Journal of Electronics and Information Science (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/jeis.2024.090313
ISSN 2371-9524 Vol. 9 Num. 3

Research on automatic segmentation and recognition of original topographic single characters based on intelligent recognition of oracle bones

Bingbing Hou

School of Mathematics and Statistics, Guangxi Normal University, Guilin, China

Keywords: Machine Vision, Image Preprocessing, K-Means, Yolov

Abstract: The study of oracle bones is of great significance to the understanding of the development of Chinese and foreign civilizations, with the development of artificial intelligence computing, the text recognition of oracle bones has a more efficient method, and the use of machine vision related technology to achieve the text segmentation and text recognition of oracle bone topography can effectively improve the efficiency of the study of oracle bones. In this paper, the establishment of a series of models from the pre-processing of oracle bone topographies, the segmentation of oracle bone text to oracle bone text recognition is investigated. In this paper, we first preprocess the image of oracle bone topographies to eliminate the elements other than topographies, such as numbers and letters in the numbering, etc., and then use commonly used machine vision techniques to filter the image to reduce the impact of interference factors on oracle bone recognition, use image enhancement and image binarization techniques to make the text elements in the image more prominent, and finally use edge extraction techniques to extract the edge information of the text in the topographies.

1. Introduction

Based on the current rapid development of artificial intelligence and big data technology, digital image processing and computer technology are used as the core to solve this problem[1]. At this stage, digital image processing and computer vision technology are utilized to solve the problem of oracle bone topography image segmentation, which in turn separates the independent text areas with distinct and complementary overlapping features from the original topography image. Modern machine algorithms and models can more accurately recognize the morphological features of oracle bone inscriptions, greatly improving the accuracy rate; through interdisciplinary cooperation, the technical and theoretical problems in the intelligent recognition of oracle bones can be better solved; at the same time, in the fields of digital humanities and cultural heritage protection, it provides a new opportunity to promote the application of intelligent recognition of oracle bone inscriptions; however, at the present stage, the recognition research still has limitations. Not only the semantic understanding is difficult, the data quality is insufficient, in the process of oracle bone topography image segmentation, it will also be affected by "point noise", "artificial texture" and "inherent texture" three types of interference. At this stage, the segmentation of oracle bone image topography has certain limitations, therefore, this paper will explore how to accurately segment the independent region from the complex background in the case of many interferences. In the image segmentation of the original oracle bone inscriptions, it will be more difficult to recognize them because they have more variant characters. In this thesis, this paper considers the image preprocessing to deal with the interference elements, and then more accurately recognize the oracle bone inscriptions[2].

2. Oracle image processing and extraction

Due to the serious influence of three types of interference elements such as point noise, artificial texture and intrinsic texture, as well as the influence of the numbering of digits underneath the oracle images, it becomes very difficult to extract the independent and clear text part from the oracle, so this paper needs to preprocess the image.

2.1 Clearing Oracle Numbers

Since each oracle image is numbered underneath, it has to be removed in this paper before extracting the separate and clear text [3]. In this paper, we have used the code to select a quarter height region from the bottom of the image and perform graying, binarization and morphological closure operations on it, replace the processed region back to the original image and display the processed image.

Visualizing this leads to Fig. 1 below.

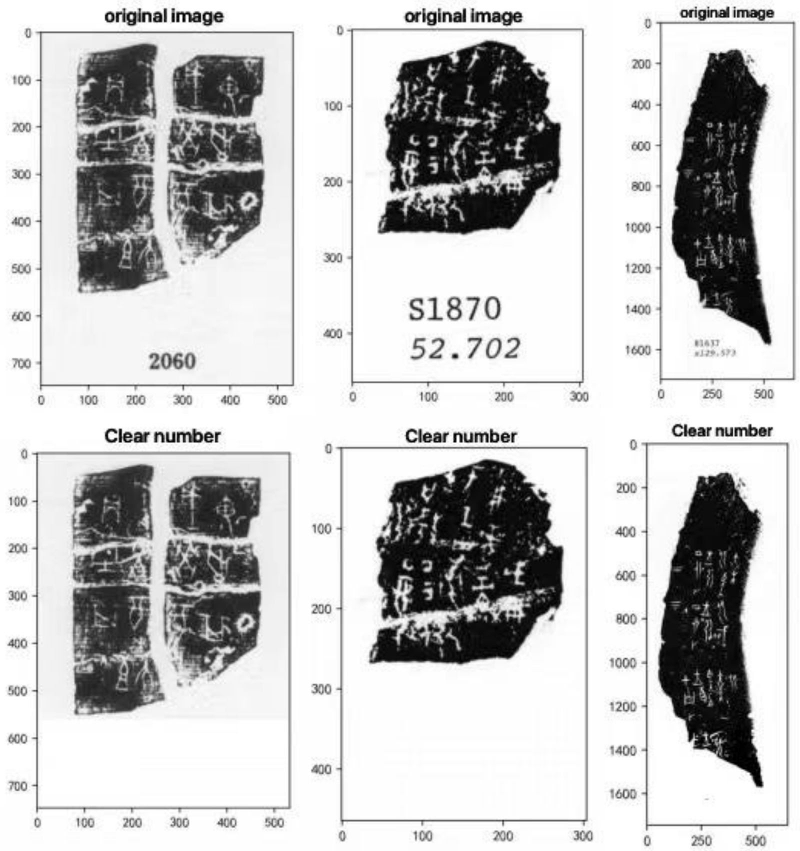

Figure 1: Processed image of the oracle

2.2 Denoising (using Gaussian filtering)

When choosing the filtering method, this paper observes the image and finds that it is a smooth image with a lot of point noise, and this paper needs to keep the details of the image in the process of denoising. Gaussian filter can effectively remove Gaussian noise and keep the edges and details at the same time, so finally this paper chooses to use Gaussian filter for denoising[4].

In this paper, a user-specified template (or convolution, mask) is used to scan each pixel in the image, and the weighted average gray value of the pixels in the neighborhood determined by the template is used to replace the value of the center pixel point of the template. The image is denoised visualized to get the following Fig. 2.

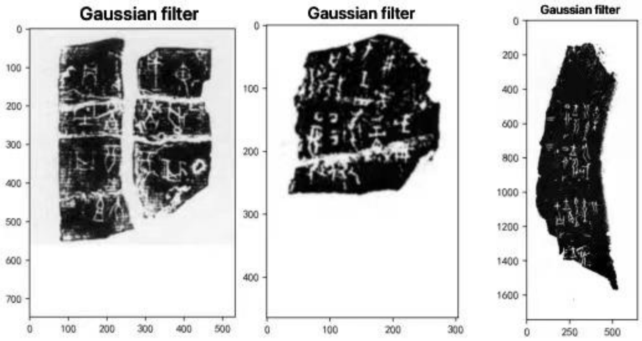

Figure 2: Image after denoising process visualization

2.3 Enhanced images

First of all, Gaussian filter is a linear smoothing filter that effectively removes Gaussian noise from an image. By weighted averaging the image, the value of each pixel point is obtained by weighted averaging the value of itself and the values of other pixels in its neighborhood, thus smoothing the image and reducing the noise[5].Next, enhancement operations are performed to further improve the visual effect of the image and the accuracy of feature extraction. This provides better input data for subsequent tasks such as image analysis, feature extraction or target detection. Since the oracle is affected by artificial and intrinsic textures, this paper chooses to use CLAHE to enhance the contrast of the image and reduce the effects of artificial and intrinsic textures.CLAHE is a contrast-constrained adaptive histogram equalization method, which is mainly used for image enhancement, it divides the image into many small blocks, and then performs histogram equalization on each block and restricts the number of equalized pixels to avoid over-enhancement. This method is mainly used for image enhancement[6].

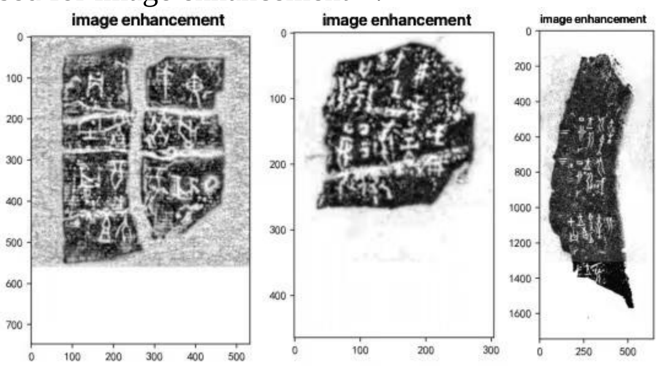

Figure 3: Enhancement of the image using clahe and visualization of the post-image

In this paper, the image is denoised and then enhanced using CLAHE and visualized as shown in

Fig. 3.

2.4 Binarization

Binarization is the process of converting an image to black and white only, highlighting the contours of the target of interest and reducing the amount of data for subsequent image analysis and processing. Binarization is performed on the basis of reduced noise and enhanced detail to ensure the accuracy and reliability of the binarization results[7].

The purpose of using Gaussian filtering followed by image enhancement and then binarization is to remove noise, enhance image details, highlight the target contours, and reduce the amount of data to improve the efficiency and accuracy of image processing. In this paper, the denoising process and clahe enhancement of the image after binarization can be visualized in Fig. 4 below.

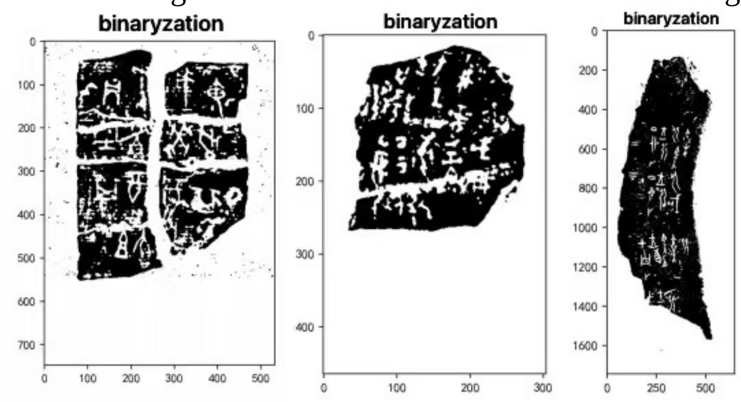

Figure 4: Image visualized after binarization using clahe enhanced image

2.5 Edge Detection

Image edges are very important features in an image, which represent discontinuities in the local characteristics of the image. Edges not only mark the end of a region, but also imply the beginning of another region, therefore, edge features can be utilized for effective image segmentation. In this paper, the direction describes the direction of the edge, while the magnitude reflects the strength and importance of the edge[8].In digital image processing, since this paper deals with discrete data points (i.e., pixels), it is not possible to directly calculate the derivative of a continuous function. Instead, discrete approximations need to be used to estimate the partial derivatives at each pixel point. This usually involves sampling the neighborhood of the pixel point and computing the difference[9].

Specifically, to compute the magnitude of the gradient of an image at the (x,y) position and in the x,y direction, we sample the point and its surrounding pixels and compute the change in their gray values in the x and y directions. In this way, a gradient vector is computed for each pixel point, which contains information about the direction and magnitude of the edge at that point. Thus the gradient magnitude of an image at (x,y) position and in x,y direction is calculated as follows

$$\frac{\partial f(x,y)}{\partial x} = \lim_{x \to 0} \frac{f(x+\varepsilon, y) - f(x,y)}{\varepsilon} \tag{1}$$

$$\frac{\partial f(x,y)}{\partial y} = \lim_{x \to 0} \frac{f(x+\varepsilon, y) - f(x,y)}{\varepsilon} \tag{2}$$

The use of Gaussian filtering is followed by image enhancement and binarization operations and then edge detection, while simplifying the image processing process. The image is visualized as shown in Fig. 5 after a series of processing for edge extraction.

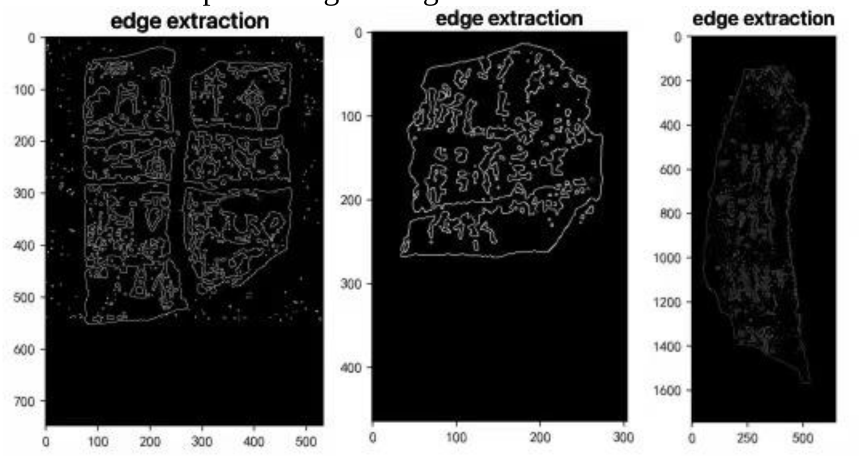

Figure 5: Image after edge extraction and visualization

2.6 Finding Contours with opencv

After the edges of the image are extracted, find Contours in opencv is used to find the contours in the image. Contours are the basis of shape analysis and are commonly used for tasks such as object detection, recognition, and shape analysis [10].The result is visualized as shown in Fig. 6.

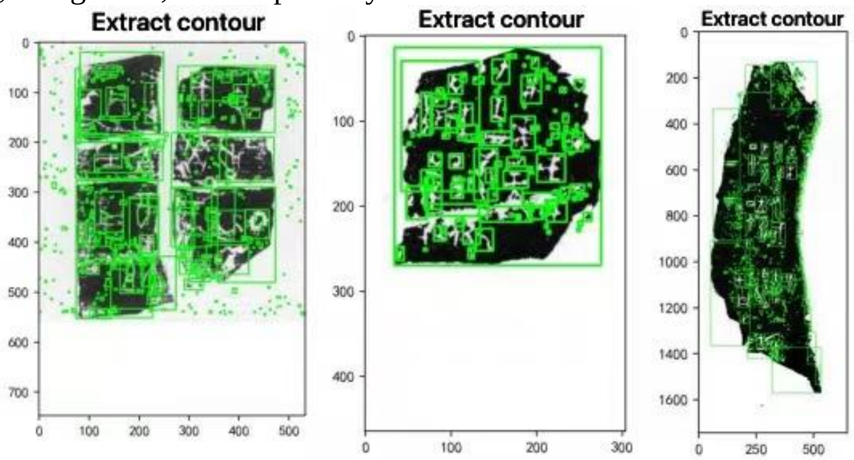

Figure 6: opencv find contour image

2.7 Extracting texture features

For images where the contours have been extracted, the choice is made to use a grayscale covariance matrix to extract the texture features in the contours.

Gray scale covariance matrix is a statistical method used to describe the texture of an image. It is defined based on the relationship between the gray value and the spatial location of pixels in an image, and it is able to extract the key information of an image by counting the condition of two pixels maintaining a certain distance from each other in an image with a certain gray level.

Specifically, the grayscale covariance matrix constructs a square matrix by calculating the number of occurrences of the pair (g1, g2) of the grayscale values of any point (x,y) and its deviation from another point (x+a, y+b) in the image, and normalizing it to the probability P(g1, g2). This square matrix reflects the spatial correlation properties of the grayscale in the image, i.e., the distribution of gray values and positions between pixels. There are five types of grayscale covariance matrix eigenvectors as follows:

(1) Contrast

The contrast feature describes the intensity and frequency of changes in pixel values in an image. High contrast means that there are many significant differences in brightness in the image, i.e., the grooves of the texture are deep and well defined, making the image look sharper and more distinct.

$$Con = \sum_i \sum_j (i-j)^2 p(i,j) \tag{3}$$

(2) Energies

It describes the degree of uniformity of the gray scale distribution of the image texture and the coarseness of the texture. The energy eigenquantity is obtained by calculating the sum of the squares of the values of all the elements in the grayscale covariance matrix. If the values of the elements in the grayscale covariance matrix are more uniformly distributed, i.e., the values of the individual elements do not differ much from each other, then the calculated energy value will be relatively small, which usually indicates that the texture of the image is more detailed.

$$Asm = \sum_i \sum_j p(i,j)^2 \tag{4}$$

(3) Entropy

The entropy value will be maximized if all the element values in the grayscale covariance matrix are equal, or if the pixel values exhibit maximum randomness.

$$Ent = -\sum_i \sum_j p(i,j) \log p(i,j) \tag{5}$$

(4) Inverse variance

If the image texture is more uniform between different regions, the inverse variance will be larger.

$$H = \sum_{i=0}^{N=1} \sum_{j=0}^{N=1} \frac{p(i,j|d,\theta)}{1+(i-j)^2} \tag{6}$$

(5) Relevance

It is used to measure the degree of similarity of the elements of the symbology matrix in the row or column direction, thus reflecting the local gray level correlation of the image. When the elements of the matrix are more uniformly equal, the correlation coefficient is larger, which means that the grayscale values of the images in the specified directions and distances are more similar.

$$Corr = \frac{\left[\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y\right]}{\sigma_x \sigma_y} \tag{7}$$

2.8 Distinguishing between interfering elements and text regions using k-means clustering

By extracting the texture features in the contour using the grayscale covariance matrix, this paper has obtained the texture features of each rectangular box, and then we have to differentiate the interference elements and text regions through the texture features of each rectangular box. In this paper, the paper choose to use the unsupervised learning algorithm Kmeans clustering, and due to the influence of point noise, artificial texture and intrinsic texture three types of interference elements, this paper decides to divide these rectangular frames into four categories, three types of interference elements and one type of text region.

$$E=\sum_{i=1}^{k}\sum_{x\in C_i}\|x-\mu_i\|_2^2 \tag{8}$$

This formula to some extent portrays the closeness of the samples in the cluster around the cluster mean vector, and the smaller the value of E, the higher the similarity of the samples in the cluster. The visualization result is shown in Fig. 7.

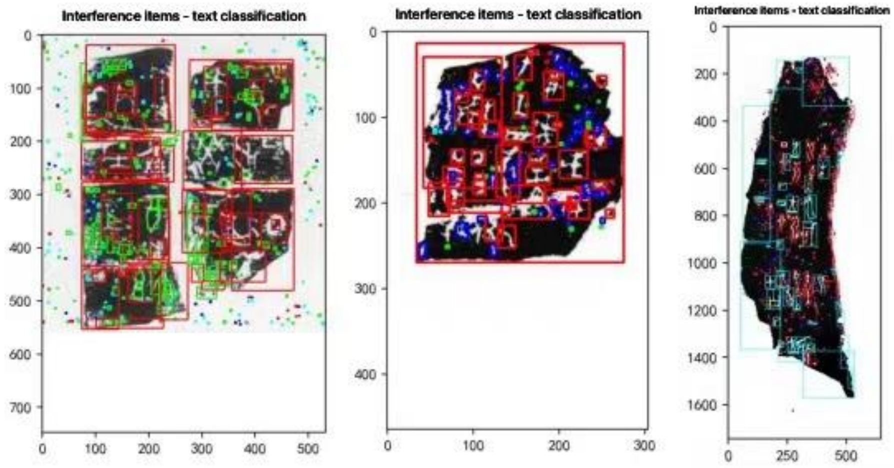

Figure 7: Image after kmeans clustering

3. Conclusions

The image preprocessing model established in this paper can well label the text in the oracle bone topography through the rectangular box, and the three types of interference elements, namely, artificial texture, intrinsic texture, and point noise, are handled better, which reduces their influence on the text segmentation. This paper uses the YOLOv8 model. It adopts the strategy of multi-scale feature fusion, which can effectively capture target information at different scales in the image. This is especially important for target detection in oracle image sets, which may contain oracle characters or symbols with different scales.

In building the image classification model, this paper uses the modified VGG16 model.The VGG16 model has a simple but effective structure. The structure of the VGG16 model is very simple and consists of 16 convolutional layers and 3 fully connected layers. Each convolutional layer uses a small convolutional kernel (3x3) and a smaller step size (1x1). The simplicity of this structure makes the model easy to understand and implement. It is also easy to train and due to the simplicity of the structure of the VGG16 model, it is easy to build and easier to converge to better results. It performs well on the oracle image classification problem and achieves a relatively high accuracy.

References

[1] Liu Guoying. Deep Learning Based Oracle Character Detection and Recognition [A]. Yindu Journal , 2020, 9(41):55-58.

[2] Men Yi, Zhang Shuangshang. Artificial Intelligence-based Oracle Recognition Technology and Glyph Database Construction[J]. Institute of Chinese Characters, 2021, 1(33):10-14.

[3] Luo Ying, Zhang Yifeng, Jiang Jiansheng. Preprocessing Technology for Ray Detection of Negative Defect Images [A].Nondestructive testing, 2024, 2(46):23-27.
[4] Qian Wenjun, Wang Jixin, Du Haoran, Zhang Xunan .Gaussian Filtering of Drill Mast X-Ray Images Based on Various Improved Methods[A]. Laser journal, 2024:2-9.
[5] Ma Pan, Yang Ziheng, Wan Hu, He Shun, Huang Yuan, Xu Shengyong .Cotton Aphid Image Recognition Algorithm and Software System Design Based on YOLOv8 Network. [A]. Journal of intelligent agricultural equipment (Chinese and English), 2023, 3(4):43-46.
[6] Yan Tao, Hu Xiaoping, Peng Qianqian, Huang Hong. Fusion of MSR and Improved CLAHE for Ampoule Image Enhancement[A].Automation and instrumentation, 2024, 5(39):85-88.
[7] Xiang Ling, Deng Xiaohua. Adaptive Evolutionary Model for Binarization of Multiplicative Noisy Text Images[A].Journal of Neijiang Normal University, 2023, 8(38):42-45.
[8] Shaomin Xie, Xinrong Li. Fast detection algorithm for small targets with weak edges in sub-pixel digital images [A]. Modern Electronic Technoiogy, 2024, 13(47):24-26.
[9] Yao F. Q., Zhang X. H., Zhang S., Zhai B.. Experimental analysis of edge detection algorithms for product size images of toothed roll crusher[A].Tianjin Science and Technology, 2024, 6(51):20-21.
[10] Zhenjie Li, Chunhua Zang, Baoyu Su, Xianggui Chen. Design of Image Processing Software Based on OpenCV[A]. Internet of Things, 2024, 6(12):47-49.