

NEURAL NETWORK & GENETIC ALGORITHM IN APPLICATION TO HANDWRITTEN CHARACTER RECOGNITION

LE HOAI BAC, LE HOANG THAI

Abstract. In recent years, many soft computing technologies have been exploited and become promising tools in solving many problems of pattern recognition. In this paper, we represent a novel hybrid technique for associating individual neural networks by using genetic algorithms. This technique is used to solve the problem of handwritten character recognition. Our experimental system shows that it gives better results (approximately 98.53% characters were recognized correctly) in comparison to other traditional techniques.

Tóm tắt. Những năm gần đây, nhiều kỹ thuật phần mềm máy tính từng được khai thác và trở thành các công cụ hứa hẹn ứng dụng cho bài toán nhận dạng mẫu.

Trong phạm vi bài báo, chúng tôi trình bày một kỹ thuật lai: sử dụng thuật giải di truyền để liên kết các mạng nơ ron cá thể. Hệ thống xây dựng từ kỹ thuật này được ứng dụng để giải quyết bài toán nhận dạng kí tự viết tay.

Các kết quả thử nghiệm cho thấy hệ thống thu được những kết quả với tỉ lệ nhận dạng cao (98,53%) tốt hơn so với các phương pháp cổ truyền khác.

1. INTRODUCTION

With the explosion of information, there are more and more requirements of the automation for text translating and recognizing process. Because of this, today, the field of pattern recognition and particularly handwritten recognition has received a lot of attention from research community in artificial intelligence. The process of text recognition includes the following primary stages:

- Digitizing text by scanner and then storing results in files.
- Raw processing such as enhancing images, extracting character regions from text, then extracting each character in the form of pixel matrix.
- Recognizing each character.
- Restoring the content of text.

Here we only discuss about the stage of character recognition, which decides the speed and exactness of a text processing and recognition system.

Until now, we have many traditional character recognition techniques such as mask matching, contrasting of weighted points through a central point, contrasting by using splitting and dividing point, exploiting the statistics of intersected points and method of structural recognition, etc...

Generally, the above traditional techniques couldn't work well in the case of the quality of image digitizing process is not so high.

In order to overcome this shortcoming, *Artificial Neural Network* has been used in pattern recognition problems. A neural network is a biological simulation of human brain by computer. It is a parallel structure constructed from many elements (neurons) associating together through weighted numbers. Each neuron is also a non-linear dynamic system which has the ability of self-learning. Thus, a neural network can learn from experiences or from a pattern set [3].

The use of neural networks in pattern recognition problems does not require pre-processing stages as thinning, contour smoothing or noise filtering.

Recently, the combination of multiple neural networks through flexible computing tools has been referred as a new way to construct pattern recognition problem-solving systems with high efficiency [4, 7]. While normal techniques select the best network from the attendant ones, this associating technique will keep all individual networks and apply an appropriate common set decision strategy.

In this paper, we propose a novel method using *Genetic Algorithm* in *Associating Individual Neural Networks*. They not only consider the differences in performance of each network during association, but also use weighted numbers (evaluating the reliability of each individual network). The genetic algorithm will define the values of these weighted numbers, and associate individual networks to obtain a suitable output result.

By comparing with traditional techniques through experiments in handwritten recognition problem, our method proves its pre-eminent quality.

The following section will discuss in detail about the problem and possible methods to associate individual neural networks.

Section 3 represents our method of *Associating Individual Neural Networks* based on *Genetic Algorithm*.

Section 4 shows experimental results in applying the proposed method to recognize handwritten characters (for vowels a, e, i, o, u).

Section 5 concludes this paper.

2. THE METHODS FOR ASSOCIATING MULTIPLE NEURAL NETWORKS

A feedforward neural network is considered as a mapping means between the sets of input and output values. It plays a role of a function f that maps the input set I into the output set O , i.e. $f : I \rightarrow O$ or $y = f(x)$, where $y \in O$ and $x \in I$. Since the problem of classification is actually a map from the space of characteristic vectors into the set of output classes, we consider neural networks as a classifier (in particular the two-layer feedforward neural network is trained by the general delta learning rule).

Given a classifier - a two layer neural network with T neurons in input layer, H neurons in hidden layer, and c neurons in output layer. Here T is the dimension of characteristic vectors, c is the number of pattern classes, and H is a properly selected number. The network will associate completely adjacent layers and its activities can be understood as a nonlinear process: input a pattern $X = (x_1, x_2, \dots, x_T)$ (its class is still unknown) and the set of classes $\Omega = \{\omega_1, \omega_2, \dots, \omega_c\}$; then each output neuron will produce y belonging to one of these classes, which is defined by

$$P(\bar{\omega}_i | X) \approx y_i = f \left\{ \sum_{k=1}^H w_{ik}^{om} f \left(\sum_{j=1}^T w_{kj}^{mi} x_j \right) \right\}. \quad (1)$$

In the expression (1), w_{kj}^{mi} is the weighted number between the j^{th} input neuron and the k^{th} hidden neuron, w_{ik}^{om} is a weighted number between the k^{th} hidden neuron and i^{th} output class, and f is sigmoid function which is defined by $f(x) = \frac{1}{1 + e^x}$. The indexes i, m, o in weighted numbers are used to show that those numbers belong to input class, hidden class, or output class of the neural network. The decision object will belong to the class which has maximum neuron.

The above-mentioned network is trained on the set of experimental patterns and discovers relationships that help to distinguish the patterns. However, a limited size network will not give high efficiency in mapping process. The increase of size and number of hidden classes doesn't provide considerable improvements. Furthermore, in complex problems, for instance, the problem of handwritten recognition, the number of both characteristics and classes is very large.

The main idea in the strategy of using simultaneously multiple networks is constructing n independent networks trained with correspondent characteristics, and applying the method of associating

networks to give decisions in general classification. Figure 1 shows the schema of associating multiple networks. The associator will combine results from trained individual networks with different characteristics. Therefore, the problem is how to synthesize results from each individual network (or could be called expert)?

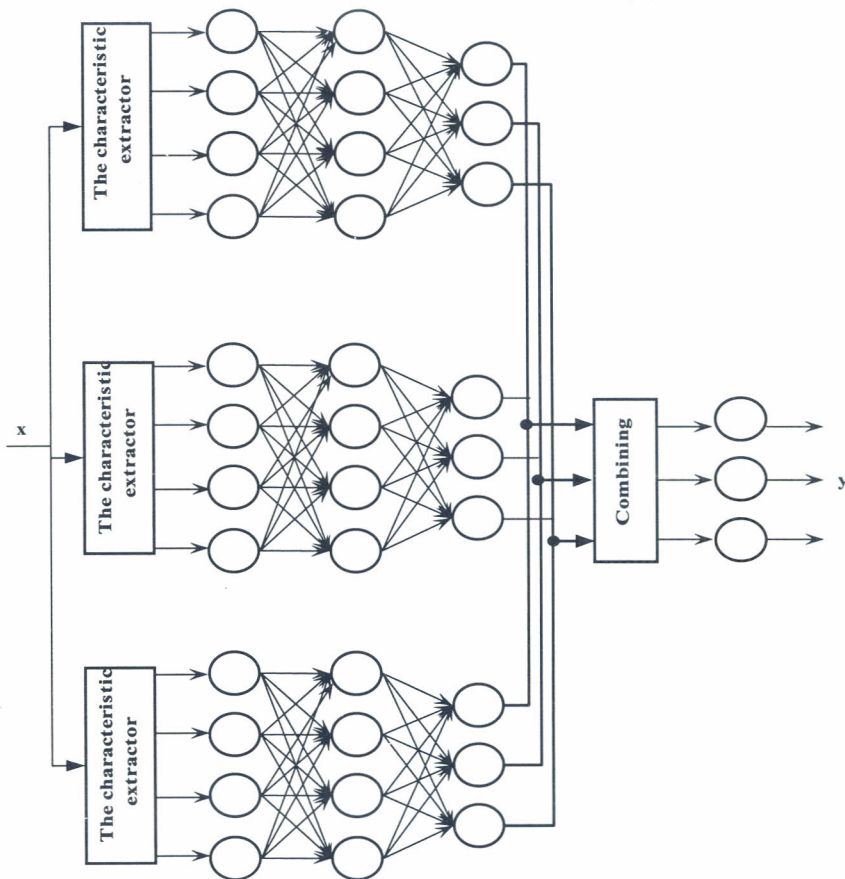


Fig. 1. The method of associating individual neural networks

There are two general methods for associating neural networks: the first one based on unifying technique, and the second one based on selection technique.

With the unifying technique, the classification of an input pattern X is based on the set of real values measured by:

$$P(\omega_i | X), \quad 1 \leq i \leq c. \tag{2}$$

They will compute the probability of the event “ X belongs to one of c original classes”. In associating scheme including n networks, each k^{th} network will be evaluated by the set of probability values:

$$P_k(\omega_i | X), \quad 1 \leq i \leq c, \quad 1 \leq k \leq n. \tag{3}$$

A simple method to associate results on the same pattern X from all n individual networks is using the following average value as a new way to evaluate associated networks:

$$P(\omega_i | X) = \frac{1}{n} \sum_{k=1}^n P_k(\omega_i | X), \quad 1 \leq i \leq c. \tag{4}$$

Thus, the above associating value can be considered as an average classification of Bayes method [7]. This evaluation will be improved if we add the ability of direction for outputs based on knowledge

about the reliability of each network:

$$P(\omega_i | X) = \sum_{k=1}^n r'_k P_k(\omega_i | X), \quad 1 \leq i \leq c, \quad (5)$$

where

$$\sum_{k=1}^n r'_k = 1. \quad (6)$$

From another point of view, the *selection technique* considers results of each network as the decision of an expert. Many selection techniques have been created and based on the theory of group decision forming, such as unanimity, Borda count etc ... (see details in [8]).

3. THE GENETIC ALGORITHM FOR COMBINING INDIVIDUAL NEURAL NETWORKS

This section will introduce generally about *Genetic Algorithm* (GA) and the application of GA to associate individual neural networks.

3.1. Fundamentals about genetic algorithm (GA) [6]

Until now, in the research and application of informatics, there have been many interesting problems without fast and appropriate solving methods appeared. The most part in those problems are optimal ones existed in applications. In order to solve this kind of problems, it is often necessary to find an efficient algorithm. However the results are just nearly optimal. A genetic algorithm (GA) is one of those tools.

The following scheme shows general structure of a genetic algorithm:

```

Procedure GA;
  Begin
    t := 0;
    Initializing class P(t);
    Evaluating class P(t);
    While not End_Condition do
      Begin
        t:=t+1;
        Select P(t) from P(t-1);
        Reconstructing P(t);
        Evaluating P(t);
      End
    End.
  
```

The basic operations of genetic algorithm for a particular problem include:

- * Representing genetic with potential solutions of the problem;
- * Creating the population of original solution;
- * Constructing evaluating function;
- * Defining genetic operators;
- * Defining parameters used by a genetic algorithm.

3.2. Applying genetic algorithm (GA) in associating neural networks

In computing process, the genetic algorithm maps the solution space of the problem into the set of strings, each string represents a potential solution.

In our problem, each string has to encode $n \times c$ real value parameters (r_i^k) in the expression (5), by this we can obtain optimal associating factors for combining individual neural networks. Each factor will be encoded into 8 bit and modified in the interval [0, 1]. After that, GA operates over encoded strings to find better solutions corresponding to each generation.

The operation cycle of GA includes:

1. Creating the population of real value encoded strings.
2. Evaluating each string with fitness function defined as recognition ratio correspond to training data set.
3. Selecting good strings (has maximum fitness value).
4. Using genetic crossovers and mutation operators to create new populations of strings.

This cycle stops when recognition ratio gains best possible level.

Note: All individuals belonging to the old population will be replaced by the new ones. The storage of the best potential solutions will help obtain the advancement after each generation [2].

Figure 2(a) depicts four stages in applying biological genetic technique. GA method proposes to take the set of factors (r_i^k) of individual neural networks to form correspondent encoded strings, as shown in Figure 2(b).

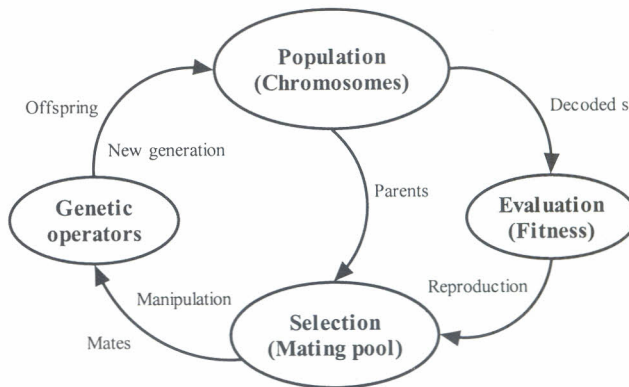


Fig. 2(a). The basic steps of a genetic algorithm

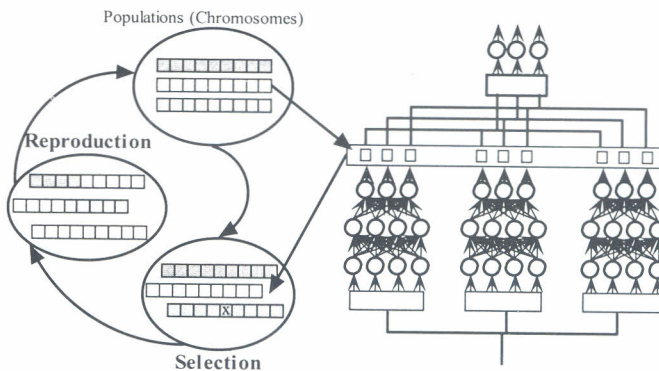


Fig. 2(b). The algorithmic schema of the proposed method

4. THE APPLICATION IN RECOGNIZING HANDWRITTEN VOWEL CHARACTERS

The above system was implemented by using Borland C language in a Pentium III PC 766 Mhz. This experiment was carried with 13 people. Each one wrote 50 times, each time they wrote 5 vowel characters (a, e, i, o, u). Thus the number of experimental characters are: $13 \times 50 \times 5 = 3250$. In those characters, 2500 characters were used for training and 750 characters were tested. Figure 3(a) and Figure 3(b) depict the sets of training patterns and testing patterns used in this experiment, respectively.

a e i o u

a e i o u

a e i o u

Fig. 3(a). Vowel patterns used for training

a e i o u

a e i o u

Fig. 3(b). Vowel patterns used for testing

4.1. Extracting characteristics

Each character, in handwritten or typewriting format, has to draw fundamental lines. Therefore the seeking of local strokes becomes a suitable method to extract characteristics. With each position in image, the information represents a stroke in a given direction will be saved in a characteristic vector [5]. Especially in this paper, Kirsch masks [1] have been used for extracting direction characteristics.

Kirsch defined non-linear algorithm to detect boundary as follows:

$$G(i, j) = \max \left\{ 1, \max_{k=0}^7 [5S_k - 3T_k] \right\}, \quad (7)$$

where

$$S_k = A_k + A_{k+1} + A_{k+2}, \quad (8)$$

$$T_k = A_{k+3} + A_{k+4} + A_{k+5} + A_{k+6} + A_{k+7}. \quad (9)$$

Here $G(i, j)$ is *gradient* of pixel (i, j) , a boundary A has the length of 8, and A_k ($k = 0, 1, 2, \dots, 7$) are 8 neighbors of pixel (i, j) as showed in Figure 4.

A_0	A_1	A_2
A_7	(i, j)	A_3
A_6	A_5	A_4

Fig. 4. 8 neighbors A_k ($k = 0, 2, \dots, 7$) of pixel (i, j)

Firstly, the data set will be pre-processed. After that each vowel will be divided into sub-regions with sizes 16×16 to preserve directions of the image. Then characteristic vectors for horizontal direction, vertical direction, right-diagonal direction, and left-diagonal direction will be formed from the divided image [1].

$$G_V(i, j) = \max (|5S_2 - 3T_2|, |5S_6 - 3T_6|),$$

$$G_H(i, j) = \max (|5S_0 - 3T_0|, |5S_4 - 3T_4|),$$

$$G_L(i, j) = \max (|5S_3 - 3T_3|, |5S_7 - 3T_7|),$$

$$G_R(i, j) = \max(|5S_1 - 3T_1|, |5S_5 - 3T_5|).$$

The final step in extracting characteristics is using median operators to compress directed vectors with size of 16×16 to vectors with size of 4×4 , these operators produce a value for 2×2 pixel by summing 4 values on $2 \times 2 = 4$ pixels divided by 4.

Besides that, the 8×8 reduced *normalized image* plays a role of a global characteristic.

At this point we also exploit a boundary characteristic. After extracting boundary from a normalized size image (8×8 image), we obtain boundary characteristics representing directions for a vowel pattern.

Thus in the final result, the characteristics used include: four 4×4 local characteristics, one global characteristic: 8×8 *normalized image*, and characteristics have structures which were extracted from boundary of vowel characters.

4.2. The application of associating method

In order to evaluate the ability of associating individual networks, we conducted our experiments with three two-layer networks. Each network uses different characteristic tuples. NN_1 is used with Kirsch characteristics, NN_2 with *normalized image*, and NN_3 with the series of boundary characteristics.

By exploiting this method, each network will form a decision through its own standard. That means, after being trained with each characteristic tuple, each NN is possible to give out its own class division (Each NN will appreciate the belonging of the received pattern in comparison to vowels a, e, i, o, u).

Each network will be trained with 2500 vowel patterns and tested over 750 vowel patterns.

We use back-propagation algorithm [2] for the training with the iteration of computing process continuing to the point where squared mean error less than 0.9 (corresponding to the set of training pattern), or when the number of iterative times equals to 1000 (the limit point of networks during training). The parameters used for training are learning factors (equals to 0.1) and an input vector which is decided to belong to a class based on maximum output value, respectively.

After training three networks with separate characteristics, we use GA to find optimal parameters for associating networks. Our original population has 100 individuals, and each contains 120 bit ($3 \times 5 \times 8$). The evolutionary parameters are used for this experiment in the following way: crossover probability at one point is 0.6, and mutation probability equals to 0.01 (1%).

The fitness value $\text{Res}[i]$ ($i = 1, \dots, \text{numclass}$, numclass is the number of class, in this case, it equals to the quantity of vowel patterns, numclass = 5) is assigned to a string by testing the recognition ratio with trained vowel patterns.

In details, assumed that the output of each network is $\text{output}[k][i]$ ($k = 1, \dots, n$, n equals to the quantity of NN . In our problem, take, for example, $n = 3$, $i = 1, \dots, \text{numclass}$). Applying genetic algorithm over encoded string of Combiner to obtain the best recognition ratio on training pattern set with the following constraint:

$$\text{Res}[i] = \sum_{k=1}^n (\text{Combiner}[k][i] * \text{Output}[k][i]), \quad i = 1, \dots, \text{NumClass}.$$

Here, combiner is the set of neural network associating factors, where $\text{Combiner}[k][i]$ is the factor (weighted number) of i th output of k^{th} network.

If $\text{Res}[i]$ is maximum, the object is decided to belong to class i .

(This applying expression is equivalent to the expression (5) in the second section.)

The cycle of evolutionary processing stops when the best fitness value of the population couldn't be improved any more in the next loops.

Through experiments, we observe that the efficiency increases gradually over generations and after original basic improvements the global fitness value will rapidly gain the convergent value (around 150 loops).

4.3. The experimental results

With the proposed method, we obtain the following result:

The number of tested vowel patterns	The number of error vowels	The recognition rate
750	11	98.53%

By applying the method of computing averages to associate three networks, we have:

The number of tested vowel patterns	The number of error vowels	The recognition rate
750	31	95.87%

By applying a neural network for recognizing the above vowel set, we obtain the following result:

The number of tested vowel patterns	The number of error vowels	The recognition rate
750	59	92.13%

From the above results, we can observe that our proposed method provides better results than applying the method of computing average or the method of using single neural network. In practice, the proposed method has exploited the advantages of recognition ratio obtained from normal methods (single neural network). We can conclude that the method of associating individual neural networks in the problem of handwritten vowel recognition is executable and ensure to obtain high recognition ratios.

5. CONCLUSION

This paper represented the method of applying a genetic algorithm in associating individual neural networks to construct a performance-improved system in the problem of pattern recognition and particularly in handwritten vowel recognition. The experimental results show that this method obtains considerable improvements. This easy-to-understand and easy-to-conduct computing method will power dramatically the field of pattern recognition. The main contribution of this paper is pointing out the potential of hybrid computing technology in applying to the problem of pattern recognition. Besides that, the proposed method also provides important improvements seeking solutions for a variety of problems.

However, all the above results are only the experimental ones at the beginning stage.

We hope that in the near future, we can declare the better results to justify clearly the reliability of the method represented in the paper.

REFERENCES

- [1] Anil K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, 1986.
- [2] Chin - TengLin and C.S. George Lee, *Neural - Fuzzy Systems: A Neuro - Fuzzy synergism to Intelligent Systems*, Prentice - Hall International, Inc, 1996.
- [3] Hoang Kiem, Le Hoai Bac, and Le Hoang Thai, A fuzzy neural network for Vietnamese character recognition, *Proceeding of ICIP'99, International Conference on Image Processing*, IEEE Signal Processing Society, JAPAN, 1999.
- [4] J. A. Benediktsson and P.H. Swain, Consensus theoretic classification methods, *IEEE Trans. On System and Man Cybernet.* **22** (1992) 418-435.

- [5] Knerr L. Personnaz, and G. Drayfus, Handwritten digit recognition by neural networks with single-layer training, *IEEE Trans. On Neural Networks* **3** (6) (1992) 962–968.
- [6] Nguyen Thanh Thuy and Le Hoang Thai, Genetic algorithm: technique and its application in data analysis, *Journal of Computer Science and Cybernetics* **15** (1) 1999.
- [7] S.-B. Cho and J.H. Kim, Combining multiple neural networks by fuzzy integral for robust classification, *IEEE Trans. On System and Man Cybernet.* **25** (1995) 380–384.
- [8] T. K. Ho, “A theory of multiple classifier systems and its application to visual word recognition”, Ph.D Dissertation, University of Buffalo, 1992.

Received June 12, 2001
Revised August 20, 2001

*Department of Information Technology,
University of Natural Sciences, Ho Chi Minh City.*