Text Line Recognition of Dai Language using Statistical Characteristics of Texture Analysis and Deep Gaussian Process

Jingying Zhao[1, 2], Na Dong[1], Hai Guo[1, *], Yifan Liu[1], Doudou Yang[1]

[1] College of Computer Science and Engineering, Dalian Minzu University,
18 LiaoheWest Road, Dalian Development Zone, Dalian, 116600,
China

[2] Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology,
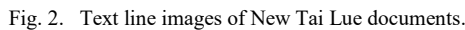Dalian, 116023,
China

Abstract—In view of the different recognition methods of Dai in different language, we proposed a novel method of text line recognition for New Tai Lue and Lanna Dai based on statistical characteristics of texture analysis and Deep Gaussian process, which can classify different Dai text lines. First, the Dai text line database is constructed, and the images are preprocessed by de-noise and size standardization. Gabor multi-scale decomposition is carried out on two Dai text line images, and then the statistical features of image entropy and average row variance feature is extracted. The multi-layers Deep Gaussian process classifier is constructed. Experiments show that the accuracy of text line classification of New Tai Lue and Lanna Dai based on Deep Gaussian process is 99.89%, the values of precision, recall and f1-score are 1, 0.9978 and 0.9989, respectively. The combination of Gabor texture analysis average row variance statistical features and Deep Gaussian process model can effectively classify the text line of New Tai Lue and Lanna Dai. Comparative experiments show that the classification accuracy of the model is superior to traditional methods, such as Gaussian Naive Bayes, Random Forest, Decision Tree, and Gaussian Process.

Keywords—Tai Lue; text line recognition; Deep Gaussian process; character recognition.

I. INTRODUCTION

DOCUMENT recognition is one of the branches of character recognition research, which aims to distinguish two or more different language document images to obtain a target document image [1]. The recognition methods of Dai in different languages need to recognize Dai documents in different languages by coarse classification, and documents in different languages involve different character cutting method. New Tai Lue characters no tone marks, it can be cutting by connected component algorithm, but this cutting method is not applicable to Lanna Dai characters with tone marks. Classification of documents in different languages will benefit the research progress of automatic character recognition [2]. Wu et al. Proposed MLC-Crnn method to solve the problem of long training time and poor effect in English text line, and achieved good results in English text line [3]. Chen et al. Proposed a multi task learning framework, which uses SepMDLSTM to recognize text images and achieves good results in English, French and IFN / ENIT [4]. In order to solve the time-consuming problem of attention mechanism, Huang et al. Proposed a fast end-to-end system called anti feature enhancement network (AFEN) for text box detection and recognition [5]. Qasem Abu al haija et al. Used convolutional neural network to study the recognition problem deeply [6,7]. Therefore, this paper studies the classification of text line in New Tai Lue and Lanna Dai documents. New Tai Lue and Lanna Dai has the following characteristics:

(a)

(b)

Fig. 1. Comparison between New Tai Lue and Lanna Dai document text line (a) New Tai Lue (b) Lanna Dai.

Image(1)

Image(2)

Image(3)

Image(4)

Image(5)

Image(6)

Image(7)

Image(8)

Image(9)

Image(10)

Fig. 2. Text line images of New Tai Lue documents.

1) New Tai Lue and Lanna Dai have high similarity. As shown in Fig. 1, the main difference between them lies in the tone marks. There are few tone marks in New Tai Lue and more tone marks in Lanna Dai. In addition, their appearance is almost the same, so it is difficult to distinguish them.

2) The shape of text line image of New Tai Lue and Lanna dai documents is not good for recognition. Traditional image recognition or handwriting recognition tasks are usually completed by extracting the image features and input them into the classifier or directly input the image into the deep model. Either way, the size of the image needs to be standardized into n*n. The text line image is a rectangle with large width and small height. If the image size is directly standardized to n*n, it will cause huge deformation of the original Dai text image, resulting in serious loss of detail, which is not good for recognition.

In view of the above characteristics, this paper proposes a recognition method based on the statistical features of texture analysis and the Deep Gaussian process (DGP) to classify the New Tai Lue and Lanna Dai text lines. This method can overcome the problems that the characters of different languages of Dai are very close to each other and it is difficult to classify documents by optical recognition, so it is more suitable for the recognition of documents of different languages of Dai. At the same time, this method solves the problem of similar characters between different characters, which is difficult to classify. Therefore, this method can not only achieve good results on Dai characters, but also can recognize and classify other characters with similar shapes. Achieve very good results and serve as a reference.

II. BUILDING AND PREPROCESSING OF DAI TEXT LINE DATABASE

A. Data collection

At present, there is no text line database can be directly used for document classification of New Tai Lue and Lanna Dai. So, text line images of handwritten New Tai Lue and Lanna Dai documents were collected by image projection line segmentation technology [8], each class 1,800 images, and some example images are shown in Fig. 2.
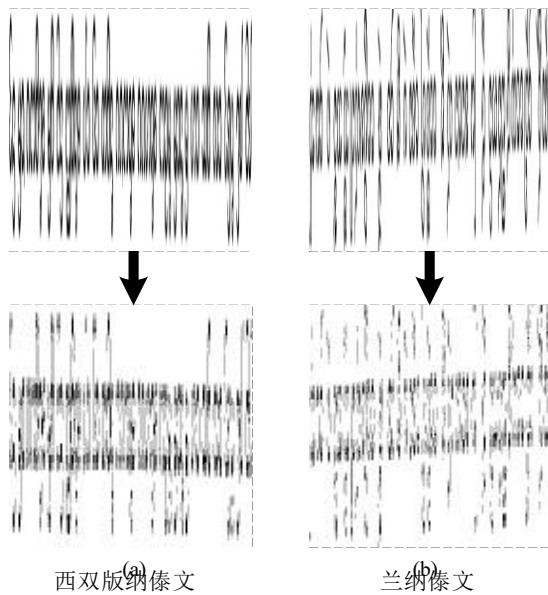

size standardization

n*n

Fig. 3. Size standardization effect of text line image.

firstly, the pictures to be processed into the specified size. Text line images of New Tai Lue and Lanna Dai are larger in width and smaller in height, so the text line images of New Tai Lue and Lanna Dai documents are compressed to n*n size at a fixed ratio based on the shorter side height. The effect of image size standardization in text line of New Tai Lue and Lanna Dai documents is shown in Fig. 3.

To strengthen the only distinction between New Tai Lue and Lanna Dai text line image, the dilation operation in image digital processing technology is applied after size standardization, and the purpose of dilation is to make the primitives of image dispersion of Lanna Dai text line more obvious [9]. As shown in Fig. 4, the difference in texture features between New Tai Lue text line and Lanna Dai text line images after dilation and size standardization is more obvious. Due to the existence of complex tone marks in the text line image of Lanna Dai document, there is obvious area of separation, and the lines are finer and the texture is more complex. Text line images of New Tai Lue documents are simpler and smoother.

西双版纳傣文 (a)   兰纳傣文 (b)

Fig. 4. Dilation processing effect of text line image on New Tai Lue and Lanna Dai. (a) New Tai Lue (b) Lanna Dai

III. TEXT LINE IMAGE FEATURE EXTRACTION

A. Gabor texture analysis

Gabor filter is a linear filter used for texture analysis, which mainly analyzes whether the image has a specific frequency

B. Preprocessing

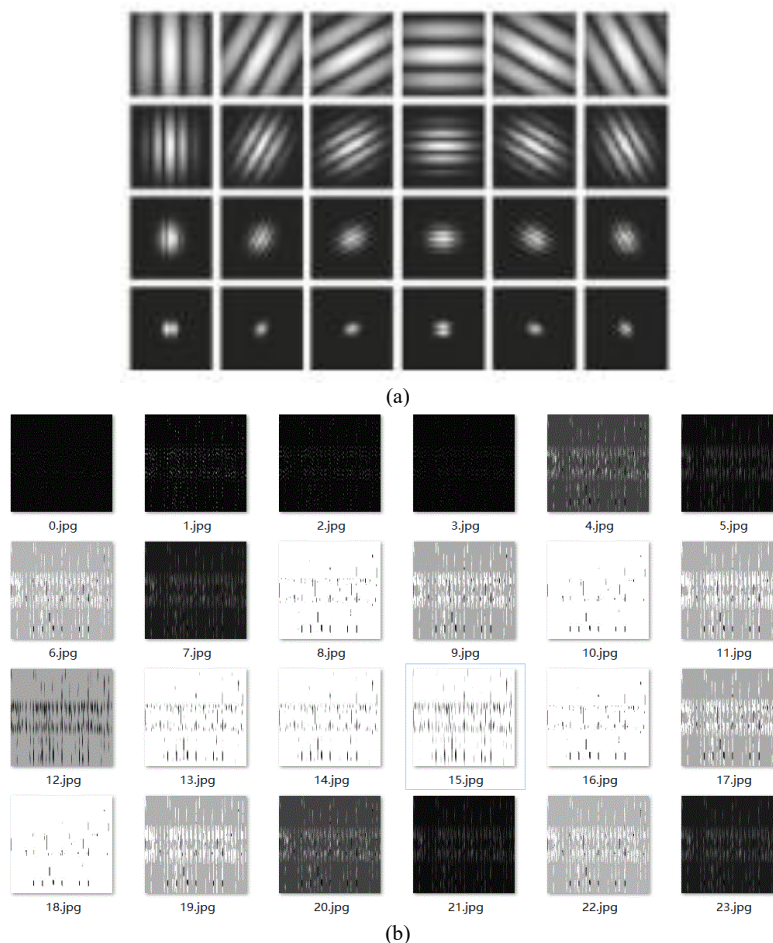In order to facilitate image processing and feature extraction,

(a)

(b)

Fig. 5. Gabor filters/ Image transformed by Gabor filter. (a) Gabor filters (b) Image.

content in a specific direction of a specific region [10]. Gabor filtering is used to analyze the texture features of the preprocessed text line images of New Tai Lue and Lanna Dai documents. By constructing multi-size and multi-direction filters, the filter analysis is carried out in different scales and directions, respectively, and Gabor filter texture analysis image of the preprocessed New Tai Lue and Lanna Dai document text line are obtained. First, Gabor filter banks are established: 4 scales and 6 directions are selected to form 24 Gabor filters, as shown in Fig. 5(a). Gabor filter banks and preprocessed New Tai Lue and Lanna Dai text line images are spatial convolution,

and each image can get 24 filter outputs, which are same size of the original image [11]. As shown in Fig. 5(b), the preprocessed text line images of New Tai Lue and Lanna Dai are images filtered by Gabor texture analysis. When Gabor filter texture acts on image, the closer the frequency of local texture under filter coverage of various scales and directions is to the frequency of filter, the larger the response will be, and vice versa.

B. Statistical features of image entropy

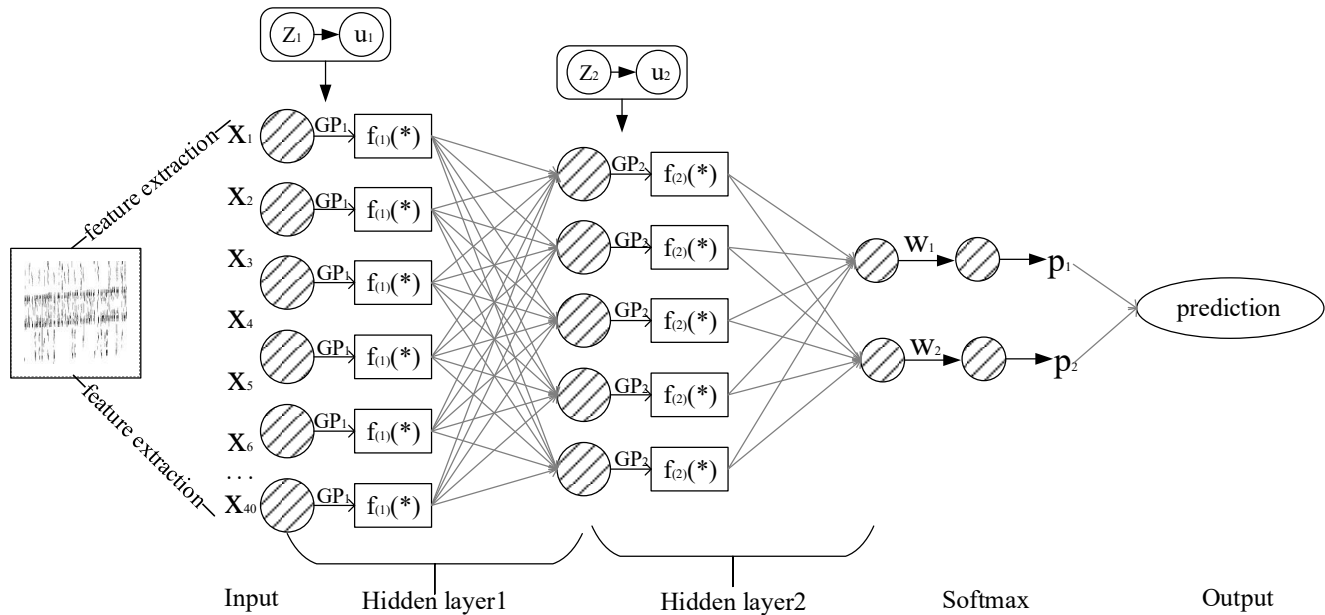Image entropy represents the aggregation features of image
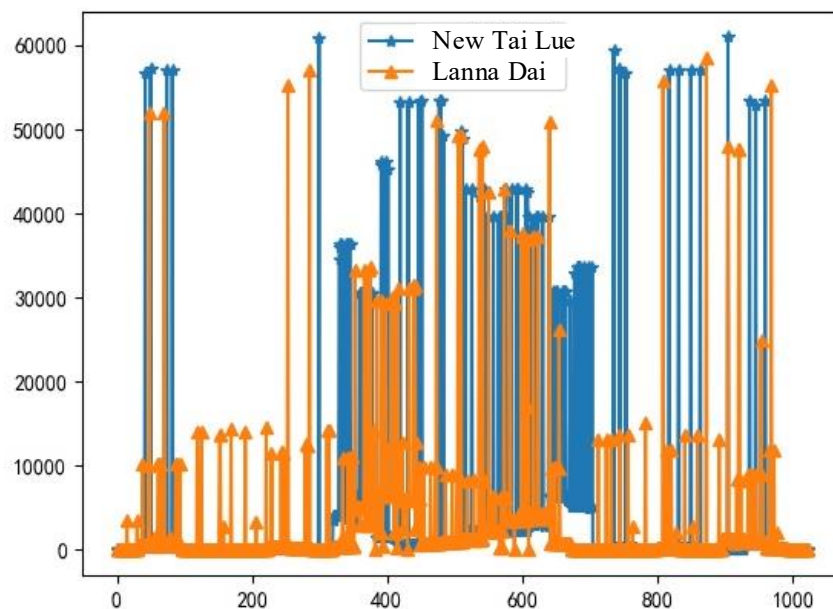

Fig. 6. Text line recognition framework

Fig. 7. Visual image of statistical feature of average row variance.

TABLE I.　　PARAMETER RANGE

Parameter	Range
Batch size	[32, 64, 128, 256, 512, 1024]
Learning rate	[0.00001, 0.00001, 0.001, 0.01, 0.1]

gray distribution, due to the image response information of the text line of New Tai Lue and Lanna Dai documents is different, the average information is different, which can lead to distinction [12]. Therefore, image entropy is one of the statistical features of Gabor filtered texture analysis of text line images of New Tai Lue and Lanna Dai documents. Image entropy is calculated for 24 images generated by Gabor filtering texture analysis [13], which is shown in equation 1.

$$H = -\sum_{i=0}^{255} p_i \log p_i \qquad (1)$$

The one-dimensional entropy of an image can represent the aggregation characteristics of the gray distribution of the image, but it cannot reflect the spatial characteristics of the gray distribution of the image. The feature quantity constitutes the two-dimensional entropy of the image. The neighborhood gray value of the image is selected as the spatial feature quantity of the gray distribution, and the pixel gray level of the image forms a feature binary group, denoted as (h, i), where i represents the gray value of the pixel ($0 <= h <= 255$), i represents the mean gray value of the neighborhood ($0 <= i <= 255$):

$$P_j = f(h,i) / N^2 \qquad (2)$$

The above formula can reflect the comprehensive characteristics of the gray value of a certain pixel position and the gray distribution of the surrounding pixels, where f(h,i) is the frequency of the feature binary group (h,i), N is the scale of the image, the definition The two-dimensional entropy of the discrete image is:

$$H = \sum_{h=0}^{255} p_{hi} \log p_{hi} \qquad (3)$$

The constructed two-dimensional entropy of the image can highlight the comprehensive characteristics of the gray information of the pixel position in the image and the gray distribution in the pixel neighborhood under the premise of the amount of information contained in the image. where p_i is the probability that a certain gray level appears in the image, which can be obtained from the gray histogram. By calculating the image of each text line of New Tai Lue and Lana Dai document, the 24-dimensional Gabor filtered texture analysis image entropy statistical features can be obtained.

C. Statistical features of average row variance

Text line in New Tai Lue and text line in Lanna Dai are very similar in appearance, the main difference lies in tone marks. New Tai Lue has almost no tone marks, while Lanna Dai has many tone marks. It can be observed that the tone marks of Lanna Dai text line only appear at 1/3 upper and 1/3 lower the image, and the pixel distribution of the whole image is more even. In New Tai Lue text line, due to there is no tone marks, there are more blank areas in the 1/3 upper and 1/3 lower image, and the pixel of the whole image are mostly concentrated in the middle. Therefore, based on Gabor filtering texture analysis, the statistical features of average row variance are extracted by using the distinguishing feature between New Tai Lue text line and Lanna Dai text line [14]. The average value of pixel points in each row of the text line image analyzed by Gabor filter texture is calculated, respectively. Each average value corresponds to a feature of a line, and the variance of all the obtained average values is calculated, and the obtained variance is the feature value of the image. Average row variance statistical features can well reflect the fluctuation of pixel characteristics in each row of New Tai Lue text line and Lanna Dai text line images, and record the main difference information [15]. After, each image will generate a feature value, we further generate Gabor eigenvalues of 3 different scales and 8 directions. and each text line image of New Tai Lue and Lanna Dai document can get the statistical features of 24-dimensional Gabor filter texture analysis average row variance. As shown in Fig. 6, it is a visualization of Gabor filtering texture analysis average row variance statistical features of text line images of New Tai Lue and Lanna Dai documents. In the figure, the green color represents the average row variance of the text line image of New Tai Lue, and the orange color represents the average row variance of the text line image of Lanna Dai. It can be observed that there is a significant gap between them.

IV.　THE DGP MODEL OF DAI TEXT LINE CLASSIFICATION

Gaussian process is equivalent to Bayesian neural network with infinite width and single hidden layer [16], while DGP is equivalent to a multilayer neural network with infinite width and hidden layers. The Gaussian process is usually considered to be the Gaussian distribution on the function. It is completely described by the mean and covariance. It is probabilistic and can consider the variance of the predicted data points. It has good prediction performance. In this article, we first assume that a Gaussian process is a priori, where we can use the mean function h(x) and the covariance function $p(x, x^1)$ to represent:

$$g(x) \sim gp(h(x), p(x, x^1)) \qquad (4)$$

At a deeper level, the Gaussian process is an infinite-dimensional multivariate Gaussian distribution, in which any label set of the data set is jointly Gaussian distributed. Therefore, according to the priori of the Gaussian process, the set of training points and test points is a joint multivariate Gaussian distribution, which can be written as：

$$E(x, x^1) = \varsigma_f^2 \exp(-\frac{1}{2\eta^2} \| x - x^1 \|^2) \qquad (5)$$

Where E is the covariance matrix, where its entries correspond to the covariance function of the observations. And writing this way, we can use the training subset for model selection.

$$H^* | X, y, X^* \sim \Lambda(\overline{f}^*, \Sigma^*) \qquad (6)$$

Deep Gaussian process is proposed by Damianou et al. It is a hierarchical extension of Gaussian process. It has the probabilistic properties of Gaussian process and can overcome the limitations of Gaussian process. The depth Gaussian process with t hidden layer extracts the value of the vector value function from the Gaussian process：

$$h = f^t(f^{t-1}(...f^2(f^1(x)))) \qquad (7)$$

DVSGP is a deep model of the extended variational sparse Gaussian process, which also uses the variational sparse framework. In the single-layer sparse Gaussian process, the input and output are X and Y respectively, and the joint probability density can be expressed as

$$x(H, R, e) = x(H|R)x(h|e)x(e) \qquad (8)$$

Among them, R represents the layer output, where e represents the induction point, the first term $x(H|R)$ represents the Gaussian likelihood, and the second term $x(h|e)x(e)$ represents the sparse Gaussian process prior Variational reasoning often produces approximate solutions that replace analytical solutions

$$\eta_{DVSGP} = G_q \left[\log(H|R) \right] - \sum_{l=1}^{L} PQ[r_l \| r_l] \qquad (9)$$

However, layers are not connected by activation function, but by Gauss mapping [17]. The model adopts variational coefficient inference, considering the low number of extracted features, so DGP structure with two hidden layers is designed. The classification model structure is shown in Fig. 7. The input is a group of 1,800 training data composed of 24-dimensional, namely the statistical features after Gabor filtering texture analysis of the text line images of New Tai Lue and Lanna Dai documents, and the output is the text line labels of New Tai Lue and Lanna Dai. The output of Gaussian process with zero mean and covariance function is directly used to simulate the observation response.

In the experiment of New Tai Lue and Lanna Dai text line classification, DGP with 2 to 5 hidden layers has been built. The hidden layer width decreases gradually from the input node to the output node, which is equivalent to completing a new feature extraction, every time the mapping is completed to obtain more advanced and compact features. Commonly used classification functions mainly include Sigmoid function and SoftMax function, in which Sigmoid function will deal with each original output value, respectively, so its results are independent of each other, and the probability sum may not be set as 1, it is suitable

for the classification task of non-mutually exclusive categories [18, 19]. The output values of the SoftMax function are
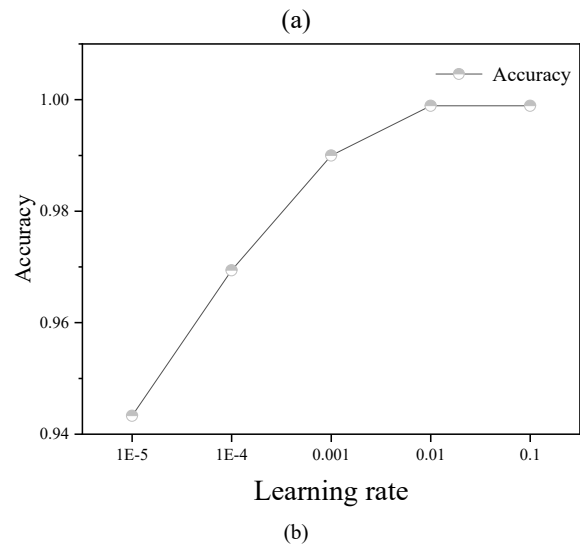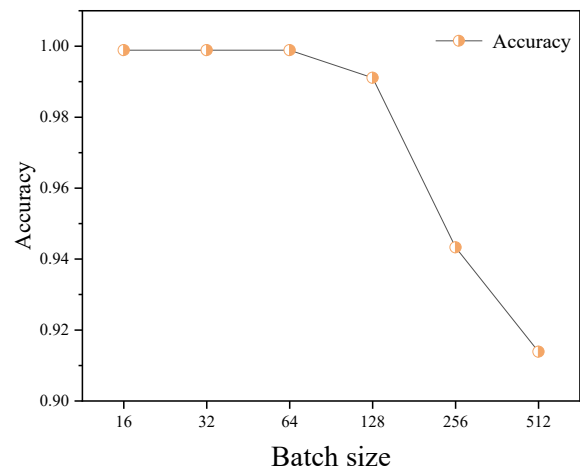
(a)

(b)

Fig. 8. Recognition accuracy with different parameters (a) batch size (b) learning rate
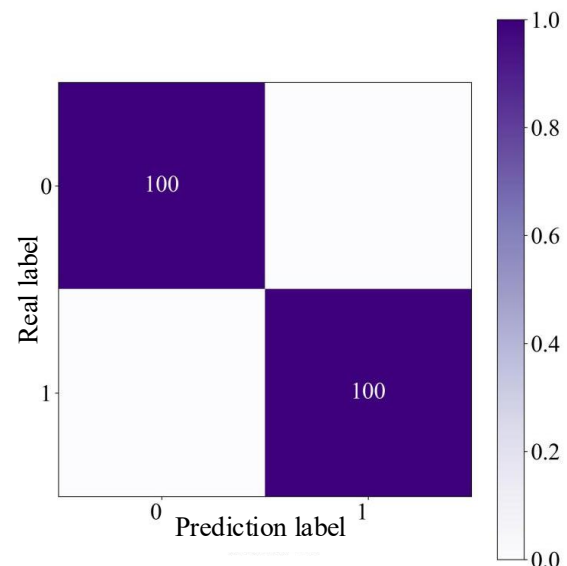

Fig. 9. Confusion matrix for handwritten New Tai Lue character recognition

correlated with each other, and the sum of their probabilities is always 1, which is suitable for the classification task of mutually exclusive categories [20-22]. In this study, the classification category was interdependence, so the output layer adopted SoftMax function, and the number of neurons is 2, text line image categories of handwritten Dai documents of different languages. The weight of each neuron in the SoftMax layer is the output of the last hidden layer of the model, its output is a given input belonging to a class of two, and the class with high probability is the prediction class.

Compared with the deep neural network, the performance of the deep neural network is mainly attributed to the weighted matrix with a large number of parameters, but the parameters often lead to the problem of over-fitting [23, 24]. The Gaussian process is based on the kernel, in which the input and output feature pairs are directly used for training and without adding the weighted matrix. Therefore, the DGP model is composed of multiple Gaussian processes, can reduce the problem of overfitting [25, 26], so the regularization method to prevent overfitting is not added to the model.

V. RESULTS AND DISCUSSION

A. Experimental setup and evaluation indexes

The DGP is trained by minimizing the differential loss of the Evidence Lower Bound (ELBO), and the optimization algorithm is the adaptive movement estimation (Adam). Independent adaptive learning rates are designed for different parameters by calculating the first and second moment estimates of the gradient. The exponential decay rate of the Adam optimizer is set as 0.9 for the first order moment estimation and 0.999 for the second order moment estimation. The training was carried out on the PyTorch platform, the hardware environment was CPU i7-9700K, and the running memory was 32G. The partition ratio between the train set and the test set was 7:3. The performance evaluation indexes included precision rate, recall rate, and F1-score. The parameter range of the experiment is shown in TABLE I.

In the classification problem, to analyze the effect of data and classifiers, evaluation indexes can be used for auxiliary analysis. In this paper, the following evaluation indexes are used to comprehensively analyze and discuss the experimental results.

Precision rate is the probability of the actual positive samples among all the predicted positive samples, which can be expressed by equation (2).

$$Precision = \frac{TP}{TP + FP} \qquad (10)$$

Recall rate is the probability of being predicted to be a positive sample in a sample that is actually positive, which can be expressed by equation (3).

$$Recall = \frac{TP}{TP + FN} \qquad (11)$$

F1-score is a weighted average of the precision and recall of the model. The closer F1-score is to 1, the better the empirical effect is. The evaluation index f1-Score can be expressed by equation (4).

$$F_1 = 2 \times \frac{Precision \times Recall}{\beta^2 \times Precision \times Recall} \qquad (12)$$

B. Experimental results and discussion

DGP model is used to test Gabor texture analysis statistical feature datasets classified by New Tai Lue text line and Lanna Dai text line, and the model is tested in the range of parameters. The experimental results within the parameter range are shown in Fig. 8. When the number of batch size is 16, 32 and 64, the classification accuracy rate reaches the maximum, and then the accuracy rate decreases as the number of batch size increases. Therefore, the number of batch size is set to 64. When the learning rate is 0.01 and 0.1, the classification accuracy reaches the maximum. However, if the learning rate is too high, information will be lost, so the learning rate is set to 0.01. Finally, the learning rate and the number of batch size is set to 0.01 and 64, respectively. Experiments show that the accuracy rate of text line classification of New Tai Lue and Lanna Dai documents based on DGP is 99.89%, the values of precision, recall and f1-score are 1, 0.9978 and 0.9989, respectively. The combination of Gabor texture analysis and DGP model can effectively classify the text line of New Tai Lue and Lanna Dai documents.

The experimental confusion matrix is shown in Fig. 9. The vertical axis scale is the real label of New Tai Lue and Lanna Dai document text line, and the vertical axis scale is the predicted label. The matrix image consists of 2*2 squares, and the values in the squares represent the percentage of correct or incorrect classifications. The squares on the diagonal in the matrix image represent the correct classification, while the rest are the wrong classification. The higher the value in the square, the darker the square. As can be seen from Fig. 9, the diagonal squares of the experimental confusion matrix images are all dark purple, and the proportion of correct classification is very high, which tends to be 100%. The experimental results show that the Gabor method combined with the DGP can effectively classify the New Tai Lue and Lanna Dai text lines.

After Gabor texture analysis of text line images of New Tai Lue and Lanna Dai documents, two recognition features are extracted, namely image entropy statistical feature and average row variance statistical feature.

To compare the effectiveness of these two recognition features, comparative experiments are carried out on different classifiers such as Gaussian Naive Bayes (Gaussian NB), Random Forest (RF), Decision Tree (Tree), Gaussian Process (GP) and DGP. Fig. 10 shows the comparison results of image entropy statistical features and average row variance statistical features of text line in New Tai Lue and Lanna Dai documents. Among them, the accuracy of the average row variance statistical feature surpassed 99.6% in traditional methods, and the highest accuracy rate of the image entropy statistical feature in the traditional method was only 56.19%, and the lowest result was 30.39% on the Tree. Recognition results far below the statistical characteristics of the average row variance. The results show that, compared with the image entropy statistical feature, the precision, recall and f1-score values identified by
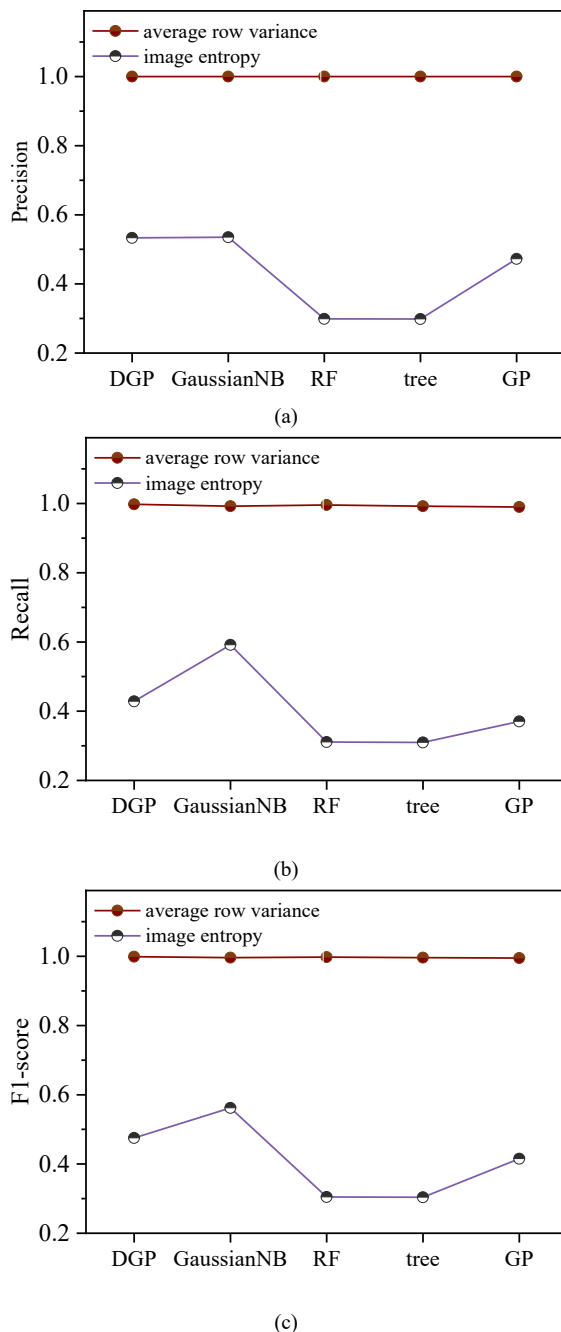
(a)

(b)

(c)

Fig. 10. Comparison of different recognition features. (a) Precision (b) Recall (c) F1-score
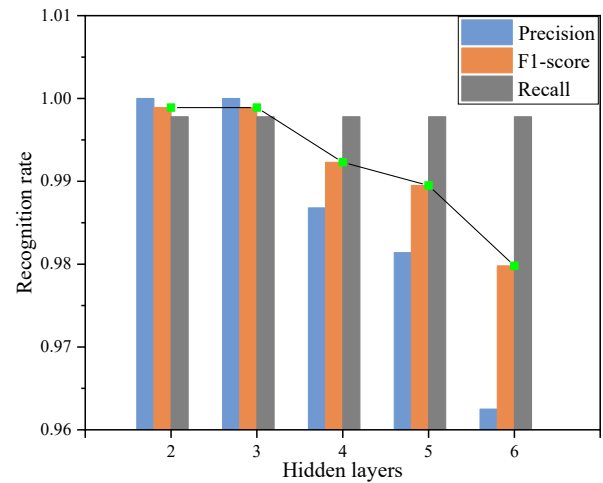

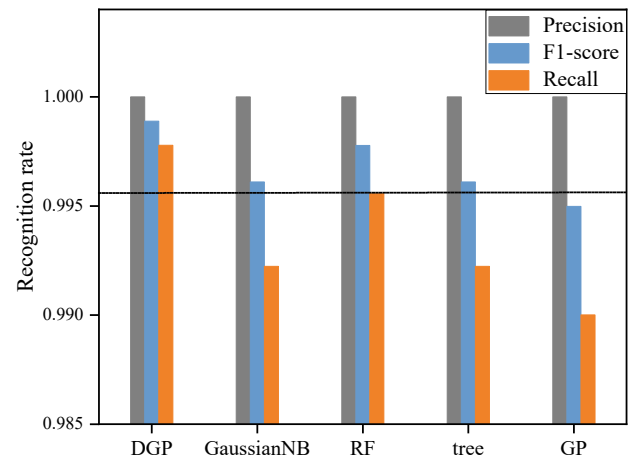Fig. 11. Variation trend of New Tai Lue and Lanna Dai documents text line recognition rate of DGP with depth growth

Fig. 12. Recognition rate comparison with other classification models

the average row variance statistical feature on five classifiers, such as Gaussian NB, RF, Tree, GP and DGP, are all closer to 1. Therefore, the statistical feature of average row variance is more suitable as the feature of text line classification of New Tai Lue and Lanna Dai documents, and the recognition effect is better.

In this paper, DGP model with 2 to 6 hidden layers is built, the performance of classify text lines of New Tai Lue and Lanna Dai documents by DGP with different depths is compared on the statistical features of the average row variance. As can be seen from Fig. 11, the recognition rate of handwritten New Tai Lue text line by the DGP with the number of hidden layers from 2 to 6 shows an overall decreasing trend. When the number of

hidden layers is 2 and 3, the recognition rate is the highest, and the precision, recall and F1-score values are 1, 0.9978 and 0.9989, respectively. When the number of hidden layers in the DGP is 4, the classification results does not increase with the increase of the number of hidden layers. Considering that when the number of hidden layers of the model exceeds 3, the overly complex system exceeds the complexity and nonlinearity of the classification problem itself at this time, leading to the degradation of system performance. Moreover, the two-layer classification efficiency is higher than the three-layer classification, so the two-layer DGP model is more suitable for the New Tai Lue and Lanna Dai text line classification.

To prove the effectiveness of the DGP, a comparative experiment was conducted with traditional classifiers such as Gaussian NB, RF, Decision Tree and GP on the same statistical features of the text line of New Tai Lue and Lanna Dai documents. Fig. 12 shows the recognition rate comparison between DGP and other classifiers. The precision, recall and

f1-score values of text line classification of New Tai Lue and Lanna Dai in DGP are 1.0000, 0.9978 and 0.9989, respectively. The values of precision, recall and f1-score of text line classification of New Tai Lue and Lanna Dai in Gaussian NB are 1.0000, 0.9922 and 0.9961, respectively. The values of precision, recall and f1-score of text line classification of New Tai Lue and Lanna Dai in RF are 1.0000, 0.9956 and 0.9978, respectively. Compared with other methods, the f1-score obtained by our DGP method is 0.28%, 0.11%, and 0.28% higher than Gaussian-NB, RF and Decision Tree Respectively. Compared with traditional classifiers such as Gaussian NB, RF, Decision Tree and GP, the precision, recall and f1-score of DGP are closer to 1, which proves that DGP is more suitable for text line classification of New Tai Lue and Lanna Dai documents which are very close in shape. Therefore, compared with the traditional classifier, the DGP is more suitable for classifying the text line of New Tai Lue and Lanna Dai documents.

VI. CONCLUSION

To solve the problem that the Dai characters of different languages are very similar in shape and the document optical recognition and classification is difficult, a new recognition method of text line for New Tai Lue and Lanna Dai languages is proposed, the method based on the statistical features of texture analysis and the DGP. The database is constructed and the features are extracted from the text line images of two Dai documents, and the DGP classifier is constructed. The results show that this method is suitable for text line recognition of Dai documents, and the comparison results show that it is better than other methods. The next step is to introduce ensemble learning to improve the classification accuracy of documents. The next step is to fuse different statistical feature methods to obtain more suitable features. In addition, ensemble learning is introduced to improve the classification accuracy of documents

References

[1] M. N. Asim, M. U. G. Khan, M. I. Malik, A. Dengel and S. Ahmed, "A Robust Hybrid Approach for Textual Document Classification," International Conference on Document Analysis and Recognition, Sydney, pp. 1390-1396, 2019.

[2] S. C. Kosaraju, M. Msaum, N. Z. Tsaku, et al., "DoT-Net: Document Layout Classification Using Texture-Based CNN," International Conference on Document Analysis and Recognition, pp. 1029-1034, 2019.

[3] K. Wu, H. Fu and W. Li, "Handwriting Text-line Detection and Recognition in Answer Sheet Composition with Few Labeled Data," 2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, pp. 129-132, 2020.

[4] Z. Chen, Y. Wu, F. Yin and C. Liu, "Simultaneous Script Identification and Handwriting Recognition via Multi-Task Learning of Recurrent Neural Networks," 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 2017, pp. 525-530, doi: 10.1109/ICDAR.2017.92.

[5] Y. Huang, Z. Xie, L. Jin, Y. Zhu and S. Zhang, "Adversarial Feature Enhancing Network for End-to-End Handwritten Paragraph Recognition," 2019 International Conference on Documsent Analysis and Recognition (ICDAR), Sydney, NSW, Australia, pp. 413-419,2019.

[6] Q. A. Al-Haija and A. Adebanjo, "Breast Cancer Diagnosis in Histopathological Images Using ResNet-50 Convolutional Neural Network," 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), Vancouver, BC, Canada, pp. 1-7,2020.

[7] Qasem Abu Al-Haija, Noor A. Jebril, A Systemic Study of Pattern Recognition System using Feedback Neural Networks, WSEAS Transactions on Computers, ISSN / E-ISSN: 1109-2750 / 2224-2872, vol 19, pp. 115-121,2020.

[8] X. Fang, N. Han, J. G. Wu, et al., "Approximate Low-Rank Projection Learning for Feature Extraction," IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 11, pp. 5228-5241, 2018.

[9] D. Tohl, J. S. Li, "Improved Image Expansion for Preserving Sharpness without Jagged Edges," TENCON 2018 - 2018 IEEE Region 10 Conference, pp. 1263-1268, 2018.

[10] S. MorenoPicot, F. J. Ferri, M. Arevalillo-Herráez, et al., "Efficient Analysis and Synthesis Using a New Factorization of the Gabor Frame Matrix," IEEE Trans. Signal Process., vol. 66, no. 17, pp. 4564-4573, 2018.

[11] S. Jia, Z. Lin, B. Deng, J. Zhu and Q. Li, "Cascade Superpixel Regularized Gabor Feature Fusion for Hyperspectral Image Classification," IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 5, pp. 1638-1652, 2020.

[12] F. Aulí-Llinàs, "Entropy-Based Evaluation of Context Models for Wavelet-Transformed Images," IEEE Trans. Image Process., vol. 24, no. 1, pp. 57-67, 2015.

[13] Y. Jiao, Y. Niu, L. Liu, G. Zhao, G. Shi, et al., "Dynamic Range Reduction of SAR Image via Global Optimum Entropy Maximization with Reflectivity-Distortion Constraint," IEEE Trans. Geosci. Remote Sensing, vol. 56, no. 5, pp. 2526-2538,2018.

[14] K. Peuwnuan, K. Woraratpanya, K. Pasupa, et al., "Local variance image-based for scene text binarization under illumination effects," 2017 2nd International Conference on Image, Vision and Computing, pp. 798-802, 2017.

[15] A. Rabinovich, A. Feuer and Z. Friedman, "Multi-line transmission combined with minimum variance beamforming in medical ultrasound imaging," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 62, no. 5, pp. 814-827, 2015.

[16] R. L. de Queiroz and P. A. Chou, "Transform Coding for Point Clouds Using a Gaussian Process Model," IEEE Trans. Image Process., vol. 26, no. 7, pp. 3507-3517, 2017.

[17] T. Koriyama and T. Kobayashi, "Statistical Parametric Speech Synthesis Using Deep Gaussian Processes,"

IEEE-ACM Trans. Audio Speech Lang., vol. 27, no. 5, pp. 948-959, 2019.

[18] C. Tsai, Y. Chih, W. H. Wong, and C. Lee, "A Hardware-Efficient Sigmoid Function with Adjustable Precision for a Neural Network System," IEEE-ACM Trans. Audio Speech Lang., vol. 62, no. 11, pp. 1073-1077, 2015.

[19] Z. Qin, Y. Qin, H. Q. Sun et al., "A Novel Approximation Methodology and Its Efficient VLSI Implementation for the Sigmoid Function," IEEE Trans. Circuits Syst. II-Express Briefs, vol. 64, no. 12, pp. 3422-3426, 2020.

[20] Z. Han, Z. B. Liu, C. Vong, et al., "Deep Spatiality: Unsupervised Learning of Spatially-Enhanced Global and Local 3D Features by Deep Neural Network with Coupled SoftMax," IEEE Trans. Image Process., vol. 27, no. 6, pp. 3049-3063, 2018.

[21] D. Zhu, S. Lu, M. Wang, J. Lin, et al., "Efficient Precision-Adjustable Architecture for Softmax Function in Deep Learning," IEEE Trans. Circuits Syst. II-Express Briefs, vol. 67, no. 12, pp. 3382-3386, 2020.

[22] X. B. Li, W. Q. Wang, "Learning discriminative features via weights-biased softmax loss," Pattern Recognit., vol. 107, 107405, 2020.

[23] Y. Jia, H. Zhang, Z. Zhao, L. Ming, "CNN-based encoder-decoder networks for salient object detection: a comprehensive review and recent advances," Inf. Sci., vol. 546, pp 835-857, 2020.

[24] H. T. Qin, R. H. Gong, X. L. Liu, X. Bai, et al., "Binary neural networks: A survey," Pattern Recognit., vol. 105, 107281, 2020.

[25] L. Chai, J. Du, Q. Liu, and C. Lee, "Using Generalized Gaussian Distributions to Improve Regression Error Modeling for Deep Learning-Based Speech Enhancement," IEEE-ACM Trans. Audio Speech Lang., vol. 27, no. 12, pp. 1919-1931, 2019.

[26] M. Kandemir, T. Kekeç, R. Yeniterzi, "Supervising topic models with Gaussian processes," Pattern Recognit., vol. 77, pp 226-236,2018.
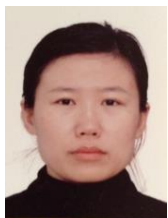
Jingying Zhao received the B.S. and M.S. degrees from School of Computer Science and Technology, Changchun University of Science and Technology, Jilin, China, in 2000 and 2003. Since 2013, she has been an Associate professor with the College of Computer Science and Engineering, Dalian Minzu University. She is currently working towards the Ph.D. degree at Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Liaoning, China. Her current research interests include pattern recognition and machine learning method and their applications.

Da Dong received the B.S. degrees from College of International economy and trade, Dalian Nationalities University, Liaoning, China, in 2018. And received the M.S. degrees from School of Computer science and Engineering, Dalian Nationalities University, Liaoning, China, in 2021.Her current research interests include artificial intelligence, deep learning, pattern recognition and image processing.

Hai Guo received the B.S. in Electronic Engineering from the Heilongjiang University, in 2000, the M.S. in Pattern Recognition and Intelligent Systems from the Kunming University of Science and Technology, in 2004, and the Ph.D. degree in Material Science from the Harbin University of Science and Technology (HUST), Harbin, China. Since 2003, he has been an Associate professor with the College of Computer Science and Engineering, Dalian Minzu University. He has authored over 30 articles in international journals and conference proceedings. His current research interests include pattern recognition and their applications.

Yifan Liu received the B.S. degrees from College of Computer, Weifang University of Science and Technology, Henan, China, in 2019. He is currently working towards the M.S. degree at Computer Science and Engineering, Dalian Minzu University. He current research interests include artificial intelligence, deep learning, and pattern recognition.

Doudou Yang received the B.S. degrees from College of Computer, Pingdingshan University, Henan, China, in 2019. He is currently working towards the M.S. degree at Computer Science and Engineering, Dalian Minzu University. He current research interests include artificial intelligence, deep learning, pattern recognition and image processing.

Author Contributions: Please, indicate the role and the contribution of each author:

Jingjing Zhao was responsible for investigating relevant literature, planning the system, and completing the simulation and optimization work.

Hai Guo is responsible for the construction of the experimental model of the fourth part.

Na Dong used Python to implement the fifth part of the experiment.

Yifan Liu and Doudou Yang were responsible for completing the experimental part.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)