Improvement of Ancient Shui Character Recognition Model Based on Convolutional Neural Network

HONGSHUAI ZHAO¹, HAOZHEN CHU¹, YUANYUAN ZHANG², AND YU JIA¹

¹School of Information Engineering, Minzu University of China, Beijing 100081, China
²College of Information Technology, Zhejiang Chinese Medical University, Hangzhou 310053, China

Corresponding author: Yuanyuan Zhang (szm0202@126.com)

ABSTRACT This study uses deep learning theory into the character recognition technology for Shui characters in ancient books, with the objectives of overcoming the instability of the high-pixel ancient Shui characters generative model and the need for large scale handwritten text data annotation among other issues. By constructing a multilayer adversarial neural network with a Laplacian structure, a clear generative model is established for original image data of Shui characters and a stable adversarial network model with multiple mapping relationships from coarse to fine is formed. Based on the analysis of the feature distance of Shui character image samples, the minimum inter-class spacing value and the optimal number of clusters are calculated. Combined with feedback from the classifier model, the optimal number of clusters in the clustering model is adjusted, an evaluation function with information entropy adjustment and clustering threshold convergence is constructed for the unsupervised labelling of Shui character image samples. In this paper, the feedback from the convolutional neural network is used to determine the algorithmic model of the hyperparameters for clustering annotation, and this structure is also designed to improve the recognition rate of handwritten Shui characters in ancient books.

INDEX TERMS Shui characters, generative adversarial network, unsupervised labeling.

I. INTRODUCTION

Shui characters (the characters of the language of the Shui nationality) are a surviving hieroglyph in China, along with the Dongba characters (the characters of the language of the Na xi nationality), their inheritance relies only on oral and handwritten forms, but most Shui characters are illegible and broken. In recent years, the use of advanced information processing methods such as machine learning and big data analysis to break through the traditional digital protection methods of ancient documents and to effectively solve the key problems of image clarity processing, image category labelling, and handwritten character recognition in the process of digital protection of Shui characters and ancient books, along with the promotion of the level of intelligent digital processing of ancient Chinese literature, it has gradually become a research hotspot and an exploration direction of many experts and scholars. Over the years, deep learning theory has made substantial progress in face recognition, natural language processing and other fields [1], [2], some researchers have leveraged deep learning to train more discriminative classifiers for image privacy prediction [3]. However, it faces new challenges in super-resolution image generation and large-scale handwritten text annotation.

This paper focuses on the difficulties and problems that are encountered in the process of Shui character recognition in ancient books in China and focuses on three key technologies: super-resolution image generation, image category labelling and handwritten character recognition. We design an approach combining the relevant research that has been conducted in China for identifying the research context, laying the foundation for the further realization of intelligent digital processing of ancient Chinese Shui books.
Image generation is a future research direction of artificial intelligence and deep learning. Domestic research on adversarial networks remains in its infancy. Adversarial neural network is a new neural network model structure and training model that was proposed by Ian J. Goodfellow in 2014 [4], a generative model and a discriminative model train each other in deception and recognition, the generative model, which is generated via training, is more effective than the traditional image generation algorithm; however, the original adversarial neural network generative model cannot be stably obtained due to the learning rate adjustment problem during the training process, P.J. Burt proposed the Laplacian pyramid structure in 1987, this structure is important for image compression and image fusion, and it is widely used in computer vision [5] and place recognition [6]. Sekiwa Daisuke combined the Laplacian pyramid structure with neural networks and proposed a neural network method for image expansion [7]. In 2015, E. Denton combined the Laplacian pyramid structure with adversarial neural networks and applied the result to handwriting recognition on the MNIST dataset [8]. Inspired by these works, our main contributions of the paper are as follows:

1) We combine the Laplacian pyramid with an adversarial neural network, and increase the stability of the adversarial network with additional information, data clarity processing of the images of ancient Shui books was conducted.

2) We use the unsupervised density clustering algorithm based on information entropy, the character images of ancient Shui books are annotated automatically, which improve the efficiency of pre-sampling.

3) We combine the clustering algorithm with the convolutional neural network. The clustering algorithm will be used as a criterion to improve the character recognition performance.

II. RELATED WORKS

A. THE STABILITY OF ADVERSARIAL NEURAL NETWORKS

On the basis of the adversarial neural network, M. Mirza proposed an adversarial neural network that is based on additional information in 2014. In the process of adversarial network training, the category information of the generated model with labels is provided and the effect of the generated model is improved [9]. To further enhance the stability of the adversarial neural network and to convert the whole network into an unsupervised network, R. Alec proposed a method for training the adversarial network using convolutional neural networks, which moves the generated image from the original global processing into local processing and converts global features into local features, the stability of deep convolution adversarial neural networks is further improved [10]. H. Thanh-Tung proposed a zero-centered gradient penalty for improving the generalization of the discriminator by pushing it toward the optimal discriminator. The penalty guarantees the generalization and convergence of adversarial neural network [11].

B. LABELLING IN IMAGE CLASSIFICATION

Image classification has been dominated by variants of convolutional neural networks [12], since it has high learning capacity and steady performance [13], although the recognition accuracy of the convolutional neural network model in image recognition is very high, it is not competitive enough in some fields like recommendation [14]. It requires many labelled image samples for providing training support, current research on image classification mainly relies on manual labelling [15], [16]. For example, most of the ImageNet databases that are used in image research are manually labelled [11]. As early as 1998, Blum and others proposed a co-training method for sample annotation. This algorithm relies on the support of annotated data, and the accuracy of the labelling must be improved [17]. For the method of collaborative training using labelled samples and unlabelled samples, Lee and others also proposed relevant research algorithms [18]. This method, which is based on cooperative training, depends on labelled samples. Then, in 2005, domestic scholar Zhou Zhihua and others proposed an improved co-training method that is based on multiple classifiers. It has higher recognition efficiency, but requires training with labelled samples, and the operation method is complicated [19]. In 2016, Su Xiangdong proposed a sample labelling method for Mongolian ancient book identification research and realized the preliminary classification of training samples via automated methods and manual elimination [20].

C. LARGE SCALE DATA ANNOTATION IN CHINESE CHARACTERS

Large scale data annotation is an urgent problem in the field of optical character recognition of ethnic minorities in China. Optical character recognition (OCR), as one of the first applications of deep learning, has made great progress since the emergence of convolutional neural networks [21], [22]. For example, for the recognition of Chinese characters [23], handwritten numbers [24], and text in natural scenes [25], it has realized highly satisfactory results. However, in the character recognition of ethnic minority languages, most of the research focuses on Tibetan [26], Uyghur [27], and Mongolian [28] characters, while Shui, Yi, and Dai characters are less frequently recognized. Even in the recognition of Tibetan, Uyghur and Mongolian characters, most methods are traditional methods of feature extraction and classification. The recognition rate cannot be as high as those for Chinese characters or English characters, and deep learning has been less extensively applied in minority language recognition research. Especially for Shui books, the current research mainly focuses on the pre-processing stage prior to identification [23] since deep learning can be used in the recognition of Chinese characters, handwritten numbers and English characters [29] due to the availability of large-scale labelled training data sets. The lack of large-scale labelled data sets is why some minority languages are rarely identified. Therefore, large-scale data annotation is an urgent problem for
deep learning in minority language recognition. Large-scale data annotation requires substantial manpower and material resources; thus, researchers expect to use unsupervised or semi-supervised algorithms to solve this problem. Su Xiangdong proposed a semi-automatic sample selection method for the identification of Mongolian ancient books [20]. This method trains a convolutional neural network with manually classified data. Then, this network is used for large-scale data classification. Finally, misclassified data are removed manually. This method improves the efficiency of large-scale data annotation; however, it still requires two rounds of manual data processing. Bhattacharya and others adopted a similar method. First, they trained a classifier with labelled data. Then, they classified unlabelled data with an offset vector. This method not only improved the utilization of unlabelled data but also required labelled data as an aid [30].

III. PROBLEM DESCRIPTION

For super resolution image generation and large-scale data annotation of ancient Shui characters, this research introduces the theory of deep learning into the image recognition of ancient Shui characters. However, a single goal has difficulty in satisfying the complex requirements of recognition [31], abstraction has long functioned as an important aspect of modeling approaches [32], multiple services are required instead, we need to divide it into several modules. As illustrated in Fig. 1, a stable Laplacian adversarial network generative model is trained via super-resolution image generation of Shui character images by high-pixel Laplacian adversarial networks, we use an unsupervised text image cluster labelling algorithm based on the density and the information entropy is studied for effectively reducing the dependence on manual labelling.

On the image annotation, the method of calculating the image distance based on the Mahalanobis distance metric is investigated. We analyse the relationship between the hyperparameter value and the clustering accuracy in a clustering algorithm with the density and the information entropy. To further improve the recognition rate of convolutional neural networks for text images of handwritten Shui characters, a convolutional neural network classifier recognition optimization method is adopted, we study the learning method of the image annotation target function and the neural network recognition rate optimization strategy, which can improve the data annotation and recognition performance on handwritten Shui characters. These works shown on Fig. 1 focus on the following aspects.

A. ADVERSARIAL NEURAL NETWORKS BASED ON A HIGH-PIXEL LAPLACIAN

The original adversarial neural network model has difficulty forming a constraint function and the “generation - discrimination” mapping relationship has difficulty producing effective and accurate data sampling output due to the widespread distribution of the data in the process of training large-scale data. Incorporating additional image information data while inputting training data can effectively facilitate adversarial neural network training and can overcome the problem of sparse data in the multi-dimensional data model training process, it causes the whole neural network model to be unconstrained. For high-pixel and multi-dimensional images, gradient level fragmentation processing is conducted and the images are transformed through layers step by step to gradually reach low-pixel and low-dimensional layers. The auxiliary information is combined with the original input data to train a stable Laplacian adversarial neural network generative model, thereby generating a stable constraint function.

B. AN UNSUPERVISED IMAGE CLUSTER LABELLING ALGORITHM

A clustering algorithm based on the density peak and the distance is proposed for text and image annotation of ancient Shui characters, which can effectively overcome the problem that the existing classification model requires substantial manual intervention when it labels the Shui character, this method calculates the image distance based on the Mahalanobis distance to realize the feature extraction of image samples. The relation between the hyperparameters and the clustering accuracy in the clustering algorithm, which is
based on the density and the information entropy, is analysed qualitatively to determine the optimal labelling rules.

C. A OPTIMIZATION METHOD FOR CONVOLUTIONAL NEURAL NETWORK CLASSIFIER

To further increase the recognition rate of the convolutional neural network for handwritten Shui characters, we analyse the potential correlation between the accuracy of cluster labelling and the structural design optimization of convolutional neural network. Aiming at overcoming the problem that is encountered in the automatic selection of hyperparameters in the cluster labelling algorithm, this paper investigates how to use the training success rate of the convolutional neural network to construct the objective function for learning the clustering hyperparameters and to improve the annotation accuracy of handwritten Shui characters. For the design of the convolutional neural network structure, structure optimization and parameter selection methods of convolutional neural networks for Shui character recognition are used to improve the accuracy of character recognition.

D. A PROTOTYPE SYSTEM FOR ANCIENT SHUI BOOKS

We establish a prototype system for ancient Shui books that is based on high-pixel, multi-dimensional ancient book image generation, single-character map classification and handwritten character recognition, use it to evaluate the stability of the Laplacian adversarial network while dealing with high-pixel images and evaluating the efficiency of the unsupervised clustering annotation of text or images based on the density and the information entropy. Furthermore, the recognition accuracy of the convolutional neural network classifier that is based on feedback optimized sample annotation is evaluated.

IV. GENERAL ARCHITECTURE

As illustrated in Fig. 2, first, through research on the adversarial network model that is based on the Laplacian pyramid structure and research on the stability of the adversarial network with additional information, data clarity processing of the images of ancient Shui books was conducted. Through research on the unsupervised density clustering algorithm that is based on the information entropy, the text images of ancient Shui books are annotated automatically. By studying the mutual enhancement of the clustering algorithm and the convolutional neural network, we continuously optimize the recognition performance of ancient Shui books and images.

We implement three modules based on the diagram, which differ in terms of the image features on which they are based. These models are based on the following three research aspects:

A. ADVERSARIAL NETWORKS WITH ADDITIONAL INFORMATION

The research on the adversarial neural network structure involves the training and establishment of a generative model. The network structure includes a generative model $G$ and a discriminative model $D$. In this research scheme, both the generative model $G$ and the discriminative model $D$ use the convolutional neural network model as the training neural network, and the back-propagation algorithm is used to adjust the model parameters.

The generative model $G$ uses a single noise vector that is extracted from the distribution $p_{\text{Noise}}(z)$ to add random noise to the original data distribution and creates generated images $\tilde{h}$ according to the random noise. The discriminative model $D$ will use randomness with equal probability $p_{\text{Data}}(z)$ between the real image $h$ from the training data and the generated image $\tilde{h}$ to train the discriminative model.

The discriminative model $D$ will evaluate the sample extraction data of two input images and output a scalar probability value that is between 0 and 1 each time to represent the image authenticity expression from each image generative model. In the initial training process, the probability value of the image data from the real image tends to 1, and the probability value of the generative model data from the generated image data that were synthesized via the expression of random noise tends to 0. The generative model and the discriminative model are trained by convolutional neural networks, and the parameters are adjusted via the back-propagation algorithm. For the adjustment method of the generative model, the probability of generating images through the discriminative model tends to 1, whereas for the discriminative model, the parameter adjustment direction is that the image judgement probability from the generated model tends to 0.

The formula for training the generative model and the discriminative model simultaneously is as follows:

$$
\min_G \max_D \text{E}_{h \sim p_{\text{Data}}(h)} [\log D(h)] + \text{E}_{z \sim p_{\text{Noise}}(z)} [\log (1 - D(G(z)))]
$$

(1)

In addition, auxiliary data $l$ which used to mark the category to which it belongs, and the additional data of the auxiliary generative model and the discriminative model are...
trained simultaneously to increase the stable structure of network model judgement and violate the original one-to-one mapping relationship. The formula for adding auxiliary information is as follows:

$$\min \max G,D \{ E_{\theta_0, I_0} \log D(h, I) + \sum_{z \in \text{Noise}(\theta_1, I_0)} \log(1 - D(G(z, I), I)) \}$$

Laplacian pyramid structure composition scheme:

As illustrated in Fig. 3, input training image $I$ and let $I_0 = I$. The original image size is $j \times j$, $I_0$ is down-sampled to $d(I_0)$. We will obtain one $\frac{j}{2} \times \frac{j}{2}$ image $I_1$, which is the input image of the next layer of the structure. $I_1$ is up-sampled as $u(I_0)$. After smooth expansion, the image pixel size is restored to the original size of $I_0$. $I_0$ is the input image of auxiliary information adversarial neural network. The original training image $I$ and $I_0$ are subtracted to obtain the high-pass image of the original image, which is denoted as $h_0$.

A high-pass image of each layer is acquired by up-sampling again according to the input data of the previous layer.

$$h_k = L_k(I) = G_k(I) - u(G_{k+1}(I)) = I_k - u(I_{k+1})$$

$$h_k = G_k(z_k, u(I_{k+1}))$$

$I_0$ is the input image of the generated model, to which noise $z_0$ is added. It is input to the first layer to generate the model $G_0$. The obtained image is the image that is generated by the generative model $h_0$.

$$I_k = u(I_{k+1}) + h_k = u(I_{k+1}) + G_k(z_k, u(I_{k+1}))$$

The first layer of the discriminative model $D_0$ randomly selects an image from real data and the image that is generated by the generative model is judged by the discriminative model based on equal probability. The discriminative model $D_0$ obtains the image corresponding probability by using the back-propagation algorithm to adjust the discriminative model and the parameters of the generative model and to adjust the adversarial neural network. Then, the next layer of the Laplacian structure combats the training of the neural network. Through the layer-refined adversarial neural network, the fine image can be adjusted on the pixel level to increase the prominence of the details of the image to obtain a super resolution image and a clear and detailed image.

**B. UNSUPERVISED DENSITY CLUSTERING ALGORITHM**

The main objective of the clustering method is to extract the unknown features. Unsupervised labelling of text and image samples was completed by solving for the number of clusters and the threshold value of the clustering algorithm based on the information entropy evaluation method. First, the centre point of the class cluster is surrounded by neighbouring points with lower local density and has a relatively large distance from any point with higher density. In the method, the distance threshold is defined as $d_c$, the distance between the Shui text and image is $d_{ij}$ and the local density is $\rho_i$. $\delta_i$ represents the minimum value of $d_{ij}$. $I$ is the set of the sample points for which $\rho_i$ is greater than $\rho_i$. For $d_{ij}$, the Mahalanobis distance is mainly used to measure the distance between images. For $\rho_i$ and $\delta_i$:

$$\rho_i = \sum_j \chi(d_{ij} - d_c)$$

$$\delta_i = \min_{j: p_j > p_i} (d_{ij})$$

The initial value of $d_c$ is 0. Let $x = d_{ij} - d_c$. If $x < 0$, it is assumed that $\chi(x) = 1$; otherwise, $\chi(x) = 0$. Hence, $\sum \chi(x)$ is the sum of the number of points that represent point $i$ over all sample points for which $d_{ij}$ is less than $d_c$. With the increase of the data set size, the dispersion degree of $\rho_i$ will increase; hence, large data sets will yield higher performance. In this algorithm, $\delta_i$ is the maximum value of $d_{ij}$ of point $i$ to all sample points for which $\rho_j$ is greater than $\rho_i$: $\delta_i = \max_{j: p_j > p_i} (d_{ij})$.

According to the local density $\rho$ and the minimum distance $d$ of the points, we can select the point with the maximum local density and the maximum distance value as the clustering centre and complete the clustering process. Based on the selected $d_c$ value, the local density $\rho$ and the minimum distance $d$ of various sample points can be obtained. The optimal value $d_c$ can be obtained by evaluating the gradient of information entropy decline and the following formula can be obtained:

$$H = -\sum_{i=1}^{n} \frac{c_i}{N} \log \left( \frac{c_i}{N} \right)$$

In the above equation, $H$ is defined as the entropy, $C_i$ represents the number of members of $I$ that belongs to the first cluster, and $N$ represents the total number of sample points. At the initial time, $H$ of the clusters is 0 and each member of the system model is a separate class; hence, the information entropy of the system is maximal, namely, the information quantity of the model is maximal. With the gradual increase of
the selection value \(d_e\), the information entropy of the system will gradually decrease. By analysing the drastic changes in the entropy value of the system under various values, the optimal value \(d_e\) can be determined.

C. A OPTIMIZATION METHOD FOR CONVOLUTIONAL NEURAL NETWORK CLASSIFIER

The classifier structure of the convolutional neural network is optimized by using the inter-class distance \(d_e\) in the clustering algorithm, that is based on information entropy optimization. First, the hyperparameters \(d_e\) of the clustering algorithm and training accuracy of the convolutional neural network are used as parameters to construct a linear regression model:

\[
y = w^T \hat{d}_c + \mathcal{E}, \text{ where } w^T = [w_0, w_1, w_2, w_3] \text{ and } \hat{d}_c = [1, d_e, d_e^2, d_e^3]^T \text{ and } p(\mathcal{E}) \text{ satisfies:}
\]

\[
p(\mathcal{E}) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{9}
\]

After that, the Gaussian likelihood function and the Gaussian logarithmic likelihood function of the objective function are derived, and the maximum value \(y\) in its range is obtained by calculating the gradient and setting the gradient equal to zero. The maximum value corresponds to the optimal value \(d_e\) of the hyperparameter in the density and entropy clustering algorithm. After solving for \(d_e\), the optimal classification result can be obtained. After a small amount of manual sorting and semantic annotation, the result can be used as a training data set for the convolutional neural network. The improvement in the clustering performance will optimize the training of the convolutional neural network and further improve the recognition rate of Shui characters by the convolutional neural network. In hieroglyphic recognition using a convolutional neural network, various achievements have been realized in the structural design and parameter optimization of the network, such as the implementation of handwriting recognition system, and the improvement of the relevant recognition rate [33]–[35].

V. EXPERIMENTS AND ANALYSIS

A. DATASET

There are few datasets for the available Shui characters, we choose to use our own dataset. During the training phase, we collect 6230 manually labelled images as sample input. The dataset images are from ancient books and contain a large number of different fonts. In the Laplacian network, we select a sampling factor of 2 for the high-resolution image downsampling, which is equal to the final upsampling factor. The noise is combined with the super-resolution image to generate the final image, and the process is referenced to Eq. 5. The size of sample images of all Shui characters is 52×52, and they are converted into grayscale images. For the evaluation, a total of 4360 images are used as the training set, and the remaining 1870 images are used as the test set.

B. TRAINING

The whole training process was done on 2 NVIDIA GTX 1080 Tis, we increase the number of samples in the image super-resolution of the adversarial neural network, and augment the training data with random horizontal flips and rotations during image pre-processing. We use a classic SCRNN network in adversarial network module which includes three layers, the size and number of features of each convolutional layer used are respectively (5, 64), (5, 64), (3, 64), each time we choose 32 images as mini-batch when batch training, the experiment uses the RMSProp method to train the network, the initial learning rate is set to 10-4, the objective function of the training generation model and the discriminant model can refer to Eq. 2. After continuously generating samples from the adversarial neural network, we need to automatically label the sample categories. The unsupervised density clustering algorithm is used here to label the image by the Mahalanobis distance as Eq. 6 and calculate the decreasing information entropy of system by Eq. 8. Then, we get an optimal classification threshold, here we set the initial threshold to 0. In the convolutional neural network classifier, we build a three convolutional and pooling layer, as illustrated in Fig. 1, the size and number of features of the kernel in three convolutional layers are respectively (5, 20), (5, 60), (3, 120), and finally the (2, 240) convolution layer and full connection layer is connected. The input here can be the sampled from the training dataset, it adds the generated sample after the automatic image annotation by the unsupervised density clustering algorithm. Finally, the objective function is used to maximize \(y\) and pass the gradient ascent method, the classification network structure can be obtained.

C. RESULTS

We use the adversarial network with the auxiliary information to produce more high resolution samples during the process of continuous iteration, as illustrated in Fig. 4, “Peak Signal to Noise Ratio” (PSNR) is an important indicator for measuring image quality in the field of super resolution, we use it to evaluate the quality of generated images, it fluctuates and finally stabilizes about 27 or so during the experiment.

We also need to evaluate the accuracy of the automatically labelled samples in the whole process, because they are mostly derived from the images generated by the adversarial neural network and have certain character features, results...
shown in Fig. 5 indicate that they grow very fast at begin, eventually stabilizing at 70%.

We find that the PSNR will change the accuracy of automatic image annotation. When the image quality is bad, classification will have more errors.

In the ablation experiment, we use the same training data and design the experimental group including only the adversarial neural network without threshold optimization and only used the classifier without generated sample, in order to show the contribution of individual module to the recognition rate. As illustrated in Fig. 6, we can observe that using the generative adversarial network will realize a higher recognition rate, and using the classifier can converge more quickly to speed up the training process. As the PSNR value and the accuracy of the automatic annotation increase, the integrated method shows a faster growth of the recognition rate.

We believe the reason for phenomenon is that the increased training amount of generated samples speeds up the classification of unsupervised clustering, which causes the neural network to learn the high level features of characters more faster and accelerate convergence. Automatic annotation can help the sample to be directly introduced into the train, there are some problems that the low accuracy of the automatic annotation, it will affect the final recognition accuracy, which will be the focus of our future work.

As summary, the image samples of the Shui characters will have higher resolution through the generative adversarial network of the Laplacian pyramid structure, and the small number of samples is also increased by this step, hence, the model can better generalize the feature in the text set. Then the unsupervised density clustering algorithm of information entropy can automatically annotate the text image for saving labor. The classifier evaluates the optimization results of convolutional neural network structure to continuously optimize the labelling effect of the text. Finally, the convolutional neural network is trained using the labelled Shui characters ancient to effectively recognize the characters.

VI. CONCLUSION AND FUTURE WORK

The adversarial neural network model based on the Laplacian structure focuses on extracting the distribution of image data, it has been applied in the field of image fusion. The adversarial network randomly obtains noise according to the data distribution and uses it to generate images. The overall process encounters difficulties with noise expression, but disordered state of the noise expression can be avoided by incorporating additional informations and they can facilitate image generation by the generative model.

Algorithms based on the density peak and the density distance can use different distance functions for feature recognition applications on various datasets. The information entropy is defined as the probability of discrete random events, it can evaluate the probability of image classification under the correct cluster, which leads to improving the accuracy of feature recognition in the clustering algorithm.

The convolutional neural network uses the gradient descent method to learn the parameters. The updating of the parameters is related to the residual size in the process of gradient descent, if there is noise in the training dataset of the convolutional neural network, the residual and the parameter values are changed in the direction opposite the optimization direction due to a great noise; hence, a large deviation of character recognition rate occurs under the same number of iterations. The training accuracy of the neural network is degraded, and the errors will add up when the noise increases. By analysing the training accuracy of the convolutional neural network under the same number of iterations, we judge whether training dataset is satisfactory or not.

Although the algorithm mentioned in this paper can play a role in ancient character recognition, there are a certain amount of errors from the character and automatic annotation, which let us choose some appropriate labelling strategies. In the future, our work focus on improving the automatic annotation accuracy of generated sample images and contributing to the improvement of some clustering algorithms.

REFERENCES

[1] T. Ching, D. S. Himmelstein, B. K. Beaulieu-Jones, A. A. Kalinin, B. T. Do, G. P. Way, E. Ferrero, P. M. Agapow, M. Zietz, and M. M. Hoffman, “Opportunities and obstacles for deep learning in biology and medicine,” J. Roy. Soc. Interface, vol. 15, no. 141, 2018, Art. no. 20170387.

[2] Y. Weng, T. Zhou, Y. Li, and X. Qiu, “NAS-Unet: Neural architecture search for medical image segmentation,” IEEE Access, vol. 7, pp. 44247–44257, 2019.
[3] J. Yu, Z. Kuang, B. Zhang, W. Zhang, D. Lin, and J. Fan, “Leveraging content sensitiveness and user trustworthiness to recommend fine-grained privacy settings for social image sharing,” IEEE Trans. Inf. Forensics Security, vol. 13, no. 5, pp. 1317–1323, May 2018.

[4] J. F. Goodfellow, J. Pouget-Abadie, M. Mirza, X. Bing, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in Proc. Int. Conf. Neural Inf. Process. Syst., 2014, pp. 2672–2680.

[5] P. J. Burt and E. H. Adelson, “The laplacian pyramid as a compact image code,” Readings Comput. Vis., vol. 31, no. 4, pp. 671–679, 1987.

[6] Y. Yu, L. Chen, J. Zhang, Q. Huang, and D. Tao, “Spatial pyramid-enhanced NetVLAD with weighted triplet loss for place recognition,” IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 2, pp. 661–674, Feb. 2020.

[7] J. Li, Y. Wong, Q. Zhao, and M. S. Kankanhalli, “Learning to learn mixture-of-experts co-trained with labeled and unlabeled data,” Neurocomputing, vol. 299, pp. 248–255.

[8] H. Gao, W. Huang, Y. Duan, X. Yang, and Q. Zou, “Research on cost-driven services composition in an uncertain environment,” J. Internet Serv. Access Mobile Netw., vol. 26, pp. 1–17, Nov. 2019.

[9] A. Blum and T. Mitchell, “Combining labeled and unlabeled data with co-training,” in Proc. 11th Annu. Conf. Comput. Learn. Theory (COLT), 1998, pp. 92–100.

[10] X. Han, A. C. Kak, and S. B. Cho, “Activity recognition with Android phone using mixture-of-experts co-trained with labeled and unlabeled data,” Neurocomputing, vol. 126, pp. 106–115, Feb. 2014.

[11] Z.-H. Zhou and M. Li, “Tri-training: Exploiting unlabeled data using three classifiers,” IEEE Trans. Knowl. Data Eng., vol. 17, no. 11, pp. 1529–1541, Nov. 2005.

[12] J. Li, L. Qi, and Y. Lin, “Research on modulation identification of digital signals based on deep learning,” in Proc. IEEE Int. Conf. Electron. Inf. Commun. Technol. (ICEICT), Aug. 2016, pp. 402–405.

[13] A. Farhat, O. Hommos, A. Al-Zawqari, A. Al-Qahtani, F. Bensala, A. Amira, and X. Zhai, “Optical character recognition on heterogeneous soc for hd automatic number plate recognition system,” EURASIP J. Image Video Process., vol. 2018, no. 1, p. 58, 2018.

[14] S. Rawls, H. Cao, S. Kumar, and P. Natarajan, “Combining convolutional neural networks and LSTMs for segmentation-free OCR,” in Proc. 14th IAPR Int. Conf. Document Anal. Recognit. (ICDAR), Nov. 2017, pp. 155–160.

[15] L. W. Jin, Z. Y. Zhong, Z. Yang, W. X. Yang, Z. C. Xie, and J. Sun, “Applications of deep learning for handwritten Chinese character recognition: A review,” Acta Automatica Sinica, vol. 42, no. 8, pp. 1125–1141, 2016.

[16] I. J. Goodfellow, Y. Bulatov, J. Ibarz, S. Arnold, and V. Shet, “Multi-digit number recognition from street view imaging using deep convolutional neural networks,” in Proc. Comput. Sci., 2014, pp. 1–13.

[17] B. Wang, J. Xu, J. Li, C. Hu, and J.-S. Pan, “Scene text recognition algorithm based on faster RCNN,” in Proc. 1st Int. Conf. Electron. Instrum. Inf. Syst. (EIIS), Jun. 2017, pp. 1–4.

[18] Y. Li, G. He, Y. Wang, and Z. Liu, “Research on printed Tibetan character recognition technology based on fractal moments,” in Proc. 3rd Int. Conf. Comput. Sci. Technol., Jul. 2010, pp. 57–60.