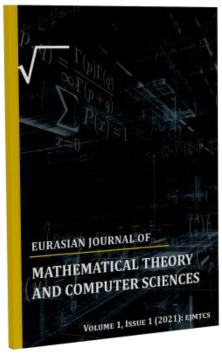
EURASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES
Innovative Academy Research Support Center
UIF = 8.3 | SJIF = 5.916
www.in-academy.uz



METHODS AND ALGORITHMS FOR SEPARATION OF TEXT WRITTEN IN BRAILLE INTO CLASSES USING NEURAL NETWORK TECHNOLOGIES

Akhatov Akmal Rustamovich[1]
akmalar@jbnuu.uz
[1]Doctor of technical sciences of Samarkand State University named after Sharof Rashidov
professor
Ulugmurodov Shokh Abbos Bakhodir ugli[2]
ushohabbos@jbnuu.uz
[2]Doctoral student of the Jizzakh branch of the National University of Uzbekistan named after Mirzo Ulugbek
https://doi.org/10.5281/zenodo.7149006

ARTICLE INFO

Received:27th September 2022
Accepted: 01st October 2022
Online: 03rd October 2022

KEY WORDS

Brayl yozuvi, CNN, kNN, Random forest, SVM, MNIST, Ntree, FC.

ABSTRACT

Introduction. Handwritten text recognition is a type of recognition embedded in character recognition technology, and some basic data processing technologies include the recognition of information written by hand or on a special writing device, such as financial statements, zip codes, braille, and various calculations. In 1998, Lekun et al. LeNetI5 proposed a handwritten digit recognition model widely used to recognize handwritten digits of US bank checks; [2], the K-nearest neighbor (KNN) algorithm was used to achieve a classification error rate of 2.83% on the MNIST dataset. Support Vector Machine (SVM) and its improved algorithms have been widely used in tasks. In 2012, Niu et al. proposed a CNN-SVM hybrid model for digital recognition [4], where CNN is used to extract object features and SVM is used as a classifier. It combined the respective advantages of both sides and achieved good experimental results in image classification tasks;

CNN is a deep learning algorithm that is widely used in many fields such as target detection, image classification, and face detection [6]. CNN learns layer by layer and each layer automatically extracts different features from the input image, which is a very good performance and is one of the leaders of general purpose image recognition systems [7]. Normally, neurons in the convolutional layer are connected to the upper layer through local perceptual fields, and the features of this local area are extracted by convolution, and the secondary features are combined in the convolutional layer. [8].

Convolutional neural networks.

A complex combination of convolutional classes, low-order models, and fully connected layers forms the simplest classical network structure of CNN. The convolutional class layer uses a linear filter kernel to perform linear convolution, and then adds a nonlinear activation function to calculate the printed features..

Figure 1 shows an example of the structure of a classical CNN neural network.

2.1. Convolution layers. Convolutional class layers are the most important component of CNNs, called filters or kernels, which are used to extract low-dimensional features from high-dimensional data. The parameters are a set of trainable convolutional class kernels, each with a relatively small size (length × width) to obtain a suitably sized feature map without losing necessary information. Each convolutional layer has a certain number of convolutional kernels, which is a hyperparameter in CNN, which should be artificially determined empirically, and each kernel is calculated. A feature map means that we have extracted some features of the input image; that is, the original 3D image is converted into a 2D feature map. The combination of all feature maps is our output, which can be used for further feature extraction or as the final feature extraction result. Multiple convolutional class kernels are used to extract different aspects of features such as color, outline, and background.

The depth of a simple neural network is mainly reflected in the deep layers of the network, which leads to a sharp increase in parameters. Compared with the most common fully connected neural networks, the most important feature of the convolutional class of layers is that the parameters can be significantly reduced even by orders of magnitude. If an n×n convolution class kernel performs a convolution class operation on an m×m image with the same image depth, we get a new image of size ((m+n)/l)+1, which is a feature map. λ in the partition is the step size we set for the convolution class. The step size can be changed to other values according to our needs, but if the division is incomplete, i.e. has defects, we will need to fill the image to take an integer number of steps, so the most important thing is that all parameters are possible to divide.

2.2. Connecting layers. A CNN usually alternates the convolutional class with a pooling layer after the layer. Its most intuitive function is dimensionality reduction, so the number of parameters in this layer is also reduced, making computation much simpler and faster, extracting important features, and conforming to invariance. The most common way to do this is to reduce the input using a 2 × 2 filter, where four pixel values are combined into one pixel value. Each max sum operation takes the largest of four numbers (2 × 2 dimensions of the input image). The depth of the image remains unchanged and the image size is reduced while preserving as much of the original information as possible. As shown in Figures 2 and 3, max summation has the advantage of not increasing the number of parameters to be tuned and generally provides more accurate feature extraction and smoother averaging than other methods.

2.3. The process of training. CNN performs supervised training and the process is approximately as follows.

2.3.1. FC layer.

We calculate the output for a fully connected layer λ as a function of (1)..

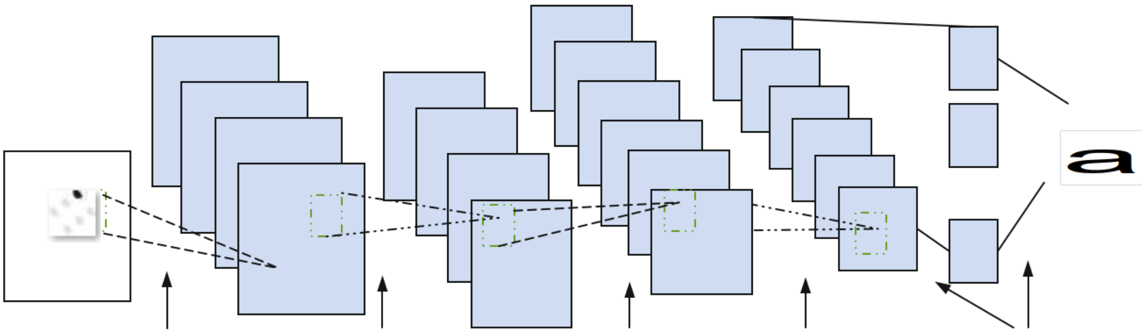
$$x^l = f(u^l), \text{ here } u^l = W^l x^{l-1} + b^l \quad (1)$$

Here f(•) represents the activation function and here we use the sigmoid function.

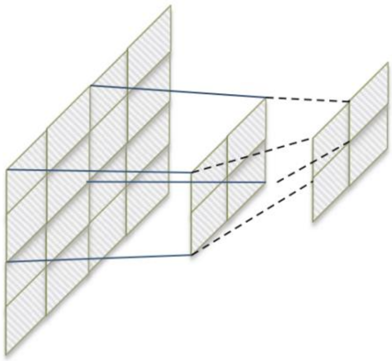
The reduction of the existing error in the experimental sample is as follows:

$$E^n = \frac{1}{2}\sum_{k=1}^{c}(t_k^n - y_k^n)^2 = \frac{1}{2}\|t^n - y^n\|_2^2, \quad (2)$$

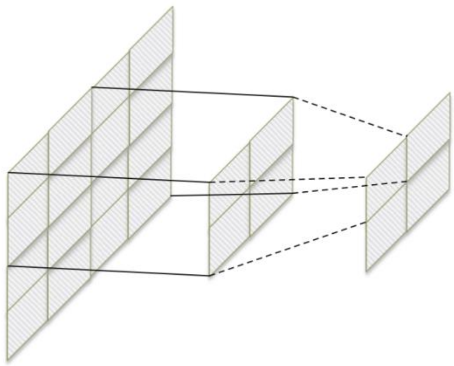
Here, c means that there are a total of c classes in multiclass problems.



1:
LeNet-5 schematic diagram of the model.



2: Maximum merge operation.



3: Average merge operation.

Many methods have been tested on the MNIST dataset, and Table 2 lists the performance of several different methods on the MNIST dataset, including CKELM (Kernel Convolutional Class Extremal Learning Machine), a convolutional neural network with random weights for number extraction. Consists of features and the classifier kernel is replaced by extremum. The MNIST data set has an error rate of 3.20%; The network structure of DAEs is 200 × 200 × 200; CNN-0 is a convolutional neural network with random weight filtering kernels because the weights are not adjusted and the table error rate of CNN-1 is a convolutional neural network after 50 training iterations.

Under the given hardware conditions, CNN takes about 190 seconds per iteration, and 50 iterations takes about 3 hours. RF itself is very fast, taking only 20 minutes to train on the MNIST dataset (when Ntree = 400), and training the hybrid model takes less time than the original RF because the data size is smaller than the original. Our model

EURASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES
Innovative Academy Research Support Center
UIF = 8.3 | SJIF = 5.916 www.in-academy.uz

significantly reduces feature extraction time and accuracy is guaranteed.

Table 1: Error rate (%) for RF and hybrid models in MNIST data set.

Ntree	RF	Hybrid-RF
100	3.02	2.16
200	3.03	2.12
300	2.95	2.06
400	2.94	2.00
500	2.90	1.98
600	2.89	2.02
700	2.85	2.08
800	2.88	2.12

4. Experiments and results.

To evaluate the classification performance of our hybrid model, we conducted experiments on the well-known MNIST and transformed MNIST datasets.

4.1. MNIST dataset. Assuming that Ntree is the number of trees in RF, when Ntree is taken too small, the classification accuracy cannot achieve the desired result. Since the RF is not prone to overfitting problem, we can make the Ntree value as large as possible to ensure the classification accuracy, but it takes a long time to generate the RF, so the Ntree value is important for performance and complexity . - RF. To avoid the long training problem of CNN, we use the method of random weights, after the CNN extracts the features, the extracted features are fed into the RF for classification; For comparison, we conduct experiments under different Ntree numbers. Table 1 shows the test error under different Ntree values.

5. Conclusion.

For the problem of image classification and recognition, we propose a hybrid model in this paper. In the hybrid-RF model, the CNN extracts randomly weighted features and then completes the combined RF classification. Thus, the model greatly reduces the time spent on the feature extraction process, effectively overcomes the problem of long CNN training time, and avoids the problem of manually selecting RF features. Sufficient experimental results show that our proposed hybrid-RF model has high performance and can effectively solve image classification and recognition problems. In the hybrid model, features are extracted by CNN with random weights and then submitted to RF to complete the classification, this model takes much less time to extract features, overcomes the long training problem of CNN, and solves the problems. Disadvantage of manual feature selection by RF.

References:

[1] Akhatov Akmal Rustamovich, Kayumov Oybek Achilovich, Ulug'murodov Shoh Abbos Baxodir ugli. Scientific and theoretical basis of development and introduction of innovative methods in inclusive education 2021// Universum: психология и образование 7 (85) // 46-48. Общество с ограниченной ответственностью «Международный центр науки и образования»

[2] M. Xin and Y. Wang, Research on image classification model based on deep convolution neural network, EURASIP Journal on Image and Video Processing, vol. 2019, no. 1, 11 pages, 2019.

[3] Akhatov, A. R., and Ulugmurodov Sh AB Qayumov OA. "Working with robot simulation using ros and gazebo in inclusion learning." Фан, таълим ва ишлаб чикариш интсграциясида ракамли иктисодиёт истикболлари" республика илмий-техник анжуман, УзМУ Жиззах филиали (2021): 5-6.

[4] E. Kremic and A. Subasi, Performance of random forest and SVM in face recognition, International Arab Journal of Infor-mation Technology, vol. 13, no. 2, pp. 287293, 2016.

[5] Ахатов, А., & Улугмуродов, Ш. А. (2022). Minimum width trees and prim algorithm using artificial intelligence. Zamonaviy innovatsion tadqiqotlarning dolzarb muammolari va rivojlanish tendensiyalari: yechimlar va istiqbollar, 1(1), 141-144.

[6] B. C. Ko, S. H. Kim, and J. Y. Nam, X-ray image classification using random forests with local wavelet-based CS-local binary patterns, Journal of Digital Imaging, vol. 24, no. 6, pp. 1141 1151, 2011.

[7] J. Xia, N. Falco, J. A. Benediktsson, P. Du, and J. Chanussot, Hyperspectral image classification with rotation random forest via KPCA, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 10, no. 4, pp. 1601 1609, 2017.

[8] M. Han, X. Zhu, and W. Yao, Remote sensing image classifi-cation based on neural network ensemble algorithm, Neuro-computing, vol. 78, no. 1, pp. 133138, 2012.