

Research on a handwritten character recognition algorithm based on an extended nonlinear kernel residual network

Zheheng Rao^{1,2}, Chunyan Zeng^{1,2}, Minghu Wu^{1,2*}, Zhifeng Wang³, Nan Zhao^{1,2}, Min Liu^{1,2}
Xiangkui Wan^{1,2}

¹Hubei Key Laboratory for High-efficiency Utilization of Solar Energy and Operation Control of Energy Storage System, Hubei University of technology,
E-mail: zh_rao123@163.com

²Hubei Collaborative Innovation Center for High-efficiency Utilization of Solar Energy, Hubei University of technology, Wuhan, 430068, China;

³School of Educational Information Technology, Central China Normal University, Wuhan, 430079, China;

*Corresponding Author:Minghu Wu
E-mail:wuxx1005@mail.hbut.edu.cn

*Received June 1, 2017; revised July 27, 2017; accepted September 16, 2017;
published January 31, 2018*

Abstract

Although the accuracy of handwritten character recognition based on deep networks has been shown to be superior to that of the traditional method, the use of an overly deep network significantly increases time consumption during parameter training. For this reason, this paper took the training time and recognition accuracy into consideration and proposed a novel handwritten character recognition algorithm with newly designed network structure, which is based on an extended nonlinear kernel residual network. This network is a non-extremely deep network, and its main design is as follows:(1) Design of an unsupervised

This research was supported by National Natural Science Foundation of China (No.61471162, No.61501178, No.61501199, No.61571182); Program of International science and technology cooperation (2015DFA10940); Science and technology support program (R & D) project of Hubei Province (2015BAA115); PhD Research Startup Foundation of Hubei University of Technology (No. BSQD13037, No. BSQD14028); Open Foundation of Hubei Collaborative Innovation Center for High-Efficiency Utilization of Solar Energy (HBSKFZD2015005, HBSKFZD2016002).

apriori algorithm for intra-class clustering, making the subsequent network training more pertinent; (2) presentation of an intermediate convolution model with a pre-processed width level of 2;(3) presentation of a composite residual structure that designs a multi-level quick link; and (4) addition of a Dropout layer after the parameter optimization. The algorithm shows superior results on MNIST and SVHN dataset, which are two character benchmark recognition datasets, and achieves better recognition accuracy and higher recognition efficiency than other deep structures with the same number of layers.

Keyword: unsupervised priori algorithm; “high-capacity” convolution layers; residual network; multi-level quick link; Dropout layer

1. Introduction

At present, handwritten characters are increasingly used in daily life. Handwritten information comes in a variety of different forms, including bills, manuscripts, documents, forms and photographed documents. Handwritten character recognition has wide application prospects, and there is great demand for it in industrial fields such as image recognition systems and handwritten text input devices as society develops and progresses [1]. In addition, more and more researchers in academic fields have begun to pay attention to and participate in studies of handwritten character recognition due to its vigorous development in industries. Especially in the field of image processing and pattern classification, handwritten character recognition has been extensively studied and developed [2]-[6].

Currently, the methods used in handwriting recognition are divided into two categories: handwritten character recognition methods based on traditional feature extraction and pattern classification and handwritten character recognition methods that originate in deep learning.

Character recognition based on traditional features involves three procedures: image pre-processing, character feature extraction [7]-[9], and character classification [10]-[12]. The pre-processing procedure basically uses methods such as greyscale, binarization and size normalization. Character feature extraction mainly relies on traditional manual feature selection from the handwritten characters. The feature extraction is based on pixel-level [13]-[14] primitive features such as colour [15]-[16], structure, texture, projection histograms and moment invariants; once feature extraction has been performed, the appropriate classifier is used to classify the output. The commonly used classification methods are template matching-based, k-nearest neighbours (K-NN) algorithms, shallow artificial neural network algorithms, support vector machine (SVM) [17]-[20], and others. The traditional method has low processing capacity, is overly dependent on manually designed low-dimensionality properties, and has a non-clear model; therefore, the classification results cannot meet the needs of practical applications.

With the increase in deep learning research, the deep network method has been gradually applied to handwritten character recognition. Among these applications, the handwritten character classification method based on a convolution neural network and the character classification method that has its origin in a residual network [21] possess representativeness. The results obtained using these methods are superior to the results of the traditional handwritten character recognition method in terms of recognition effect [22]-[26]. However, the current deep-learning methods also present some issues that have not yet been resolved. (1) With the rapid development of industries and academic fields, the demand for recognition accuracy and recognition efficiency of handwritten character recognition is becoming higher and higher. The current deep-learning method still has room for further development. (2) The deep model network (especially the extremely deep network) parameter training is extremely time-consuming, and the latest research shows that extremely deep networks not the factor that most impacts the overall performance but its use affects the other components of the network. Therefore, whether a structure that can exhibit excellent performance when used on networks of certain depths (non-extremely deep) can be built is the question to be answered.

In view of the issues discussed above [27]-[29], this study presents a handwritten character recognition algorithm based on an extended nonlinear kernel residual network. The use of this algorithm improves recognition accuracy, reduces training time, and partially resolves the conflicting issues of improvement of the recognition accuracy of deep models and training time reduction. First, an unsupervised apriori algorithm for intra-class clustering is designed prior to handwritten character recognition, and the number of clusters is automatically detected based on the characteristics of the dataset, obtain good result values, and reduce the parameter training time to allow the subsequent network training to be more pertinent, the computation is centred on the known clustering centre. In addition, this algorithm also proposes the following three network structure designs in the kernel structure: (1) presenting an intermediate convolution operation model with a pre-processed width level of 2; (2) designing a composite residual structure by adding a multi-level quick link to the newly proposed residual learning building block; and (3) adding a Dropout layer to the newly proposed residual learning building block after parameter optimization. Through experimental tests, it is verified that this algorithm has better character recognition accuracy and higher recognition efficiency than handwritten character classification methods based on convolution neural networks and character classification methods based on residual networks under the same conditions.

In the remaining sections of this paper, the framework for handwritten character recognition is first briefly introduced; then, the innovation and design of this paper are explained, and the handwritten character recognition algorithm based on an extended nonlinear kernel residual network is presented. The algorithm provides a better resolution of categorization issues in handwritten numeral recognition. Next, an experimental plan is

designed to compare the extended nonlinear kernel residual network-based handwritten character recognition algorithm with algorithms based on convolution neural networks and residual networks. The performance comparison resulted in some useful conclusions and demonstrated the perspectives for further research.

2. Handwritten Character Recognition Framework

The traditional handwritten character recognition task flowchart is shown below in [Fig. 1](#). Almost all of its parts rely on the participation and assistance of the designer, the amount of data that can be processed is small, it cannot be applied extensively, and its generalization ability is not outstanding. In addition, its processing features are superficial low-dimensional features, and the information on handwritten characters is not expressed in sufficient detail. The manual intervention required by this method is very intense, and the classification model and classification effect cannot meet the requirements of practical applications.

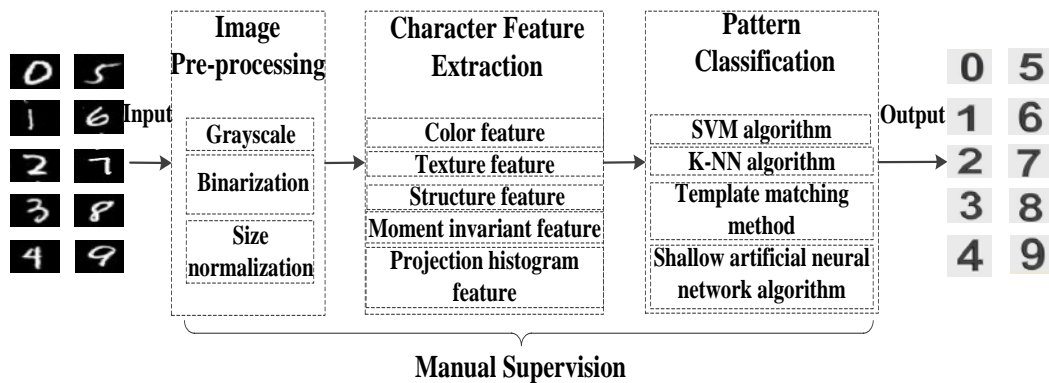


Fig. 1. Traditional handwritten character recognition method

In recent years, the study of handwritten character recognition frameworks based on the traditional “pre-treatment + feature extraction + classifier” has been progressing slowly, and research reports of performance breakthroughs in this area are rare. Nevertheless, the increased use of deeplearning methods has brought new vitality and extremely effective solutions to handwritten character recognition; for instance, the winners of two consecutive International Conference on Document Analysis and Recognition(ICDAR) handwritten character recognition contests all applied deep learning or neural network methods.

A handwritten character recognition task flowchart based on a deep network [30]-[33] is shown below in [Fig. 2](#). It can be seen that the recognition process based on the deep network is more complete with respect to pre-processing, feature extraction and model training and that the structure mainly focuses on the automatic (non-manually adjusted) training of internal parameters with respect to the deep network. When the structural design of the deep

network is completed, the designer only needs to input the handwritten character that is to be detected to the network, and the desired result can be obtained (end-to-end recognition), making the method very convenient.

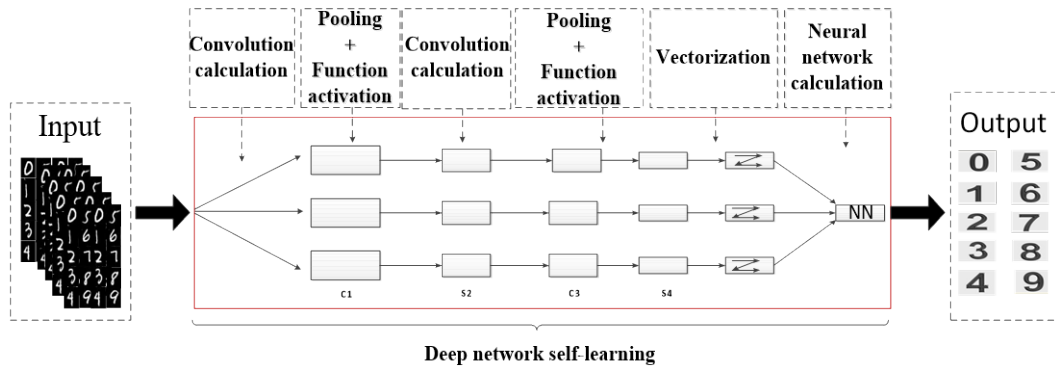


Fig. 2. Handwritten character recognition methods based on deep networks

Deep learning comes from the study of artificial neural networks. The method involves continuously learning, training, and discovering the distributed feature representation of data. The training process of deep learning forms a high-level feature representation that is more abstract by combining the relevant low-level features in a process, which is similar to the process used by human brains in object identification. The network completes the classification based on these high-level features, resulting in a better classification recognition effect. Compared with handwritten character classification based on traditional features, the deep-learning-based handwritten character classification method permits a larger amount of data processing and involves automatic learning, end-to-end modelling, and the use of high-dimensional features to perform classification in a process, which is similar to that used by the human brain; thus, the advantages of high classification accuracy are particularly evident.

However, as the development and study of identification applications continues in industries and academic fields, the requirements of enterprises and research institutions for handwritten character recognition have become increasingly stringent. With regard to current development, the deep learning structure has some room for development in the recognition accuracy and recognition efficiency of handwritten character recognition applications.

3. Handwritten Character Recognition Algorithm Based on an Extended Nonlinear Kernel Residual Network

There are some issues that urgently need to be resolved for existing deep learning methods. (1) The network parameter training of deep models (and that of extremely deep networks in particular) consumes large amounts of time. The latest research indicates not

only that the use of an extremely deep network is not the factor that most affects overall performance but also that it influences the other components of the network.(2) With the rapid development of industries and academic fields, the requirements for recognition accuracy and recognition efficiency in handwritten character recognition have become increasingly higher. Therefore, the deep-learning-based method awaits further development. This study attempts to construct a non-extremely deep structure that has superior performance to that of deep structures under the same conditions. For this reason, this paper presents a framework that uses a handwritten character recognition algorithm based on an extended nonlinear kernel residual network. Its structural design is shown in Fig. 3. The characteristics of this algorithm are as follows:

- (1) Data pre-processing: the design of unsupervised apriori knowledge is used for clustering. The number of clustering centres is automatically detected based on the characteristics of the dataset, producing n known clustering centres. Intra-class optimization clustering is performed around the known centres during operations, thus ultimately yielding good result values and a reasonable parameter training time.
- (2) Convolution neural network framework for the extended nonlinear kernel: the feature extraction and morphological classification of handwritten characters are performed by means of sparse connection of the convolution neural network and weight sharing. This method can use spatial relationships to reduce the number of parameters that must be learned, thus improving the training performance of the general forward back-propagation (BP) algorithm.
- (3) The extended nonlinear kernel model: to improve the result of the model, a composite residual structure is introduced by means of design a multi-level quick link, the convolution operation model is presented with a width level of 2, and a Dropout layer is added after parameter optimization, thereby yielding good test results while improving the training efficiency.

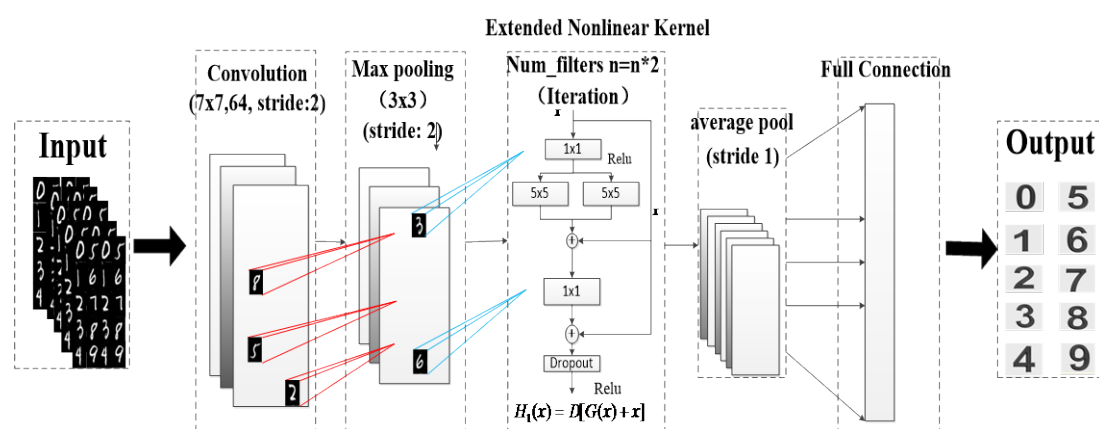


Fig. 3. Framework of the handwritten character recognition algorithm based on an extended nonlinear kernel residual network

3.1 Data pre-processing

This study introduces an unsupervised clustering algorithm for pre-processing experimental data in the proposed handwritten character recognition algorithm based on an extended nonlinear kernel residual network. Before entering the parameter training network, the number of clusters n is automatically detected (n is the number of categories of known datasets) according to the demand, and the clustering operation is performed to generate n known clustering centres. Intra-class optimization clustering is performed around known centres during operations (non-traditional clustering methods are searched throughout the data sets); the goal of these operations is to obtain good result values while greatly reducing the training time for deep network parameters.

The algorithm includes four steps, including sample allocation and intra-class updating of clustering centres.

Step 1: Initialize the clustering centres $c_1^{(0)}, c_2^{(0)}, \dots, c_n^{(0)}$, and select the representative sample of n (known) categories as the initial clustering number and clustering centre according to the characteristics of the data set. (do not have to select clustering centres through a large number of iterations)

Step 2: Assign each sample x_j to the adjacent clustering set. The sample distribution is based on

$$S_i^{(t)} = \{x_j \mid \|x_j - c_i^{(t)}\|_2^2 \leq \|x_j - c_p^{(t)}\|_2^2\}, \text{ in which } i = 1, 2, \dots, k, \quad p \neq j. \quad (1)$$

Step 3: Update each clustering centre within the corresponding category according to the distribution result of step 2:

$$c_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j \quad (2)$$

Step 4:

$$\|c_i^{(t+1)} - c_i^{(t)}\|_2^2 < \varepsilon \quad (3)$$

If the iteration reaches the maximum number of iterations or the difference between two adjacent iterations is less than the present threshold ε (i.e., Equation (3) is satisfied), then the algorithm ends; otherwise, repeat step 2.

When the clustering algorithm designed in this paper is used, the operation design of the updated clustering centre of step 3 is clustered in a certain category (class). In updating the clustering centre of the entire data set, the parameter training time is greatly shortened in comparison with other clustering algorithms.

3.2 Convolution Neural Network Framework for the Extended Nonlinear Kernel

The convolution network structure is used to construct the algorithm framework, and the initial values of the parameters in the network are adjusted appropriately according to the requirements of the extended nonlinear kernel model. The weight sharing of the convolution structure is used to reduce the complexity of the network model and decrease the number of weights, in a manner similar to the network structure of a biological neural network. The algorithm used here introduces this structure; the advantage is more evident when the input of the network is a multidimensional image, as the image can be directly input into the network to avoid performing the complex processes of feature extraction and data reconstruction in the traditional recognition algorithm. In addition, the convolution network has a high degree of invariance for translation, scaling, tilting, and other types of deformations.

3.3 The Extended Nonlinear Kernel Model

This paper presents a handwritten character recognition algorithm network based on an extended nonlinear kernel residual network. The main network design innovations of the extended nonlinear kernel model are as follows:

- (1) presenting the intermediate convolution operation model with a pre-processed width level of 2;
- (2) presenting a composite residual structure that designs a multi-level quick link;
- (3) adding a Dropout layer after the parameter optimization.

Furthermore, in practical application, taking the cost into consideration for the computational optimization of the kernel structure, the two convolution layers $n \times n$ used in the network are discarded and replaced by the new kernel learning building block $1 \times 1 + (n \times n + n \times n) + 1 \times 1$. As shown in [Fig. 4](#) below, the middle double-channel “high-capacity” convolution layers of the new kernel structure are first decreased in a 1×1 convolution layer of reduced dimensions and then restored under another 1×1 convolution layer, thereby maintaining the same level of accuracy and reducing the amount of calculation required. The optimization results are shown in [Fig. 4](#); the left portion of the figure shows the two $n \times n$ structures used by the original network, and the right portion of the figure displays the optimized new kernel learning building block.

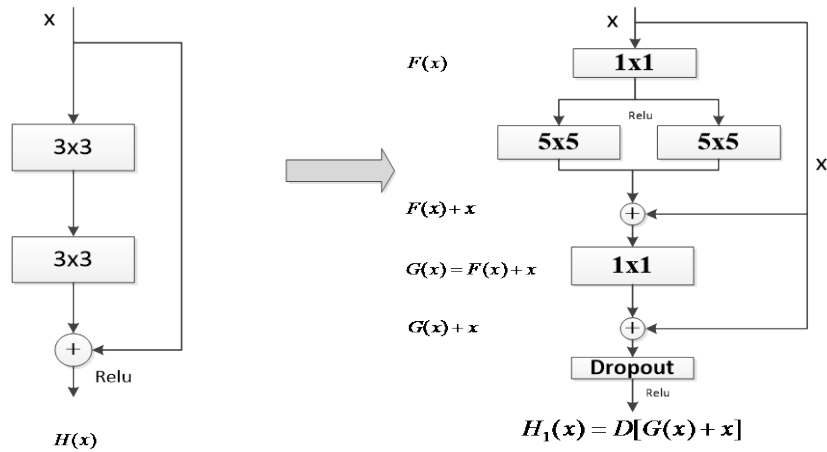


Fig. 4. Comparison of the original kernel learning building block and the new kernel learning building block

This study presents an extended nonlinear kernel model based on the proposed design plan shown above. The structure of the framework is shown in **Fig. 5.**

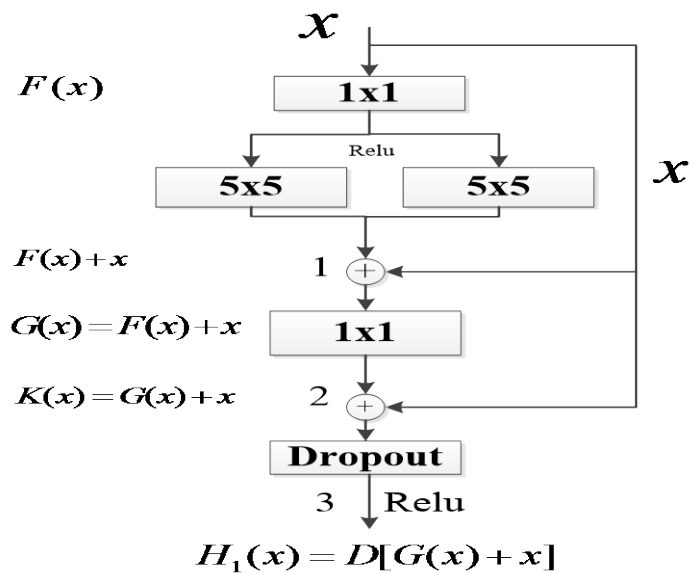


Fig. 5. Extended nonlinear residual kernel structure framework

The description of the formulae used in **Fig. 5** (the input of the first layer of the extended nonlinear residual kernel is expressed as x) is as follows:

Node 1:

$$G(x) = F(x) + x \tag{4}$$

in which $F(x)$ is the normal deep network structure and x is the first-level quick link network.
Node 2:

$$K(x) = G(x) + x \quad (5)$$

in which $G(x)$ is the normal depth of the network structure and x is the second-level quick link network.

Node 3:

$$H_1(x) = D[G(x) + x] \quad (6)$$

in which $D(x)$ is the Dropout layer operation and x in $D(x)$ (the setting of x is generally in a proportional form) can be adjusted according to the settings of different parameters.

$G(x) + x$, corresponding to Node 3, is the normal deep structure.

Suppose to remove Node 2 from **Fig. 5**, only taking Node 1 into consideration (or remove Node 1 and only consider Node 2); the structure can then be seen as a simple residual learning module which has a single-level quick link, and it can be defined as

$$y = D[F(x, \{W_i\}) + x] \quad (7)$$

in which x and y take the input and output vectors of the layers into account and $D(x)$ is the operation of the Dropout layer. The function $F(x, \{W_i\})$ represents the learned residual function (the function $F(x, \{W_i\})$ represents multiple convolution layers). There are three convolution layers, as shown in **Fig. 5**, in which σ in $F = \sigma\{D[W_3[W_2\sigma(W_1x)]]\}$ represents an ReLU function. To simplify the annotation, the extended nonlinear kernel model ignores the offset in the description. The operation $F+x$ is completed by the quick link and by element-by-element addition. The network uses a second non-linear characteristic after the increment (e.g., $\sigma(y)$) as shown in **Fig. 5**.

Assume that each layer function increases the input variable by one dimension. Then, a single-level quick link, such as the application of $x+b$ (b is a constant), can be used to approximate the actual complex output function. Because the use of a composite quick link is equivalent to the use of x^2+x+b (b is a constant) to approximate the actual complex output function, the error in the network approximation of complex functions is reduced, thereby enhancing the results of learning on the characteristics of the non-extremely deep network.

In Equation (7), the dimensions of x and F must be the same. If the dimensions are not the same (for instance, if the input and output channels change), the dimensions can be matched through the quick-link linear projection W_s :

$$y = F(x, \{W_i\}) + W_s x \quad (8)$$

In summary, the handwritten character recognition algorithm based on the extended nonlinear kernel residual network is advantageous in that it presents an intermediate

convolution operation model with a pre-processed width level of 2, optimizes the middle convolution operation, widens the network structure, and increases the computing weight so that the ability to express a complex function is enhanced for network structures with a certain depth. However, the degree of widening the network structure must be appropriate; if the network structure is widened excessively, the number of saddle points and local solutions in the network will increase dramatically. Through experimental analysis and comparison, this study selected the model with a width level of 2 for the simulation experiment. At the same time, in the proposed extended nonlinear kernel model, the composite residual structure, which designed a multi-level quick link, was adopted. Formally, we set $H_1(x)$ as the required basic mapping and put forward composite residual mapping $\begin{cases} F(x)+x \\ G(x)+x \end{cases}$ by

adding another non-parameter shortcut of identity mapping between the matching dimension convolution layer and the extraction feature convolution layer. Because the multilevel residual network is more resourceful in residual learning, the recognition ability of the network improves after the addition of structure.

Finally, we add a Dropout layer after the convolution operation of each extended nonlinear kernel model is completed: $H_1(x) = D[G(x) + x]$; x in $D(x)$ (the setting of x is generally in a proportional form) can be adjusted according to the settings of different parameters. At this point, the weights of some hidden layer nodes in the network are randomly disabled during model training; those disabled nodes can be temporarily considered not to form part of the network structure, but their weights are retained. The issues of network over-fitting and excess time consumption during network parameter training are resolved after the introduction of the Dropout layer into the model.

4 Experimental Results and Analysis

4.1 Experimental data and platform

To verify the effectiveness of the proposed algorithm, the following three experiments were designed, and the experimental results were analysed. The experiments were based on the Ubuntu16.04 system and the equipped with Intel i7-4770 CPU and Nvidia GTX TITAN X GPU. The benchmark recognition datasets MNIST and SVHN were used to verify the experimental results for different network structures with the same number of network layers and using the same data set.

The MNIST dataset consists of 70,000 handwritten digits from 0 to 9, centered on a 28×28 square canvas. Each pixel represents the gray-scale in the range of $[0, 255]$. We split the dataset into 50,000 training samples, 10,000 validation samples and 10,000 test samples, following the standard split.

The Street View House Numbers (SVHN) dataset consists of cropped images representing house numbers captured by Google Street View vehicles as a part of the Google

Maps mapping process. These images consist of digits 0 through 9 with values in the range of [0, 255] in each of 3 red-green-blue color channels. Each image is 32 pixels wide and 32 pixels high giving a sample dimensionality (32, 32, 3). The number of samples we used for training, valid, and test sets is 543,949, 60,439, and 26,032 respectively.

In experiment 1, the handwritten character recognition results obtained using an ordinary convolution network, an ordinary residual network structure, and an extended nonlinear kernel residual network were compared on the MNIST dataset.

In experiment 2, the effectiveness of the extended nonlinear kernel residual network algorithm design was verified on the MNIST dataset.

In experiment 3, the effectiveness of the extended nonlinear kernel residual network algorithm design is verified on SVHN dataset in the character recognition.

4.2 Experimental1: Experimental comparisons

4.2.1 Experimental design and training process

Experiment 1 uses the extended nonlinear kernel residual network-based structure (Fig. 3) as well as the classical convolution neural network and a residual network structure using a residual kernel structure for the classification experiments. To analyse the difference in performance of the three networks, the experiment adopts the same parameters with respect to filtering, convolution step length and down-sampling precision. With respect to the network structure, the fact that the extended nonlinear kernel and the residual kernel are multilayer structures was taken into account; the number of layers cannot be completely consistent, but the difference in the number of training network layers is minimized as much as possible in the experiment. Figs. 6 and Fig. 7 below show the training process of the extended nonlinear residual network structure. The horizontal axes represent the number of training clusters, and the vertical axes represent the numerical values of accuracy and cross entropy in the training process in Figs. 6 and Fig. 7, respectively.



Fig. 6. Dependence of the accuracy curve on the number of training clusters in the training process using the extended nonlinear kernel residual network model

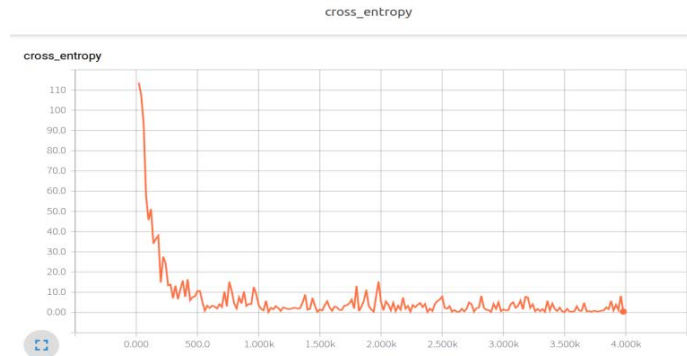


Fig. 7. Dependence of cross entropy on the number of training clusters in the training process using the extended nonlinear kernel residual network model

4.2.2 Analysis and discussion of experimental results

Based on the MNIST handwritten character test data library, the experimental test accuracies of the different models were compared; the results are shown in **Fig. 8**, **Fig. 9**, and **Table 1**. Although the recognition rate of the extremely deep network is high, its parameter training process is complex, which consumes a large amount of training time, whereas the performance result of the deep (not the extremely deep) network with respect to recognition accuracy is still inadequate. This paper proposes a new deep (not extremely deep) network structure that displays better recognition accuracy and higher recognition efficiency for moderate-size character recognition training sets, simulating the actual situation of lacking fully labelled samples and structures with the same numbers of layers. Based on this, this paper selects an MNIST handwritten character library for the experiment. It should be noted that through data set training and under the same conditions, the difference between the test accuracy of the ordinary convolution neural network model and those of the residual network model and the extended nonlinear kernel residual network model is large. Thus, the extended training data set of the convolution neural network in the experiment was 20k, which made the training of the ordinary convolution network model more thorough. Nevertheless, the number of training clusters in the residual network model and the extended nonlinear kernel residual model was still 4k.

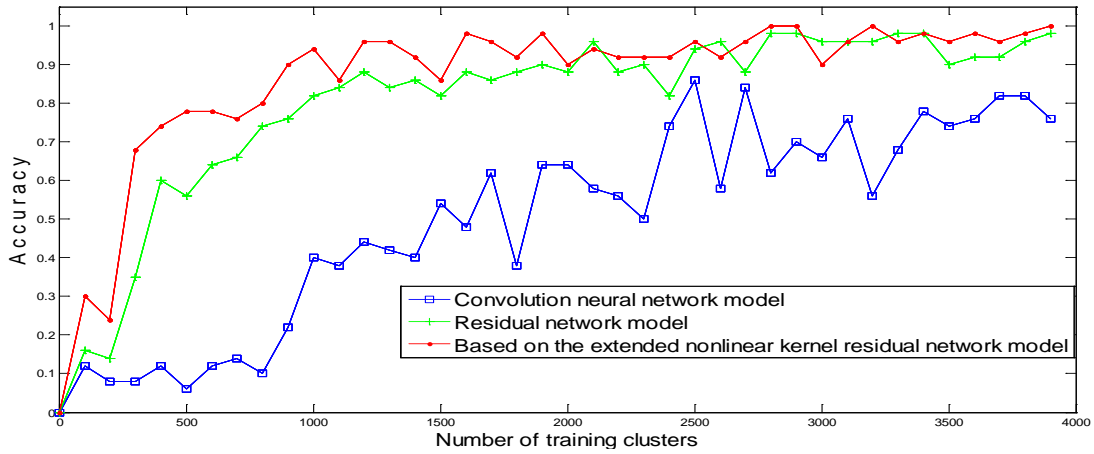


Fig. 8. Experimental training accuracy curves of different models in 4K training clusters

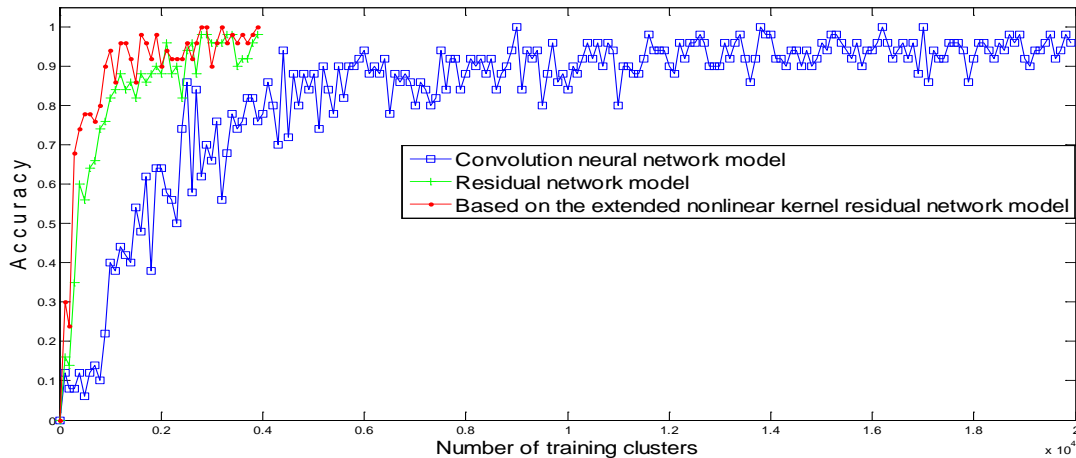


Fig. 9. Experimental training accuracy curves of different models in the training cluster of convolution neural network is increased to 20K

Table 1. Comparison of the experimental results obtained with different models on the MNIST handwritten library experiment

Method	Number of training clusters (k)	Test accuracy (%) (mean \pm standard deviation)
Convolution neural network model	20	94.65
Residual network model	4	98.4375
Extended nonlinear kernel residual network-based model	4	99.6094

It can be seen from **Table 1** that the extended nonlinear kernel residual network-based model yields better recognition results on the MNIST dataset. Based on this experimental result, this paper proposes the extended nonlinear kernel residual network-based handwritten

character recognition algorithm. In the comparative experiment that follows, the effectiveness of the extended nonlinear kernel residual network-based handwritten character recognition design will be specifically verified.

4.3 Experimental 2: Effectiveness of the Extended Nonlinear Kernel Residual Network Design

The effectiveness of the extended nonlinear kernel residual network algorithm was verified in Section 4.2. The recognition result obtained using the extended nonlinear kernel residual network-based handwritten character recognition algorithm model is better than those obtained with the ordinary convolution neural network model and the residual learning network model. In addition to the introduction of an appropriate unsupervised intra-class clustering method, the reasons for the better recognition result also include the following:

- (1) A Dropout layer was added to every kernel unit after parameter optimization; this randomly disables the weighting of some hidden layer nodes (but keeps their weight values) during the model training process, thereby preventing the network from being trapped in local optima and resolving the network over-fitting problem. Moreover, the random elimination of a large amount of data during the Dropout process also improves the issue of excessive time consumption during network parameter training.
- (2) An intermediate convolution computation operation model with a width level of 2 with pre-processing was proposed; this enhances the ability of the non-extremely deep network to express complex functions.
- (3) A composite residual structure with a multi-level quick link was designed. The multi-level structure results in a smaller error when the network calculates complex functions and allows more thorough study of features by the non-extremely deep network.

4.3.1 Verification of the Effectiveness of Handwritten Character Recognition after Adding a Dropout Layer Design

Keeping the rest of the structures in the extended nonlinear kernel model network fixed, we only add a Dropout layer after each kernel learning process and randomly disable the weight of some of the hidden layer nodes in the network; the effectiveness of these changes are tested experimentally.

The new kernel of the extended nonlinear kernel model after the addition of a Dropout layer is shown in [Fig. 10](#); the recognition accuracy rate is 99.2188%.

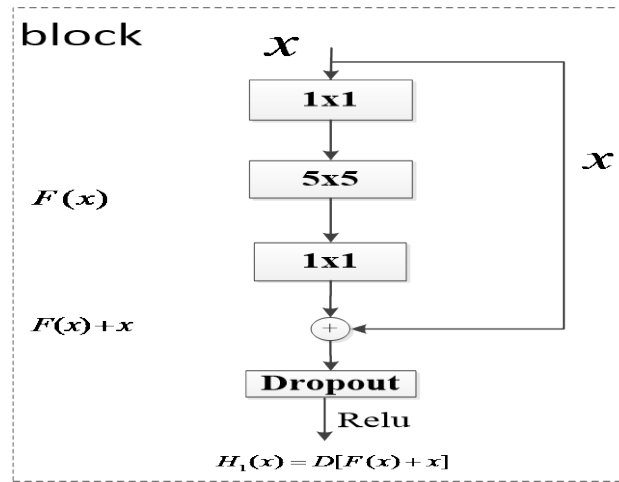


Fig. 10. Design of the extended nonlinear kernel model with a Dropout layer

4.3.2 Design of the Model with a Convolution Layer of Width Level 2 and Verification of Its Effectiveness in Handwritten Character Recognition

Maintaining the rest of the structures in the extended nonlinear kernel model network fixed, we add an intermediate convolution computation model a width level of 2 to each kernel learning process which has been pre-processed; the effectiveness of the experimental results is observed.

A convolution layer with a width level of 2 in the intermediate convolution layer is added to the extended nonlinear kernel model, as shown in **Fig. 11** below; the recognition accuracy rate is 98.8679%.

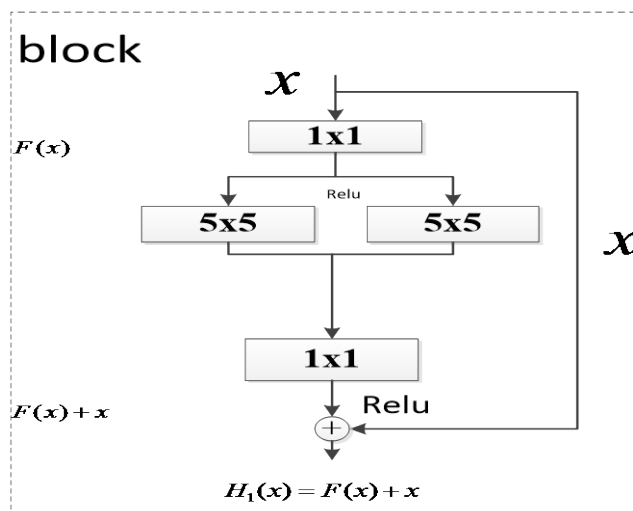


Fig. 11. Design of the extended nonlinear kernel model with a convolution layer of width level of 2

4.3.3 Verification of the Effectiveness of Handwritten Character Recognition after the Introduction of a Composite Residue Structure with a Multi-Level Quick Link

Maintaining the rest of the structures in the extended nonlinear kernel model network fixed, we design a composite residue structure with a multi-level quick link in each kernel learning process. The effectiveness of the experimental results is observed.

The new kernel of the extended nonlinear kernel model after the introduction of a composite residue structure is shown in Fig. 12 below; the recognition accuracy rate is 98.8281%.

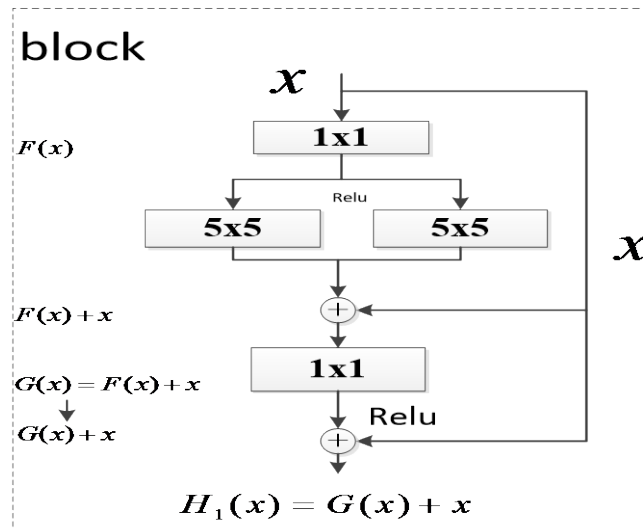


Fig. 12. Design of the extended nonlinear kernel model witha multi-level quick link

4.3.4 Experimental Results

The results of the extended nonlinear kernel model design experiments are shown in Table 2.

Table 2. Results of the extended nonlinear kernel design experiments

Method	Test accuracy (%) (mean \pm standard deviation)
the character classification method that has its origin in a residual network	98.4375
Residue kernel model(adding a priori knowledge) + Dropout layer design	99.2188
Residue kernel model(adding a priori knowledge) + intermediate convolution layer of width level 2 design	98.8679
Residue kernel model(adding a priori knowledge) + composite residue structure design	98.8281

4.3.5 Analysis of Extended Nonlinear Kernel Model Design Experiments

Sections 4.3.1, 4.3.2, and 4.3.3 describe experiments with controlled variables. These experiments proved that the three network structures proposed in this paper ((1) the addition of a Dropout layer after parameter optimization; (2) presentation of an intermediate convolution model with a pre-processed width level of 2; and (3) the design of a composite residual structure with a multi-level quick link) can all provide a certain degree of optimization of the recognition results and the recognition efficiency of the network. The results presented in Section 4.2 show that the effect of the overall algorithm combining the three designs is better than the individual effect of each design and better than the effect of the other two methods alone. Based on the above experimental results, this paper proposes an extended nonlinear kernel residual network-based handwritten character recognition algorithm.

4.4 Experimental 3: Effectiveness of the Extended Nonlinear Kernel Residual Network Algorithm Design is verified on SVHN dataset in the character recognition

Maintaining the structures of three models in the experimental 1, we changed the recognition dataset to SVHN, and also joined the recent popular networks for experimental comparison. The experimental results has also shown that our method has some improvement on the dataset, shown in [Table 3](#).

Table 3. Comparison of the experimental results obtained with different models on the SVHN character library experiment

Method	Test accuracy (%) (mean \pm standard deviation)
HOG	85.0
Stacked Sparse Auto-Encoders	89.7
Convolution neural network model	95.1
Stochastic Pooling	97.1
Network in Network	97.65
Residual network-based model	97.70
Extended nonlinear kernel residual network-based model	97.72

5 Conclusions

This paper proposed an extended nonlinear kernel residual network-based handwritten character recognition algorithm model. The main idea of this work are two designs, which one is the design of an unsupervised intra-class clustering algorithm to perform targeted pre-processing of the input data. The other is the three network structure designs for the

extended nonlinear kernel were proposed in this paper: (1) presentation of an intermediate convolution model with a pre-processed width level of 2; (2) a composite residue structure with the design of a multi-level quick link; and (3) addition of a Dropout layer after parameter optimization in the extended nonlinear kernel model. Compared with those convolution neural network and residual network structures under the same conditions, the proposed model has good recognition results and data training time. In future work, on the one hand, we will try to enhance the robustness of our algorithm; on the other hand, more simple and effective classification methods would be sought to optimize and improve the network.

References

- [1] LeCun Y, Bengio Y, and Hinton G., "Deep learning," *Nature*, 521(7553), 436-444, 2015. [Article \(CrossRef Link\)](#).
- [2] Zhang X Y, Bengio Y, Liu C L., "Online and offline handwritten Chinese character recognition: A comprehensive study and new benchmark," *Pattern Recognition*, 61, 348-360, 2017. [Article \(CrossRef Link\)](#).
- [3] Liu C L, Yin F, Wang D H, et al., "Online and offline handwritten Chinese character recognition: benchmarking on new databases," *Pattern Recognition*, 46(1), 155-162, 2013. [Article \(CrossRef Link\)](#).
- [4] Foggia P, Percannella G, Vento M., "Graph matching and learning in pattern recognition in the last 10 years," *International Journal of Pattern Recognition and Artificial Intelligence*, 28(01), 2014. [Article \(CrossRef Link\)](#).
- [5] Zhang L, Tan J, Han D, et al., "From machine learning to deep learning: progress in machine intelligence for rational drug discovery," *Drug Discovery Today*, 1680-1685, 2017. [Article \(CrossRef Link\)](#).
- [6] Bottou L., "From machine learning to machine reasoning," *Machine learning*, 94(2), 133-149, 2014. [Article \(CrossRef Link\)](#).
- [7] Yang J, Jiao Y, Xiong N, et al., "Fast Face Gender Recognition by Using Local Ternary Pattern and Extreme Learning Machine," *KSII Transactions on Internet & Information Systems*, 7(7), 2013. [Article \(CrossRef Link\)](#).
- [8] Zheng Y, Liu J, Liu H, et al., "Integrated Method for Text Detection in Natural Scene Images," *KSII Transactions on Internet & Information Systems*, 10(11), 2016. [Article \(CrossRef Link\)](#).
- [9] Ouellet S, Michaud F., "Enhanced automated body feature extraction from a 2D image using anthropomorphic measures for silhouette analysis," *Expert Systems with Applications*, 270-276, 2017. [Article \(CrossRef Link\)](#).
- [10] Yi C, Tian Y., "Scene text recognition in mobile applications by character descriptor and structure configuration," *IEEE transactions on image processing*, 23(7), 2972-2982, 2014. [Article \(CrossRef Link\)](#).

- [11] Yin X C, Yin X, Huang K, et al., "Accurate and robust text detection: A step-in for text retrieval in natural scene images," *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval. ACM*, 1091-1092, 2013.
[Article \(CrossRef Link\)](#).
- [12] Luo H, Zhao F, Chen S, et al., "A Tree Regularized Classifier--Exploiting Hierarchical Structure Information in Feature Vector for Human Action Recognition," *KSII Transactions on Internet & Information Systems*, 11(3), 2017.[Article \(CrossRef Link\)](#).
- [13] Delaye A, Liu C L., "Contextual text/non-text stroke classification in online handwritten notes with conditional random fields," *Pattern Recognition*, 47(3), 959-968, 2014.
[Article \(CrossRef Link\)](#).
- [14] Li W, Zhang T, Zhu Z, et al., "Detection of LSB Matching Revisited Using Pixel Difference Feature," *KSII Transactions on Internet & Information Systems*, 7(10), 2013.
[Article \(CrossRef Link\)](#).
- [15] Mussarat Y, Muhammad S, Sajjad M, et al., "Content based image retrieval using combined features of shape, color and relevance feedback," *KSII Transactions on internet and information systems*, 7(12), 3149-3165, 2013. [Article \(CrossRef Link\)](#).
- [16] Shrivastava V K, Londhe N D, Sonawane R S, et al., "Exploring the color feature power for psoriasis risk stratification and classification: A data mining paradigm," *Computers in biology and medicine*, 65, 54-68, 2015. [Article \(CrossRef Link\)](#).
- [17] Kim S, Yu Z, Kil R M, et al., "Deep learning of support vector machines with class probability output networks," *Neural Networks*, 64, 19-28, 2015. [Article \(CrossRef Link\)](#).
- [18] Wong W K, Cheung D W, Kao B, et al., "Secure kNN computation on encrypted databases," *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data. ACM*, 139-152, 2009.[Article \(CrossRef Link\)](#).
- [19] Aisha A, Muhammad S, Hussain S J, et al., "Face recognition invariant to partial occlusions," *KSII Transactions on Internet and Information Systems*, 8(7), 2496-2511, 2014.
[Article \(CrossRef Link\)](#).
- [20] Chu J, Liang H, Tong Z, et al., "Slow Feature Analysis for Mitotic Event Recognition," *KSII Transactions on Internet & Information Systems*, 11(3), 1670-1683, 2017.
[Article \(CrossRef Link\)](#).
- [21] He K, Zhang X, Ren S, et al., "Deep residual learning for image recognition," in *Proc. of Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on*, 27-30 June 2016.
[Article \(CrossRef Link\)](#).
- [22] Huang G, Sun Y, Liu Z, et al., "Deep networks with stochastic depth," *European Conference on Computer Vision. Springer International Publishing*, 646-661, 2016.[Article \(CrossRef Link\)](#).
- [23] Luo X, Shen R, Hu J, et al., "A Deep Convolution Neural Network Model for Vehicle Recognition and Face Recognition," *Procedia Computer Science*, 107, 715-720, 2017.
[Article \(CrossRef Link\)](#).

- [24] Lee S J, Yun J P, Koo G, et al., "End-to-end recognition of slab identification numbers using a deep convolutional neural network," *Knowledge-Based Systems*, 132, 1-10, 2017. [Article \(CrossRef Link\)](#).
- [25] Kawano Y, Yanai K., "Food image recognition with deep convolutional features," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM, 589-593, 2014. [Article \(CrossRef Link\)](#).
- [26] Girshick R., "Fast r-cnn," in *Proc. of the IEEE International Conference on Computer Vision*, 1440-1448, 2015. [Article \(CrossRef Link\)](#).
- [27] Kumar M, Jindal M K, Sharma R K., "Segmentation of isolated and touching characters in offline handwritten Gurmukhi script recognition," *International Journal of Information Technology and Computer Science (IJITCS)*, 6(2), 58-63, 2014. [Article \(CrossRef Link\)](#).
- [28] Kamble P M, Hegadi R S., "Handwritten Marathi character recognition using R-HOG Feature," *Procedia Computer Science*, 45, 266-274, 2015. [Article \(CrossRef Link\)](#).
- [29] Liang M, Hu X., "Recurrent convolutional neural network for object recognition," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, 3367-3375, 2015. [Article \(CrossRef Link\)](#).
- [30] Schmidhuber J., "Deep learning in neural networks: An overview," *Neural Networks*, 61, 85-117, 2015. [Article \(CrossRef Link\)](#).
- [31] Zhang Y, Shi B M., "Improving pooling method for regularization of convolutional networks based on the failure probability density," *Optik-International Journal for Light and Electron Optics*, 145, 258-265, 2017. [Article \(CrossRef Link\)](#)
- [32] Affonso C, Rossi A L D, Vieira F H A, et al., "Deep learning for biological image classification," *Expert Systems with Applications*, 85, 114-122, 2017. [Article \(CrossRef Link\)](#)
- [33] LeCun Y, Bottou L, Bengio Y, et al., "Gradient-based learning applied to document recognition," in *Proc. of the IEEE*, 86(11), 2278-2324, 1998. [Article \(CrossRef Link\)](#)



Zheheng Rao: He is currently pursuing a master degree at the School of Electrical and Electronic Engineering in Hubei University of Technology, Wuhan, China. His current research interests include the deep learning and image processing.



Chunyan Zeng: received her M.S. degree and PhD degree in Electronic Engineering from South China University of Technology, Guangzhou, P.R. China, in 2013. She is now an lecturer in the School of Hubei University of Technology. Her research interests include compressed sensing and deep learning.



Minghu Wu: He received the B.S. degree from Communication University of China, Beijing, China and the M.S. degree from Huazhong University of Science and Technology in 1998 and 2002, respectively. He received PhD degree from Nanjing University of Post and Telecommunications in 2014. His major research interests include signal processing, video coding and compressive sensing.



Zhifeng Wang: He received the BEng degree in Electronic Engineering from China University of Geosciences in 2008, and the PhD degree in Electronic Engineering from South China University of Technology, China, in 2013. He was a joint training PhD student in the computer science department of Carnegie Mellon University, Pittsburg, PA, during 2010 to 2011. He is now an associate professor in the School of Educational Information Technology of Central China Normal University. His research interests include deep learning, data mining, machine learning, signal processing, and digital forensics.



Nan Zhao: She received the PhD. degrees from the Wuhan University, Wuhan, China in 2013. She is currently an associate professor in the Hubei University of Technology, Wuhan, China. Her research interests include wireless network and cognitive radio.



Min Liu: She received the B.S. degree from Zhengzhou University, Zhengzhou, China in 2004. She received the M.S. degree and PhD degree from Huazhong University of Science and Technology in 2007 and 2013, respectively. Her research interests include the video processing and computer vision.



Xiangkui Wan: He received the M.S. and Ph.D. degree in mechatronic engineering from Chongqing University in 2002 and 2005, respectively. He is currently a professor in the School of Electrical and Electronic Engineering, Hubei University of Technology, P. R. China. His research interests include digital signal processing, biomedical signal processing, with applications to electrocardiography signal processing.