4:1. The average accuracy, sensitivity and specificity was calculated by 5-fold and 10-fold cross-validation, respectively. Batch size was set to 30 according to the size of the dataset. Learning rate was 0.01 according to the change of error rate in the training process. C-SVC algorithm and radial kernel function were used for SVM training. Grid search provided by Pycharm are used to find the optimal penalty parameter C and the kernel functionγ. Results are shown in Figs. 17 and 18.

Figs. 17 and 18 showed the accuracy of CNN, CNN-SVM and strong feature CNN-SVM using 5-fold and 10-fold cross-validation on private dataset. It is concluded that the accuracy of strong feature CNN-SVM model was higher than those of the other two models. The standard deviation of strong feature CNN-SVM model was smaller than that of others. It has been proved that the performance and generalization ability of strong feature CNN-SVM are better improved by combining with SVM. Moreover, the accuracy of classification under the condition of 10-fold cross-validation was slightly higher than that under the condition of 5-fold cross-validation.

In the acquisition of medical image data, unbalanced samples are ubiquitous. Confusion matrix were applied to visualize the

Table 6
Accuracy of classifiers under different ratios of α, β.

Weight Number	[1,9]	[2,8]	[3,7]	[4,6]	[5,5]	[6,4]	[7,3]	[8,2]	[9,1]
1	0.899	0.912	0.938	**0.939**	0.932	0.912	0.909	0.799	0.681
2	0.907	0.918	0.928	**0.932**	0.922	0.913	0.901	0.782	0.672
3	0.899	0.916	0.928	**0.926**	0.921	0.922	0.898	0.799	0.692
4	0.903	0.932	0.933	**0.934**	0.929	0.921	0.902	0.801	0.601
5	0.902	0.920	0.939	**0.948**	0.939	0.913	0.901	0.803	0.703
6	0.903	0.913	0.925	**0.930**	0.921	0.920	0.910	0.798	0.718
7	0.901	0.913	0.932	**0.938**	0.931	0.901	0.903	0.795	0.693
8	0.902	0.916	0.920	**0.921**	0.921	0.913	0.902	0.803	0.702
9	0.890	0.911	0.931	**0.934**	0.924	0.911	0.898	0.801	0.691
10	0.906	0.910	0.921	**0.932**	0.925	0.909	0.902	0.799	0.679
Mean value	0.901	0.916	0.930	**0.933**	0.927	0.914	0.903	0.798	0.683

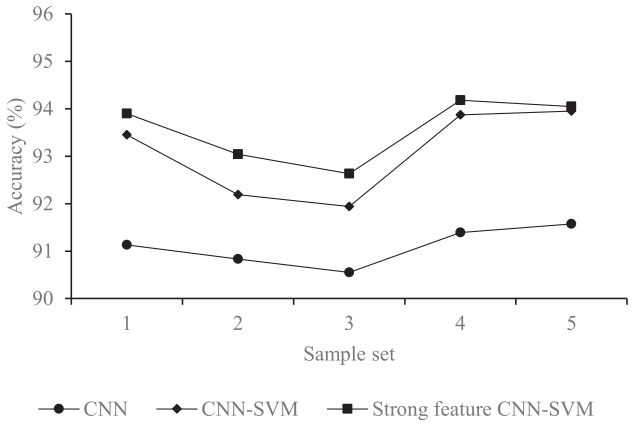

Fig. 17. Accuracy of 3 models using 5-fold cross-validation.

	Positive	Negative
Positive	**92.8**	7.2
Negative	7.8	**92.2**

CNN

	Positive	Negative
Positive	**94.7**	5.3
Negative	6.0	**94.0**

CNN-SVM

	Positive	Negative
Positive	**95.2**	4.8
Negative	5.2	**94.8**

Strong feature CNN-SVM

Fig. 19. Confusion matrix of different models.

performance of different classifier algorithms. If the elements off the diagonal are all 0, then the classifier performs well. Confusion matrix experiment was conducted on the private dataset using 10-fold cross-validation, other conditions were kept the same.

Fig. 19 reported the confusion matrix of three models on the private dataset. Data of the confusion matrix presented the average accuracy of model in the form of percentage. It is concluded that the accuracy of CNN was lower than the other two methods. Performance of CNN-SVM was slightly worse than strong feature CNN-SVM. Hybrid model behaved better than the single model, the addition of strong feature extraction method also improved the performance. Moreover, it has been proved that the proposed model has more obvious advantages in identifying positive cells.

Besides, three classification methods were also compared on the sensitivity and specificity using the same dataset and initialization parameters. Results were shown in Table 9. The sensitivity and specificity contrast curve of three networks were concluded in Figs. 20 and 21.

It is concluded from the Table 9 and Figs. 20 and 21 that strong feature CNN-SVM classifier outperformed in accuracy than the other two models. As for the sensitivity curve, the proposed method gained an improvement of 4.9 and 2.9 percent point comparing with others, respectively. This means a higher accuracy for positive cells. Besides, the proposed model achieved an improvement of 4.1 and 2.7 in specificity than CNN and CNN-SVM model, which signified a higher probability of diagnosis for cancerous cells. It was proved that adding strong feature paths with CNN-SVM model greatly improved the sensitivity and specificity meanwhile guaranteeing the classification accuracy.

4.6. Comparison evaluation with state-of-the-art models and baselines

Comparison of performance was conducted using metrics of Acc, Sn and Sp. As for the proposed method in this paper, private and Herlev datasets were used to evaluate for 2,7 class classification, respectively. The experimental results are as shown in Table 10. Gautam et al. [25], Nguyen et al. [38], Zhang et al. [18], Jith et al. [24] and Lin et al. [39] applied the Herlev dataset for model assessment. The others used private dataset to identify the performance of their methods. We adopted the experimental results rooting from their published work into the table for comparison.

It is concluded for the Table 10 that our proposed method achieved the second highest accuracy in 2 class classification with a 0.3 percent point lower than the peak, and the best performance in 7 class classification applying Herlev dataset. As for the private

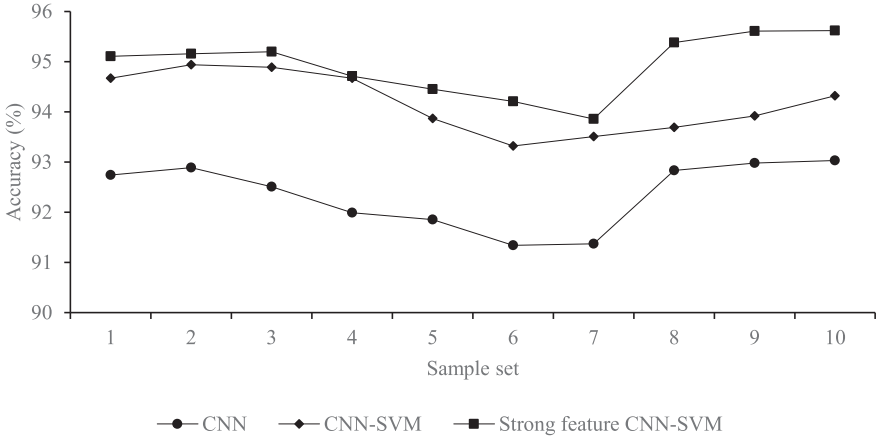

Fig. 18. Accuracy of three models using 10-fold cross-validation.

Table 9
Performance evaluation among three different classifiers.

Evaluation metrics	Acc (%)	Sn (%)	Sp (%)
Classic CNN network	92.4	88.4	89.2
CNN + SVM	94.2	90.4	90.6
Strong feature CNN + SVM	**94.9**	**93.3**	**93.3**

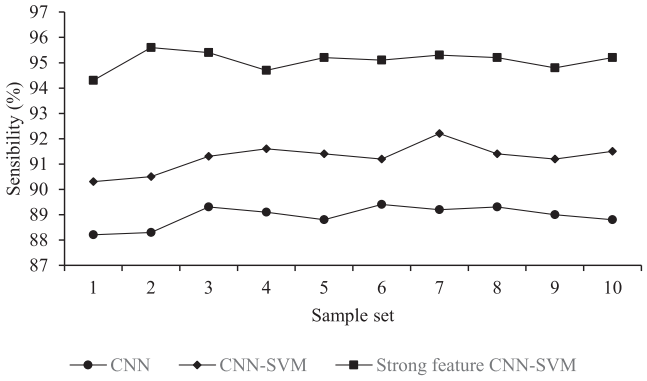

Fig. 20. Sensitivity curves of 3 networks.

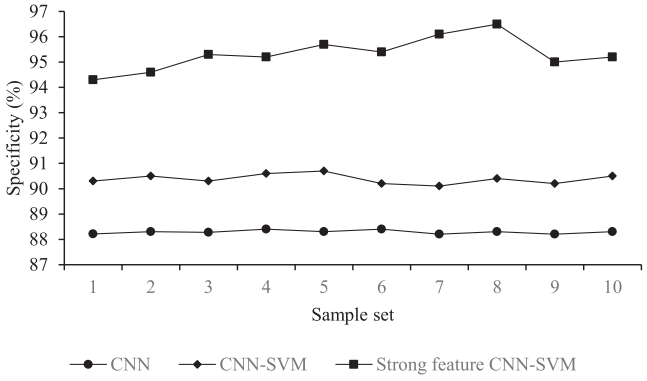

Fig. 21. Specificity curves of 3 networks.

dataset, the proposed method was 1.7 percent higher than the method of Zhang et al. [40]. The approach in this paper outperformed in all metrics than the method of [18,38,39]. Higher accuracy suggested that the system effectively recognizes the total cells. The improvement of sensitivity and specificity meant that the positive and negative samples can be classified better to a great extent. Comparing with method of [18], the sensitivity and specificity of the which are relatively poor, it is easy to produce false and missed classification. The proposed approach was 0.3 point lower than the model of [24], while this part of the gap is not too big. Their model adopted the fine-tuned method, which is more complex than ours in terms of computational complexity. The proposed model can be further improved by optimizing network parameters and enriching datasets. Therefore, the superiority of the proposed method is clearly demonstrated comparing with the other state-of-the-art models using two different datasets and three metrics.

Existing CNN-SVM approaches were applied as baselines for comparison aiming at characterizing the performance of the proposed method on cervical cell classification. The accuracy was used as the evaluation index, and the comparison results are shown in the Fig. 22.

Approaches of [26–28,41] were based on CNN-SVM model, while methods of [22,42] were traditional CNN models, comparison among which were conducted to demonstrate the superiority of the proposed model. Our model outperformed than others in accuracy, with an improvement of 0.11 percentage point over the second best method [27]. As for the sensitivity and specificity, method [27] and [41] were a little higher than the proposed model as 0.61 and 0.6 percentage point, respectively. Accuracy has more practical clinical significance than sensitivity and specificity in the classification of cervical cells. The weakness of the proposed model in other indicators can be further improved by parameter adjusting and dataset enlarging. Therefore, the strong feature method in this paper achieved certain improvement comparing with the baselines.

4.7. Analysis of iteration times and model speed optimization

4.7.1. Relation between iteration times and model accuracy

Following experiment was conducted to study the relationship between training time and accuracy, which adopted the 10-fold cross validation. The experiment was repeated two times, so each model was trained 20 times and iterated 500 times before reaching stable, hardware parameters and specific conditions were kept same. Finally, results of CNN, CNN-SVM and strong feature CNN-SVM model during 1–256 iterations were obtained. LeNet-5 model was used to construct the hidden layer of three models. Experimental results were shown in Fig. 23.

It was concluded that the convergence speed of strong feature CNN-SVM and CNN-SVM model were faster, accuracies of these were always higher than that of the traditional CNN model. Strong feature CNN-SVM model performed slightly better than CNN-SVM in terms of accuracy but the convergence speed of both models was close. In the early stage of iteration, the accuracy of CNN differed greatly from the other two methods, because the initializa-

Table 10
Comparison of classification performance among different methods (Sensitivity (Sn), Specificity (Sp), Accuracy (Acc)).

Ref	Method	Class	Data	Result (%)
[25]	a CNN-based feature analysis and a transfer learning-based approach	2,7	Herlev and private dataset	2-class: Acc = 99.3 7-class: Acc = 93.75
[38]	Feature concatenation and ensemble CNNs	7	Herlev dataset	Acc = 93.51
[18]	Nuclei centered patched-based CNN throughTransfer Learning	2	Herlev dataset	Acc = 98.3, Sn = 98.2, Sp = 98.3
[24]	CNN based on fine-tuned AlexNet	2	Herlev and private dataset	Acc = 99.6
[39]	Pre-trained CNN using a five-channel input	2,7	Herlev dataset	2-class: Acc = 94.5; Sn = 97.4; Sp = 90.4; 7-class: Acc = 64.5
[40]	A size-sensitive fully convolutional network (R-FCN)	–	Private dataset	Abnormal region: Acc = 93.2
–	Strong feature &CNN-SVM	2,7	Herlev and private dataset	Herlev 2-class: Acc = 99.3, Sn = 98.9, Sp = 99.4 7-class: Acc = 93.8, Sn = 93.7, Sp = 93.7 Private 2-class: Acc = 94.9, Sn = 93.3, Sp = 93.3

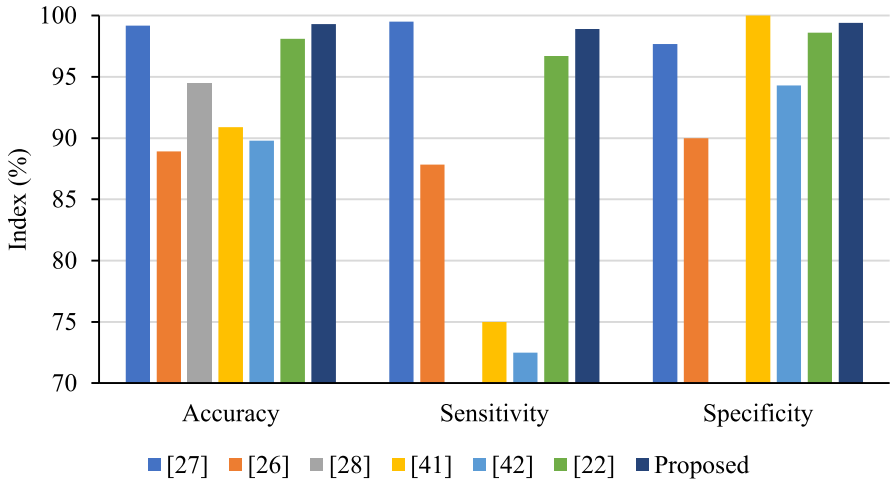

Fig. 22. Baseline comparison experiment.

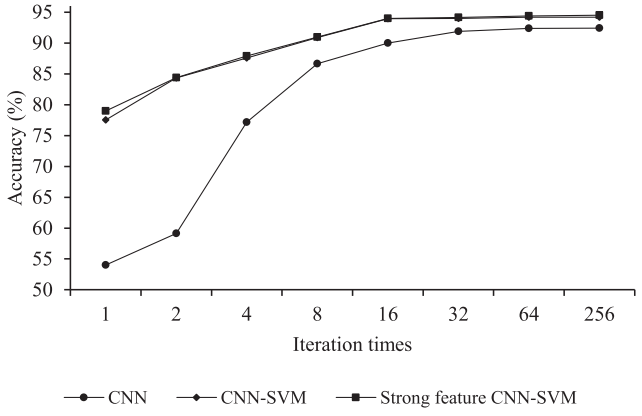

Fig. 23. Relation between iteration times and model accuracy.

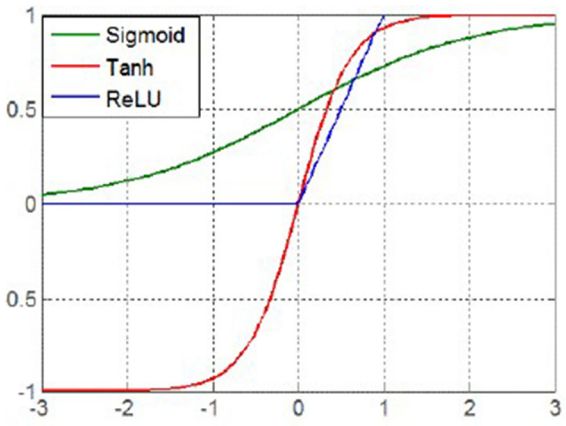

Fig. 24. Model curve of activation function.

tion of network weights of CNN is random when not fully trained. And the SVM layer of the other two models can complete the data mapping of pictures, which can ensure the accuracy. The accuracy of the proposed model is higher than that of CNN-SVM because some strong features were selected based on prior knowledge. When the iteration is almost finished, performances of three models were better, while the proposed approach kept better than the others. The model accuracy of CNN-SVM was 94.2%, while that of the strong feature model was 94.5%, slightly higher. It was been proved that proposed model can effectively extract strong features in the process of classification under the same times of iteration, thus achieving higher accuracy in the process of calculating network weights.

4.7.2. Model speed optimization evaluation

As for deeper neural networks, a non-linear activation function is usually needed for being connected. Commonly used activation functions are as follows: *Sigmoid*, *ReLU*, *Tanh*,and so on. The functional model is shown in the Fig. 24.

Sigmoid maps the input to the range of $(0, 1)$, while *Tanh* maps the input to the range of $(-1, 1)$. As a saturated function,*Sigmoid* has the disadvantage of large computation, which leads to the slow response of the model in the below comparative experiments. *ReLU* is a kind of unsaturated non-linear function, which can set the negative value to 0, thus ensuring the sparsity of the model. *ReLU* also adjusts the number of non-zero values in the network through gradient training, aiming at transferring gradients better and giving the model more robustness when extracting features. The original *Sigmoid* function in the system was replaced with *ReLU*. A comparative experiment on model processing time under the same conditions was carried out using the private dataset. Experimental results were described in Table 11. It was concluded that the processing efficiency of *ReLU* outperformed than the other two models, which provided experimental proof for the selection of activation function.

Table 11
Comparative experiment on model processing time.

Model	Sigmoid (s)	Tanh (s)	ReLU (s)
Strong feature CNN-SVM model	47	68	**41**

5. Conclusions

A strong feature CNN-SVM approach for cervical epithelial cell classification was proposed in this paper. The designed framework included three main parts: feature extraction, feature fusion and classification. We fused the strong features extracted from Gabor + GLCM with the abstract ones in hidden layer of CNN and input them into the SVM for classification, by which a superior performance was obtained. The method was evaluated on two inde-

pendent datasets. Our proposed approach outperformed than the state-of-the-art models and baselines with the Acc, Sn, and Sp of 99.3, 98.9, 99.4 for 2-class detection in the mass, respectively. Besides, the method in this paper behaved well in the classification of positive cells, which demonstrated the efficiency of our approach in identifying cancerous cells. We will keep focusing on optimizing the feature extraction frame of both strong features and neural network, aiming at further improving the accuracy and efficiency of 7 class classification on cervical cells in the future work. The collection and labeling of private datasets will be promoted in order to further expand the size and variety. We hope to open the dataset to other researchers for modeling and learning in the future, ultimately further help the study in the field.

CRediT authorship contribution statement

A. Dongyao Jia: Conceptualization, Methodology, Data curation, Writing - original draft. **B. Zhengyi Li:** Software, Validation. **C. Chuanwang Zhang:** Conceptualization, Methodology, Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neucom.2020.06.006.

References

[1] F. Bray et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA 68 (6) (2018) 394–424.
[2] W. Small Jr et al., Cervical cancer: a global health crisis, Cancer 123 (13) (2017) 2404–2412.
[3] T.A. Kessler, Cervical cancer: prevention and early detection, Seminars in Oncology Nursing, Elsevier, 2017.
[4] J. Yessi, S. Ng, N. Osman, Intelligent screening systems for cervical cancer, Sci. World J. 2018 (2014) 810368.
[5] K. Bora et al., Automated classification of Pap smear images to detect cervical dysplasia, Comput. Methods Programs Biomed. 138 (2017) 31–47.
[6] T. Chankong, N. Theera-Umpon, S. Auephanwiriyakul, Automatic cervical cell segmentation and classification in Pap smears, Comput. Methods Programs Biomed. 113 (2) (2014) 539–556.
[7] Y.-F. Chen et al., Semi-automatic segmentation and classification of pap smear cells, IEEE J. Biomed. Health Inf. 18 (1) (2013) 94–108.
[8] E. Bengtsson, P. Malm, Screening for cervical cancer using automated analysis of PAP-smears, Comput. Math. Meth. Med. 2014 (2014.).
[9] M.E. Plissiti, C. Nikou, Cervical cell classification based exclusively on nucleus features, International Conference Image Analysis and Recognition, Springer, 2012.
[10] L. Zhang et al., Automation-assisted cervical cancer screening in manual liquid-based cytology with hematoxylin and eosin staining, Cytometry Part A 85 (3) (2014) 214–230.
[11] N. Zeng et al., A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease, Neurocomputing 320 (2018) 195–202.
[12] P.-C. Huang et al., Quantitative assessment of Pap smear cells by PC-based cytopathologic image analysis system and support vector machine, International Conference on Medical Biometrics, Springer, 2008.
[13] J. Hyeon et al., Diagnosing cervical cell images using pre-trained convolutional neural network as feature extractor, 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), IEEE, 2017.
[14] K. Bora et al., Pap smear image classification using convolutional neural network, Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, 2016.
[15] R. Gupta, A. Sarwar, V. Sharma, Screening of cervical cancer by artificial intelligence based analysis of digitized papanicolaou-smear images, Int. J. Contemp. Med. Res. 4 (5) (2017).
[16] A. GençTav, S. Aksoy, S. ÖNder, Unsupervised segmentation and classification of cervical cell images, Pattern Recogn. 45 (12) (2012) 4151–4168.
[17] Y. Marinakis, M. Marinaki, G. Dounias, Particle swarm optimization for pap-smear diagnosis, Expert Syst. Appl. 35 (4) (2008) 1645–1656.
[18] L. Zhang et al., DeepPap: deep convolutional networks for cervical cell classification, IEEE J. Biomed. Health Inf. 21 (6) (2017) 1633–1643.
[19] T. Mahmood et al., Artificial intelligence-based mitosis detection in breast cancer histopathology images using faster R-CNN and deep CNNs, J. Clin. Med. 9 (3) (2020) 749.
[20] N. Zeng et al., An improved particle filter with a novel hybrid proposal distribution for quantitative analysis of gold immunochromatographic strips, IEEE Trans. Nanotechnol. 18 (2019) 819–829.
[21] N. Zeng et al., Image-based quantitative analysis of gold immunochromatographic strip via cellular neural network approach, IEEE Trans. Med. Imaging 33 (5) (2014) 1129–1136.
[22] Kurnianingsih et al., Segmentation and classification of cervical cells using deep learning, IEEE Access 7 (2019) 116925–116941.
[23] N. Kaur, N. Panigrahi, A. Mittal, Automated cervical cancer screening using transfer learning, Proceedings of International Conference on Recent Advances in Engineering Science and Management, 2017.
[24] O.N. Jith et al., DeepCerv: Deep neural network for segmentation free robust cervical cell classification, in: Computational Pathology and Ophthalmic Medical Image Analysis, Springer, 2018, pp. 86–94.
[25] S. Gautam et al., Considerations for a PAP smear image analysis system with CNN features. arXiv preprint arXiv:1806.09025, 2018.
[26] Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. Lecture Notes in Computer Science. 2016.
[27] Medical Image Understanding and Analysis. Communications in Computer and Information Science. 2017.
[28] M.E. Plissiti et al., SIPAKMED: A new dataset for feature and image based classification of normal and pathological cervical cells in Pap smear images, 2018 25th IEEE International Conference on Image Processing (ICIP), IEEE, 2018.
[29] Y.H. Liu, Feature extraction and image recognition with convolutional neural networks, J. Phys.: Conf. Series (2018), IOP Publishing.
[30] M. Arora, S. Dhawan, K. Singh, Exploring deep convolution neural networks with transfer learning for transformation zone type prediction in cervical cancer, in: Soft Computing: Theories and Applications, Springer, 2020, pp. 1127–1138.
[31] J. Shijie et al., Research on data augmentation for image classification based on convolution neural networks, 2017 Chinese Automation Congress (CAC), IEEE, 2017.
[32] Y. Guo, J. Hou, M. Yu, et al., Content-aware image resizing based on random-carving with probability, J. Commun. 38 (06) (2017) 30–38.
[33] J. Sang, S. Wang, X. Niu, Quantum realization of the nearest-neighbor interpolation method for FRQI and NEQR, Quantum Inf. Process. 15 (1) (2016) 37–64.
[34] Y. LeCun, LeNet-5, convolutional neural networks. URL: http://yann. lecun. com/exdb/lenet, 2015. 20: p. 5.
[35] Z. Wu, D. Lin, X. Tang, Deep markov random field for image modeling, European Conference on Computer Vision, Springer, 2016.
[36] L.O. Chua, CNN: A paradigm for complexity, World Scientific, 1998.
[37] A. Sharif Razavian et al., CNN features off-the-shelf: an astounding baseline for recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2014.
[38] L.D. Nguyen et al., Biomedical image classification based on a feature concatenation and ensemble of deep CNNs, J. Ambient Intell. Hum. Comput. (2019) 1–13.
[39] H. Lin et al., Fine-grained classification of cervical cells using morphological and appearance based convolutional neural networks, IEEE Access 7 (2019) 71541–71549.
[40] J. Zhang et al., Abnormal region detection in cervical smear images based on fully convolutional network, IET Image Proc. 13 (4) (2018) 583–590.
[41] D. Selvathi, W.R. Sharmila, P.S. Sankari, Advanced computational intelligence techniques based computer aided diagnosis system for cervical cancer detection using pap smear images, in: Classification in BioApps, 2018, pp. 295–322.
[42] N. Sompawong et al., Automated pap smear cervical cancer screening using deep learning, 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2019.

Dongyao Jia was born in 1974 in Henan Province, China. He graduated from Tsinghua University in 2006 with a PhD in Engineering. Now he is the professor of Beijing Jiaotong University as well as the Ph.D. supervisor. Main research directions: new detection technology in traffic field, medical image processing and data mining. He has presided or participated in 863 sub-projects, 863 major projects, major projects of the National Natural Science Foundation, the "Science and Technology Support" project of the Ministry of Science and Technology, the third-class Hospital, metro corporation as well as other engineering and scientific research projects. Academic work: (1) Jia Dongyao, Hu Po,Zou Shengxiong,Changzhu Zhang. Optimization Method for Crop Growth Characteristics Based on Improved Locality Preserving

Projection[J]. Journal of Applied Mathematics,2014,2014. (2) Jia Dongyao,Ai Yanke, Zou Shengxiong. Reduction of Multidimensional Image Characteristics Based on Improved KICA[J]. Journal of Applied Mathematics,2014,2014. (3) Dongyao Jia,Meng Li,Huaihua Zhu,Bing Zhang. Layer-Cluster Topology Sensor Node Deployment for Large-Scale Multi-Nodes of WSN[J]. Wireless Personal Communications,2017,94(4). (4) Jia Dongyao,Zou Shengxiong,Li Meng,Zhu Huaihua. Adaptive multi-path routing based on an improved leapfrog algorithm[J]. Information Sciences,2016,367-368. (5) Dongyao Jia,Huaihua Zhu,Shengxiong Zou,Ke Huang. Recognition method based on Green Associative Mechanism for weak contrast vehicle targets[J]. Neurocomputing,2016,203. (6) Jia dongyao, Wang TingTing, ChenXi, A control system of indoor air quality based on modified smith predictor, Proceedings of 2010 CrossStrait Conference on Information Science And Technology, 2010, QinHuangDao, China, pp. 718–722. (7) Jia dongyao, Yang wenkao, Liu ze, Metamerism breakdown characteristic and its application in detection of foreign materials, Spectroscopy and Spectral Analysis, 2008, 28(11), 2596–2600.

Zhengyi Li was born in 1993 in Shandong Province, China. He graduated from the Department of Automation, Beijing Jiaotong University and obtained his master's degree in 2018. After graduation, he joined the national laboratory of digital multimedia in HIKVISION and worked as a senior algorithmic engineer. His main research content is medical images detection and recognition, tracking algorithms based on deep learning. His main work is to use deep learning technology to intellectualize traditional household appliances.

Chuanwang Zhang was born in 1995 in Hebei Province, China. He received his bachelor and master's degrees from Beijing Jiaotong University and is currently preparing to pursue his doctorate. His main research direction is medical image detection. He has been responsible for several cervical cancer and breast cancer detection projects. He used to develope a prototype machine which can be popularized in the market and possessed rich experience in the field of medical image processing.