Research On Oracle Image Segmentation and Recognition Based on LBP And R-CNN

Yuxiang Ma *, Suhang Zhu, Jiaxu Liao

School Of Traffic and Transportation, Beijing Jiaotong University, Beijing, China, 100044

* Corresponding Author Email: 22251173@bjtu.edu.cn

Abstract. With the development of today's science and technology, we are constantly exploring the application of computer algorithms and other technologies in the study of oracle bones. In this paper, for the original carrier image of oracle bone, first filtering and denoising to reduce or attenuate the many interference factors in the original topographic image of the oracle bone, and then the local binary mode of feature extraction on the three images to realize the feature enhancement of the image, and then use the support vector machine to optimally classify the image features, so as to achieve the accurate identification of the interference factors, and finally through the feature confusion matrix to prove the accuracy of the oracle bone. After completing the noise processing and classification of the image, the R-CNN algorithm is used to combine deep learning and target detection, iteratively learning the training set and validation set, and automatically adjusting the parameters until the optimal model is derived to realize image segmentation and recognition.

Keywords: Image Segmentation, Text Recognition, Local Binary Battern, R-CNN, Support Vector Machine.

1. Introduction

How to effectively realize the recognition and segmentation of single characters in oracle bone topographies has become one of the important challenges in the current research on intelligent recognition of oracle bones. Lv Xiaoqing et al. of Peking University used Fourier descriptors obtained from histograms of contour line curvature to represent the shape of oracle bone characters and used support vector machines for oracle bone character recognition [1]. Prof. Li Qingsheng of Anyang Normal College performed oracle bone character recognition based on graph isomorphism determination algorithm [2]. Liu Yongge et al. performed oracle bone character recognition based on support vector machine [3]. L. Meng of Ritsumeikan University, Japan, used Hough transform and clustering for the straight-line feature points extracted from the oracle bone character image to be recognized and the template oracle bone image, respectively, and used the corresponding minimum distance for the first level of recognition, and then used the template matching method for the optimization of the recognition results [4]. With the development of deep learning technology, especially the rapid development of target detection technology in natural scene images, this paper considers combining the automatic learning and recognition function of convolutional neural networks to realize the segmentation and recognition of reinforced oracle bones, which can capture the complex and varied strokes and structural features of the oracle bones, which overcomes the limitations of the traditional methods in dealing with the highly complex and varied features to a certain extent, and improves the recognition This overcomes the limitations of traditional methods in dealing with highly complex and diverse features and improves the recognition efficiency.

2. Research content and methodology

2.1. Image preprocessing and feature extraction

2.1.1 Filtering and noise reduction

In order to improve the accuracy of the target text segmentation and obtain better image processing results, we first pre-process the original oracle bone topography to realize the image feature enhancement by denoising. In this section, Gaussian Filter is adopted to reduce the noise of the image.

Gaussian Filter weights all pixel values in the template and assigns different weights to different pixel values according to the law of monotonically decreasing with the distance from the center point. This method overcomes the boundary effect and does not affect the image edges too much, and the filtered image is better [5]. First define a template and set one, calculate the weight matrix of this 2D template with the following formula:

$$(x,y) = \left(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}\right)^{-1} exp$$
$$= \left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_1)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right] \quad (1)$$

The parameter size also has a significant effect on the result of filter denoising. If the template range is large, the noise can be effectively removed, but it will blur the image and lose some feature details; if the template range is small, the original image features can be better preserved, but the denoising effect is not good. After continuous attempts and discovery of the law, we gradually adjusted and set the parameters with relatively better results. After completing the noise reduction process of Gaussian filtering on the oracle image, we transformed the image into a grayscale map with a skeleton map for preliminary feature extraction of the image shown in Figure 1:

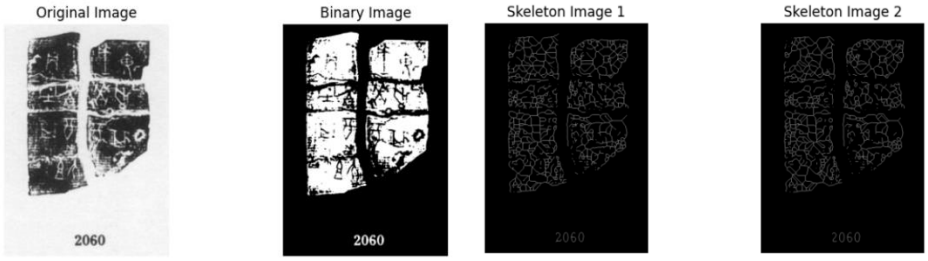

Figure 1. Noise-canceled gray-scale and skeleton maps of oracle bones

2.1.2 Local binary patterns

Local binary patterns are a visual operator used for classification in the field of computer vision, which plays an important role in the extraction and comparison of local features such as image texture features [6].

First a 3*3 window is selected and the gray values of the pixel in the center are compared with those of its neighboring pixels and the comparison results are expressed in binary form. The neighboring pixels are assigned weights according to their positions, and the sum is the LBP value of the pixel in the window. The schematic diagram is shown in Figure 2:

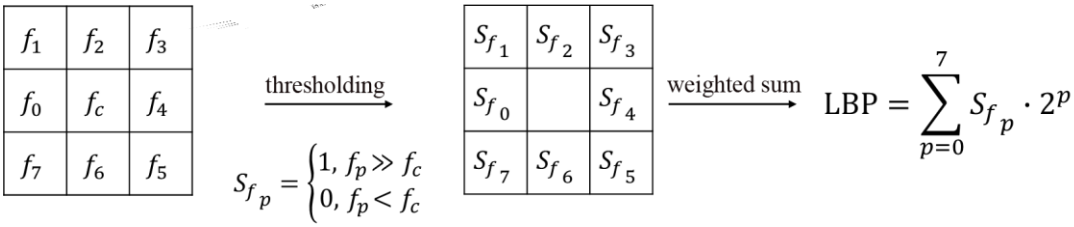

Figure 2. Schematic diagram of the Local binary patterns

Since the local binary model takes the selected local region as the object of study instead of studying a pixel point separately, and can analyze random texture and periodic texture at the same time, it can show the regional texture features more accurately, and has a better effect of feature extraction and recognition. We input all the pixel gray values of the original topographic map of the oracle bone after denoising into this model, that is, we let this local window traverse in the image, and do preliminary texture feature extraction on the preprocessed image.

2.1.3 Support Vector Machines

Support vector machines, developed from the optimal classification surface in the linearly separable case, are a generalized linear classification method for binary classification of data based on statistical learning theory, where the decision boundary is the maximum margin hyperplane solved for the learning samples[7].

For linearly separable training samples , there exists a classification surface that can separate the two classes of samples correctly. And the optimal classification surface can be obtained when the classification interval between the two classes of samples is maximum. The optimal classification function can be found by applying the quadratic programming method as:

$$f(x) = Sgn \sum_{i=1}^{l} y_i \alpha_i^* \, (x_i \cdot x) + b^* \tag{2}$$

As for the nonlinear problem, it is transformed into a linear problem in a high-dimensional space by mapping the input vectors to a high-dimensional eigenvector space, and then searching for the optimal classification surface in the transformed space[8]. The kernel function that satisfies the Mercer condition is defined to correspond to the inner product function in the transformed space, and thus the optimal classification function is:

$$f(x) = Sgn \sum_{i=1}^{l} y_i \alpha_i^* \, K(x_i \cdot x) + b^* \tag{3}$$

On the basis of the image features have been enhanced, and then use the SVM method, can realize the optimal classification of different texture features, improve the classification accuracy, this advantage will also be reflected in the subsequent target text extraction process.

In order to verify the accuracy of the above SVM classification, the confusion matrix can be used for evaluation. Confusion matrix is widely used in evaluating the classification effect, which can visualize the relationship between the real attributes of the analyzed objects and the classification results, and intuitively reflect the performance of a classification model. We find all the feature textures in the topological map and the SVM classification results corresponding to that texture respectively, and determine whether there is classification confusion among them. The precision, recall, and f1 scores of the classification are represented in the form of a matrix, as shown in Table 1:

Table 1. SVM classification feature confusion matrix

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1
Accuracy			1.00	1
macro avg	1.00	1.00	1.00	1
weighted avg	1.00	1.00	1.00	1

The precision column is 1.00, which indicates that all the samples predicted as category 0 by the model actually belong to category 0, i.e., the classification is completely accurate; the recall column is also 1.00, which indicates that all the samples belonging to category 0 are correctly categorized by the SVM classification model; the precision and recall are averaged to obtain the F1 score, and the closer it is to 1, the better the classification effect of the model is. The F1 score in the matrix is 1.00 and the accuracy is also 1, indicating that the classification results of the texture features by the SVM used are in full compliance with the real attributes and have excellent classification performance.

2.2. Oracle Image Segmentation and Recognition

2.2.1 Convolutional Neural Network

Convolutional Neural Network mainly includes convolutional layer, pooling layer, activation function, fully connected layer and output layer.The principle of CNN is "end-to-end" data processing. The convolutional layers are locally connected and share weights with the previous layer to reduce the number of parameters needed to extract features from the network, and use the inter-layer activation function to improve the nonlinear performance of the network[9]. After extracting the features, it is combined with the fully connected layer classification so that the weights of each classifier are continuously improved during the iterative process and gradually converge after obtaining the optimal values.

The structure of convolutional neural network is shown in Figure3:

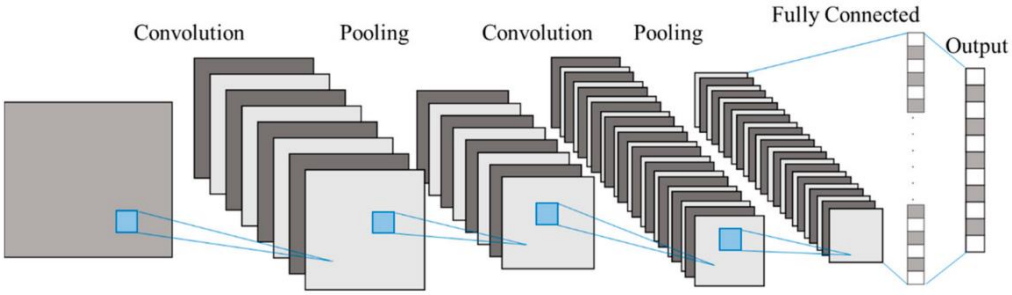

Figure 3. Structure of convolutional neural network

CNN mainly realizes parameter update by forward propagation and back propagation. Among them, forward propagation means passing image features layer by layer from input to output, and using the loss function to reflect the difference between the feature and the real image label. Loss function formula:

$$L = -\frac{1}{m}\sum_{k=1}^{m}\sum_{j} p_k^j \log q_k^j \tag{4}$$

Backpropagation passes the classification error back to the previous layers layer by layer through the loss function, and then adjusts the weights of each neuron node to gradually reduce the model classification error.

2.2.2 R-CNN

R-CNN (Region-CNN) is an algorithm that applies deep learning to target detection based on the algorithmic foundations of convolutional neural networks (CNN), linear regression, and support vector machines (SVMs), etc[10]. The model construction steps of R-CNN are as follows:

(1) Select 2000 candidate region frames from the image according to the selective search algorithm, each candidate region frame may contain target text. All the candidate region frames are uniformly set to a fixed size, and then input them into the convolutional neural network model for feature extraction.

(2) The feature vectors of each candidate region box are input to SVM for classification, and the SVM classifier is able to predict the probability that the image features contained in the candidate region boxes belong to each class. After training the corresponding SVM classifier for each class, the image features in the candidate region box can be classified by the feature vectors.

(3) After the above classifier processing, each region is assigned a corresponding score, but there may be a high degree of overlap between each region box and other region boxes. Here, the non-extremely large value suppression algorithm is used to suppress the non-extremely large value in the local region, so as to more significantly reflect the location of the obvious feature changes, and to reduce the probability of re-detection and misdetection.

(4) At the same time, since the predicted candidate region box and the real box are not completely overlapped, and the intersection ratio between the two is not 1, it is necessary to carry out the border regression to adjust the position of the predicted candidate box, so as to make it closer to the real box, with more accurate detection results[11].

2.2.3 Model Training Process

Firstly, the training set path is set and an instance is created to pre-process the training data such as pixel value scaling, data division, etc. Here, 80% of the data is divided into the training set and 20% into the validation set. A generator for training and validation data is created, i.e., data is loaded from the training set path and training and validation data are generated in batch. Then construct the sequential model, which here consists of a linear stack of five layers: convolutional layer, batch normalization layer, ReLU excitation layer, flattening layer, and fully connected layer. We specify stochastic gradient descent (SGD) as the optimization algorithm to achieve parameter tuning, and choose classification cross entropy as the loss function to measure the model classification performance. The model evaluation metric is accuracy.

Using the training data generator and validation data generator created above, the model is provided with data to start training, and 5 iteration cycles are set.CNN, based on its combination of forward propagation and back propagation, can continuously carry out automatic parameter tuning, so that the loss function decreases continuously, i.e., the accuracy rate increases continuously until the classification accuracy of the validation set is high and stable. The output of the model training results is shown in Table2:

Table 2. Neural network training result

Epoch	accuracy	loss	val_accuracy	val_loss
1/5	0.2526	3.4162	0.3432	2.8462
2/5	0.6377	1.4214	0.4160	2.4569
3/5	0.8235	0.7405	0.4164	2.5038
4/5	0.9241	0.4119	0.4476	2.3958
5/5	0.9622	0.2657	0.4637	2.3606

To make the performance more intuitive, this training result data is visualized as shown in Figure 4 and Figure 5:

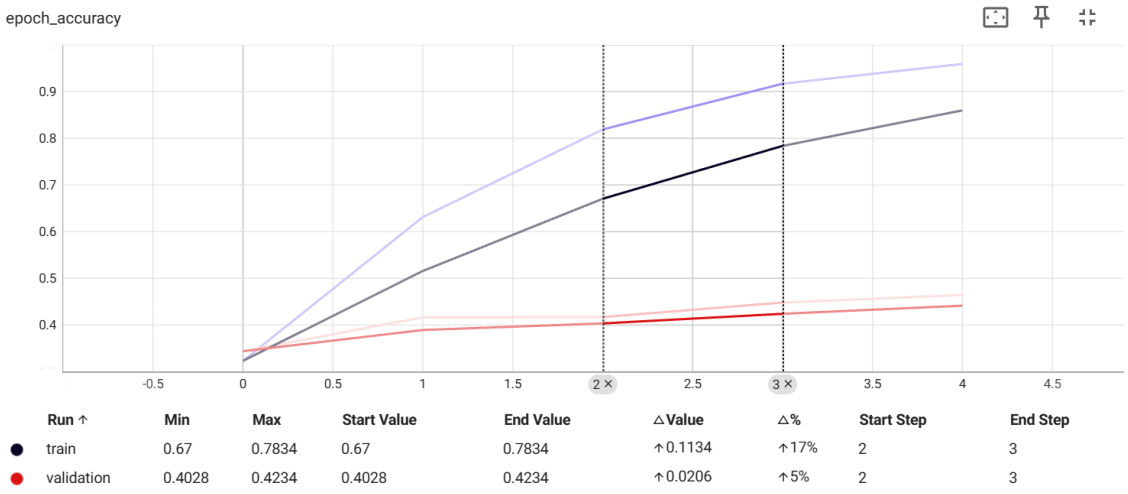

Figure 4. Accuracy for training set (black) and validation set (red)

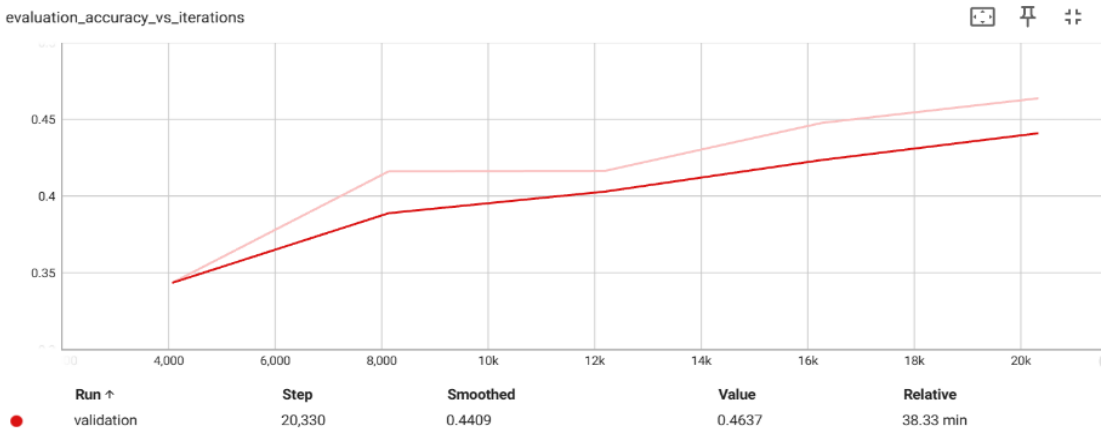

Figure 5. Accuracy changes in the validation set throughout the training process

According to the output results, with the increase of iteration number, the accuracy of the model on the training set gradually increases, and reaches 0.9622 at the end of training. The accuracy on the validation set is also increasing, but it only reaches 0.4637 at the end of training.Since all the training process There is no growth of loss function in the whole training process, so the model is not overfitting. The lower accuracy rate in the validation set here is mainly due to the fact that effective segmentation and single-word extraction of the original topographic image of the oracle bone was not realized before the training of the model, resulting in the dataset containing more interferences, which affects the model's text recognition effect.

3. Conclusions

Text recognition and segmentation of oracle bone has been an ongoing research issue in the scientific community, based on the emerging artificial intelligence technology, this paper proposes a method for oracle bone image segmentation and recognition based on LBP and R-CNN, and the performance of the proposed model is evaluated and analyzed. The experimental results show that the model achieves good results in oracle bone image noise reduction processing, image segmentation and text recognition. The research of this paper can provide new ideas and methods for the study of oracle bones and promote the in-depth development of oracle bone research. At the same time, the method proposed in this paper can also provide reference for other similar image processing and recognition.

References

[1] Lv Xiaoqing, Li Mo'nan, Cai Kaiwei, Wang Xiao, Tang Yingmin. An oracle bone classification method based on graphical recognition [J]. Journal of Beijing University of Information Science and Technology (Natural Science Edition), 2010, 25(S2): 92-96

[2] Li Qingsheng,Yang Yuxing,WANG Aimin. A graph isomorphism approach for oracle recognition [J]. Computer Engineering and Applications,2011,47(08): 112-114

[3] Liu Yongge,Liu Guoying. Recognition of oracle bone characters based on SVM [J]. Journal of Anyang Normal College,2017(2):54-56

[4] Meng L. Two-stage recognition for oracle bone inscriptions[C]// International Conference on Image Analysis and Processing. Springer, Cham, 2017: 672-682

[5] Han Yuxuan, GAO Yin, LI Qiming, et al. Image texture filtering based on direction-adaptive relative Gaussian bootstrapping [J]. Computer Engineering and Design,2024

[6] Pan Hui,Yang Jinrun. Texture image classification based on double refinement local binary mode [J]. Value Engineering,2024,43(19):135-138

[7] Zhangqian Wu,Qing Wang. Segmentation of leaf image based on support vector machine [J]. Software Engineering,2022,25(06): 1-3.DOI: 10.19644/j.cnki.issn2096-1472.2022.006.001.

[8] Shanyi Lin, Ma Yan, Huang Hui, et al. An interactive image segmentation algorithm incorporating multiple histograms and SVM [J]. Computer and Digital Engineering,2024,52(06):1604-1611

[9] Luo Fugui, Song Qian, Qin Yunchu, et al. Research on the application of convolutional neural network based on image recognition [J]. Computer and Information Technology,2024,32(03): 51-54.DOI: 10.19414/j.cnki.1005-1228.2024.03.032

[10] HOU Shijie, LI Yu, LI Xiang, et al. Improved Mask R-CNN insulator detection method based on image segmentation [J]. Industrial Control Computer,2023,36(06):68-70

[11] Yafei Mao. Research on oracle bone character recognition based on convolutional neural network [D]. Harbin Engineering University, 2023.DOI: 10.27060/d.cnki.ghbcu.2023.002991