

Handwriting Digits Recognition Using New Neuro-Markovian Approach

Boukrouh Mehdi^{1*}, Mohamed Lashab² Ahcene Goutas¹ And Salim Ouchtati¹

¹ Université 20 Aout 55, Skikda, Algérie.

² Larbi Ben M'hidi University, Oum El Bouaghi, Algeria

(Received 09 February 2019 - Accepted 21 July 2019 – Published 30 July 2019)

Abstract. This paper presents a hybrid approach based on the Multi Layer Perceptron (MLP) neural network and the Hidden Markov Model (HMM), in addition of some feature extraction methods for the recognition of handwriting digits. The features extraction methods are composed of the entropy method, the object follow-up method, the dropping method, the average line, the column method and finally the descriptors method of the normal Gaussian law. Initially, the learning for each recognition system is used by a 2000 images database where the MLP neural network is based on the gradient backpropagation learning algorithm, and the HMM system is based on the Baum-Welch backpropagation learning algorithm. Furthermore, for each primitive vector, the number of the recognized digits is calculated by the MLP system in which the non-recognized digits are tested by the HMM system. Finally, the recognized digits obtained by both systems are summed for the calculation of the recognition rate. The application of the proposed approach using Matlab gave an impressive recognition rate of 98.8%

Keywords: : Multi Layer Perceptron (MLP), Hidden Markov Model (HMM), Baum-Welch learning algorithm, Gradient learning algorithm

1. Introduction

Due to the importance of writing over the past centuries in collection, storage, and transmission of information, handwriting is still an essential way of communication. Today it is not only used for communication between humans, but also for machine learning [1] and human signature recognition. More recently, the recognition of handwriting is a field that took several interests for researchers. The objective is to develop algorithms with computers softwares that can solve forms recognition problems [2]. The major issues of numeric handwriting recognition is the dimensions and the quality of writing of digits which may change in different variation according to the person, this variation creates problems on the discrimination of digits for the machine, There are two types of recognition that can be distinguished, on-line recognition and off-line recognition. First of all, the data are entered during the writing process by a special electronic surface pen. In the second, the data is scanned after the writing process. Off-line and on-line recognition systems are also discriminated by the applications for which they are intended. The off-line recognition is much more used for bank check processing, mail sorting, and reading business forms, etc., while the on-line recognition is mainly used in computer industries and fields such as signatures verification and authors authentication [1]. Optical characters recognition (OCR) is one of the effective utilizations of handwriting recognition, this field has been a theme of many researches for a long time. At the beginning, the recognition of isolated handwritten characters was explored [3] [4]. While the entire words recognition was addressed later [5]. The greatest part of the writing frameworks that have been announced till today consider compelled recognition issues in view of vocabularies from particular spaces, e.g. the recognition of handwritten check amounts [6], or postal addresses [7] [8].

* Corresponding author.

E-mail: lashabmoh@yahoo.fr (Boukrouh M.).

Address: Université 20 Aout 55, Skikda, Algérie

The recognition rate of such systems is still low, and there is a need to enhance it. Character and handwriting recognition has an extraordinary potential in information and word handling, for example, automated postal address and ZIP code perusing, bank information security, content voice changes, and so on. Therefore, escalated innovative work endeavors and frameworks are accessible for English language [9] [10], Chinese language [11], Arabic language [12], and handwritten numerals, etc [11]. This paper presents a novel recognition system using a hybrid neural network/hidden Markov model for the recognition of the handwritten digits, this system is composed of the three following steps: The first step is the preprocessing, where we have applied the methods of binarization, the object framing method, the pixel inversion method, and the grayscale image normalization. The second step deals with the features extraction, in which five methods of primitives extraction on the output image have been applied. These methods are: the entropy method, the object follow-up method, the dropping method, the average line and column method, and the descriptors method of the normal Gaussian law. Finally is the recognition step, which is divided into two parts; the first is for the system recognition learning and the second is for system testing.

2. Related work

The field of writing recognition had several interests for researchers in recent years. The choice of primitive extraction methods and the choice of classifiers are very important to build a system that recognizes the writing, as they are considered as the articulation of the recognition system. The extraction of the parameters makes it possible to define the form with their primitives' vector, after the classifier uses these parameters to classify the form among several ones in the database to recognize it. Some research works have focused on using a single classifier for recognition, while others have used hybrid systems to achieve recognition of writing; these systems use two or more classifiers for recognition at the same time. [1] tried to use the MLP with two primitive extraction methods called bar featur and the distribution sequence for the recognition of numeric digit strings. [2] in their Arabic writing recognition work combined the MLP and the HMM to achieve recognition. [5] applied a HMM for Recognition of Large Numbers of Writing Vocabulary based on the SDLBA method. [7] combined two classifiers for the recognition of handwritten envelopes using the HWR classifier (Holistic Word Recognizer) and the AWR classifier (Analytical Word Recognizer). [9] used HMMs to build a recognition system. On other hands, our system combines the MLP and the HMM for the recognition of handwritten digits with the most recent extraction methods.

3. The Recognition System

In our work, we have used: The MLP system in neural network, Based on the gradient backpropagation algorithm, and the HMM system based on the Baum- Welch backpropagation algorithm.

3.1. The Neural Network

An artificial neural network is a group of perceptrons connected to each other by an elementary processor operating in parallel, each perceptron calculates the output signal based on the received information from its input [13]

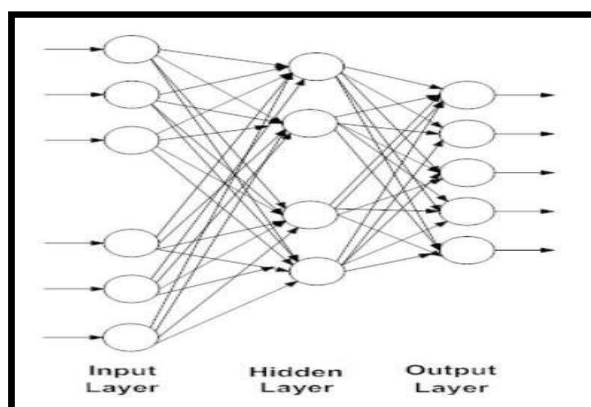


Fig.1. The MLP neural network architecture [14]

The MLP is one of the neural network types, this network is composed of three layers, The first one is the input layer (the system input), the second is the hidden layer (one or more), the third is the output layer (the system output). Figure 1 [1] represents the MLP architecture. The MLP uses the gradient algorithm of Back propagation which modifies the weights of the system to minimize the squared error. This modification of the weights propagates from the output layer to the input layer.

3.2. The Hidden Markov Model (HMM)

HMM (Hidden Markov Model) is a set of states and transitions between them, The HMM is a random process displaced from a state to another by a transition probability and generates an observation.

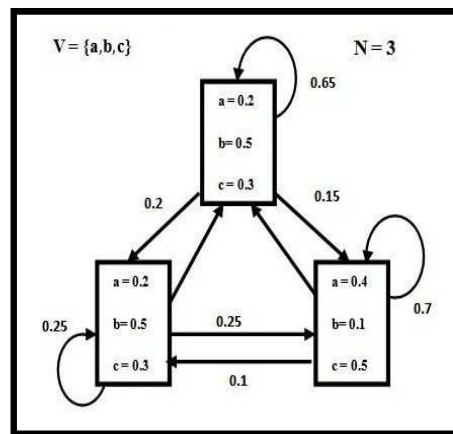


Fig.2. An HMM Representation

In other words, the states are invisible but the observations generated by the states are observed, i.e. at the instant t the system is in the state q_t (invisible), and generates the observation O_t (visible). All of these observations, called the observation sequence can be a finite set of discrete values or symbols (for example the finite alphabet characters), or in a continuous and infinite set (signal frequency, temperature).

- S : A set of N states.
- V : Alphabet of M symbols.
- $T = \{a_{ij}\}$ The matrix of probabilities of transitions between states. a_{ij} Is the probability of transition from state i to j .
- G : The matrix of probabilities of generation of symbols, $P(y_i|q_i)$ is The probability of generating the symbol y observed at time t from the state
- q . [15]. Figure 2 shows the HMM representation.

The learning of Markov models consists to search the matrix parameters T and G of the HMM which allow to maximize for each class of numbers the generated probability sequence, this training has been carried out by Russel and Levinson [16] [17] The idea is to Re-estimate the parameters T and G matrixes to maximize the generated probability sequence. up to an end criterion according to the following steps.

- Choose the initial set of HMM A_0 parameters.
- Calculate A_1 from A_0 , then A_2 from A_1 , etc.
- Repeat this process until an end criterion
- For each learning step p , we dispose A_p and calculate A_{p+1} . Verifying $P(O|A_{p+1}) \geq P(O|A_p)$ [15].

In our work, ten classes of manuscripts digits are used (0,1,2,3,4,5,6,7,8,9), therefore, ten HMMs are constructed using the Buam-Welch algorithm. The Baum-Welch algorithm re-estimates the probability of T and G matrix for each HMM corresponding to their classes till the convergence. When it's finished for all HMMs that are assumed to be equiprobable, the recognition is carried by searching the most discriminated model, i.e. we calculated the probability of emission of the form for all the HMM, and the form to be recognized is carried out in the class of the model which gives the highest probability compared to the other models by: [18]

$$\lambda^* = \operatorname{argmax}_{\lambda \in A} P(o/\lambda)$$

Where A denotes the set of Markov models and λ denotes the discriminated Model

4. Database

Our database content is composed of 3000 digits, which was made at the Algerian University of Skikda with the collaboration of several students, and after writing on a paper we scanned the contents of the paper to convert it into jpg images of 60×60 pixels for use in the learning and testing. Figure 3 presents some examples of different types of digits handwriting

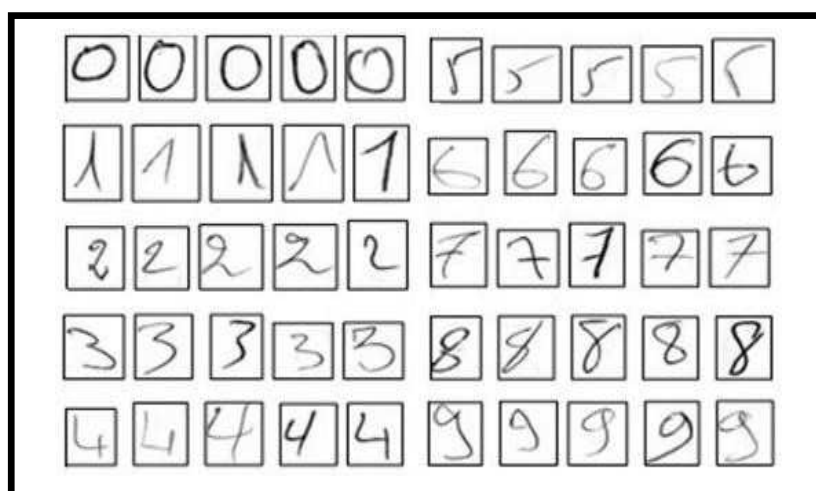


Fig. 3. Some samples of the used database

5. The handwriting recognition system

Our system of work is represented in figure 4 this system is divided into three main parts.

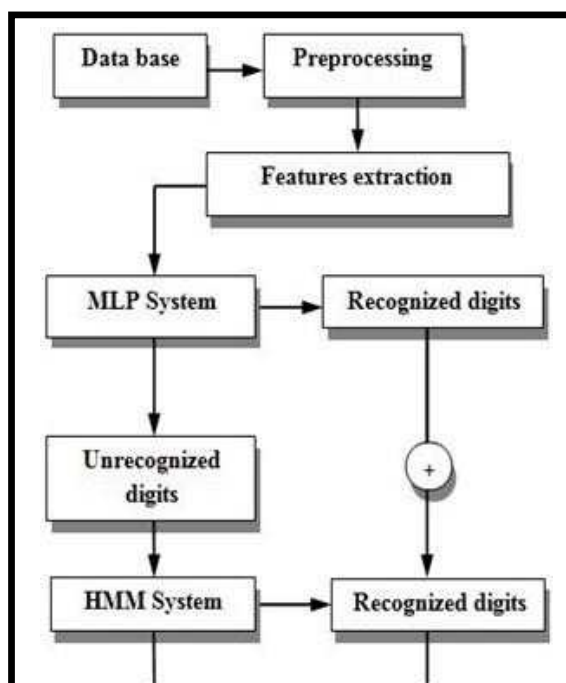


Fig. 4. General Schema of our digits handwritten recognition system

The first part is the Preprocessing, the second is the Features extraction, is the last part the recognition.

5.1. Preprocessing

This step consists of using some operations in the image processing in order to facilitate the subsequent processing, such as the Features extraction and the recognition in our work it includes the binarization, inversion of pixels, the framing of the object, and the normalization.

- The binarization: In this stage we have transferred a 225 level grey image to a double level grey image, in which the black pixels objects take the value 0 and the other pixels take the value 1. .
- The inversion of pixel: This method consists of inverting the pixels of the image, in which the object presented by white pixels, and the background of the image presented by black pixels.
- The framing of the object: This method consists to put the object of image in a frame
- The Normalization: This method consists in normalizing all the images of the database of the same size, A size of (60 60) has been chosen.

Figure 5 represents some example of the preprocessing steps.

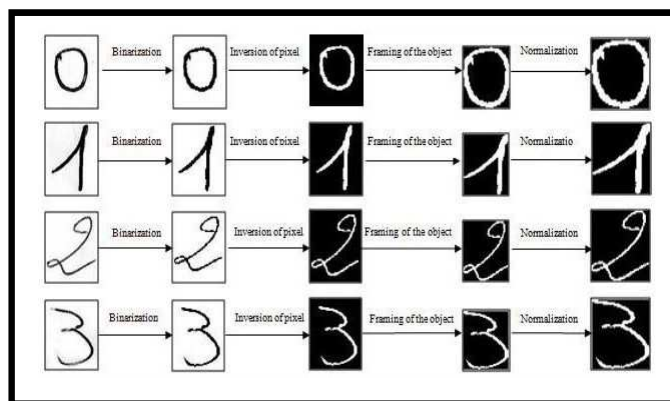


Fig.5: Example of the different preprocessing operations used for our system

5.2. The proposed Features extraction method

The method of feature extraction is applied after the Preprocessing operation to extract the most relevant characteristic of the character to be recognized. This method consists of representing the image object having the size ($N \times N$) by a vector V of the numeric parameters whose size is ($1 \times N$). Features extraction has been a field of intensive research recently, a considerable number of features extraction methods can be found in literature [1] [19], the main objective of using features extraction in recognition is not only to find several features extraction methods but which method is the best for recognition [1]. In our work, we have detailed five methods that will be presented in this section.

5.2.1. The Entropy Method

The entropy value is a mathematical function used to calculate the amount of information contained or provided by its source. This source can be a written text, an electrical signal or any other computer file. Considering a discrete source denoted S having a set of n different alphabets can be expressed in the following representation:

$$A_1, A_2, \dots, A_n$$

For each of these symbols, an emission probability associated is written as follows:

$$p(A_1), p(A_2), \dots, p(A_n)$$

The entropy of the source is then defined by:

$$H(s) = \sum_{i=1}^n p(A_i) \cdot \log(p(A_i))$$

- $H(S)$: The entropy value
- $P(A_i)$: The probability of the symbol A_i

In the proposed method, we have divided the global image into block of pixels of (6×6) and for each block we have calculated the entropy value for the pixels object and take it as a parameter for the primitives vector. The chosen image size is (60×60) , so the vector of primitives contains 100 primitives value. The function « entropy » of Matlab allows calculating the entropy value of each block. The proposed entropy method can be represented in figure 6.

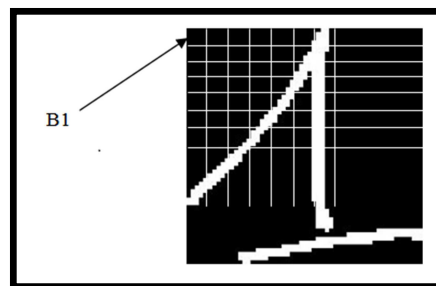


Fig.6: The proposed entropy method

We have;

$$H1 = \text{entropy}(B1)$$

Accordingly, the vector of primitives can be developed as follows:

$$V = [H1, H2, \dots, Hn]$$

Figure 7 represents the primitive vector parameters of the number “1” represented in figure 6.

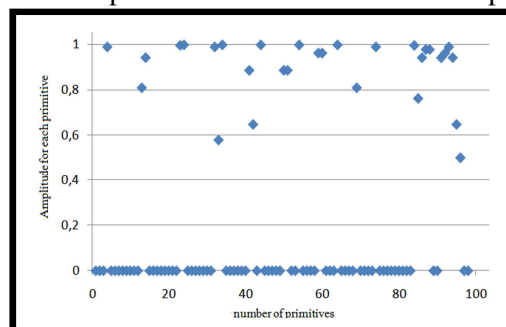


Fig.7: The vector of primitives for number “1” by the entropy method

5.2.2. The object follow-up method

This method consists of following the object as how it was drawn in the image concerning the four profiles as a function of lines and columns. For example: for the left profile of the image, we make a left to right scanning operation for all the image lines, and in each line-column intersection where a pixel object was found, we take the pixel column as a value of initial primitive vector V_i as shown below.

$$V_i = [X_1, X_2, X_3, \dots, X_n]$$

The vector of the final primitives V_f can be obtained with the following method. Each time, we compare the digit X_n with the digit X_{n+1} of V_i and for each:

- For $X_n < X_{n+1}$ we have $V_f(n) = 1$
- For $X_n > X_{n+1}$ we have $V_f(n) = 2$
- For $X_n = X_{n+1}$ we have $V_f(n) = 3$

At the end, the vector of the final primitives represents the object in the image with the digits 1, 2, and 3. The same work is done for the other profiles but for the high and low profiles, the initial vector V_i takes the line values.

Figure 8 explains how to extract the vector V_f for the left scanning by the object follow-up method.

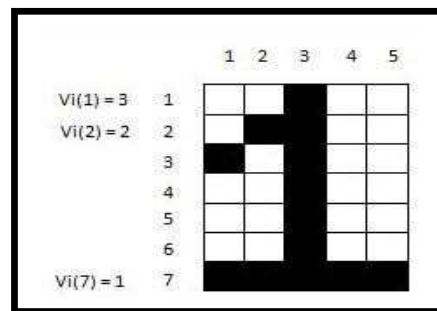


Fig.8: Example of extraction of the primitives by the object follow-up method

Tab.1: The initial vector V_i of the left profile

lignes	1	2	3	4	5	6	7
V_i	3	2	1	3	3	3	1

Tab.2: The final primitive vector of the left profile.

V_f	2	2	1	3	3	2
-------	---	---	---	---	---	---

In our study, the image size is (60 X 60), so for each profile, we have a vector of 59 primitives. Hence, the global vector contains 236 primitives. Figure 9 represents the vector primitive parameters of the number "1" given in figure 6.

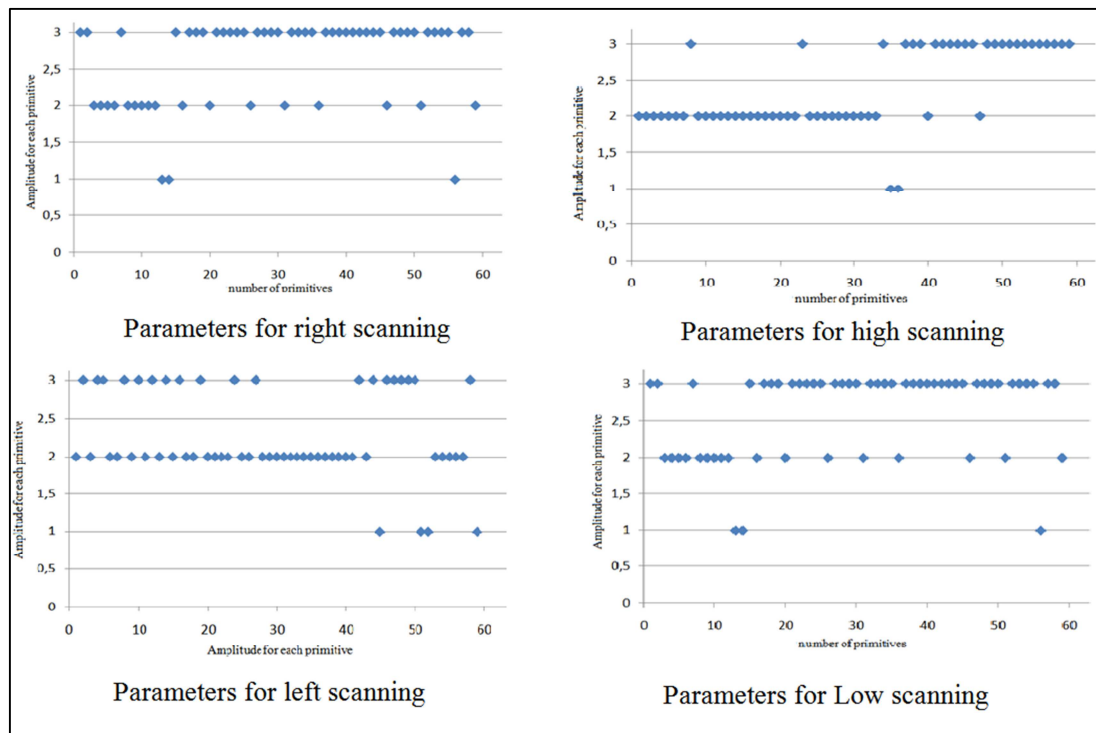


Fig.9: The vector of primitives for each profile the image for number “1” by the object follow-up method

5.2.3. The Dropping Method

In this method, we apply the dropping method for the four profiles of the image as illustrated in figure 10.

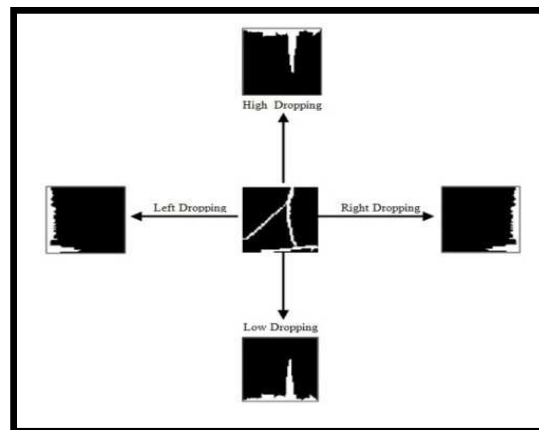


Fig.10: The Dropping Method

The primitive vectors take the values of pixel lines of the highest and lowest summits which represents the information for the two verticals droppings. In the same way, the pixels values of the top right and left columns which represent the information in the two horizontal droppings are presented. Accordingly, as in our work the image size is (60×60) , we have obtained a vector of 60 primitives for each profile, the global vector contains 240 primitives. Figure 11 represents the vector primitive parameters of the number “1” given in figure 6.

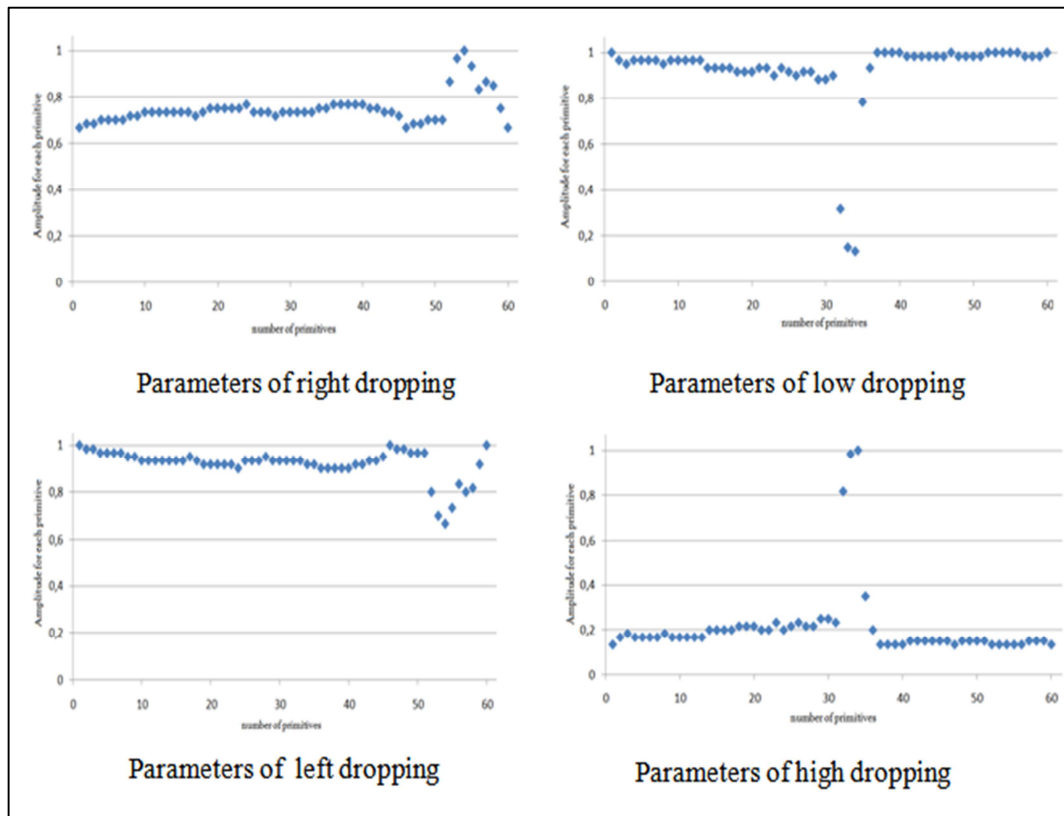


Fig.11.d Parameters of high dropping

5.2.4.The Average Line and Column Method

In this Method we have calculated for each line of image, the average values of the pixels columns values where the information is found. And for each column we have calculated the average value of the pixels lines values where the information is found. So, initially, we have made a line by line scanning and in each line we have taken the chain values of pixels columns where the information is found, then we have calculated the average value of this pixels chain and taken it as a primitive's value.

The same procedure is done in the column-by-column sweeping. However, in this case, we have calculated the average value of the chain values of pixels lines for each column. At the end, we have obtained, in each scanning, a vector of 60 primitives. Thus, the whole primitive's vector in this method contains 120 primitives. Figure 12 represents the vector of primitive parameters of the number “1” given in figure 6.

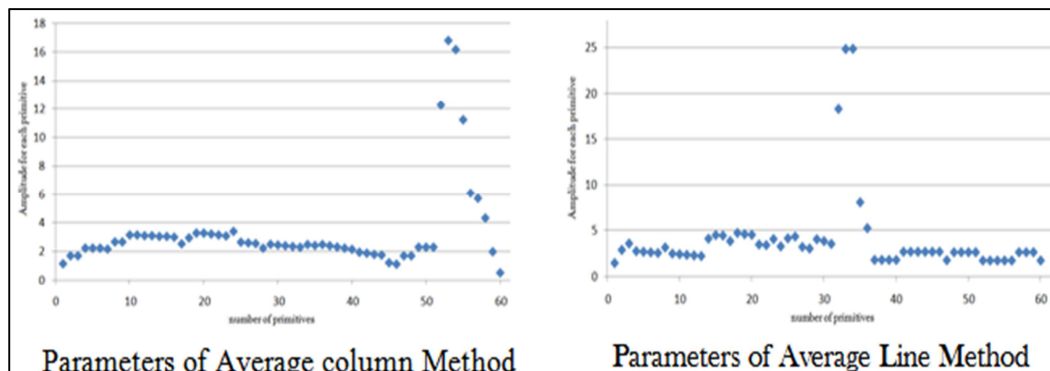


Fig.12: The vector of primitives for number “1” by the Average line and column Method

5.2.5. The Descriptors Method of the Normal Gaussian Law

The Normal Gaussian Law is an absolutely continuous probability law, which depends on two parameters, the expectation denoted μ , and the standard deviation denoted σ . The probability density of the normal distribution is given by:

$$f(x) = \frac{1}{\sigma \times 2\sqrt{\pi}} \times \exp\left(-0.5 \times \frac{(x - \mu)^2}{\sigma^2}\right)$$

In our method we have used the normal law of gauss in order to calculate the probability density of a centered and reduced random value ($\sigma=1$, $\mu=0$). Figure 13 represents the vector of primitives parameters of the number “1” as represented in figure 6.

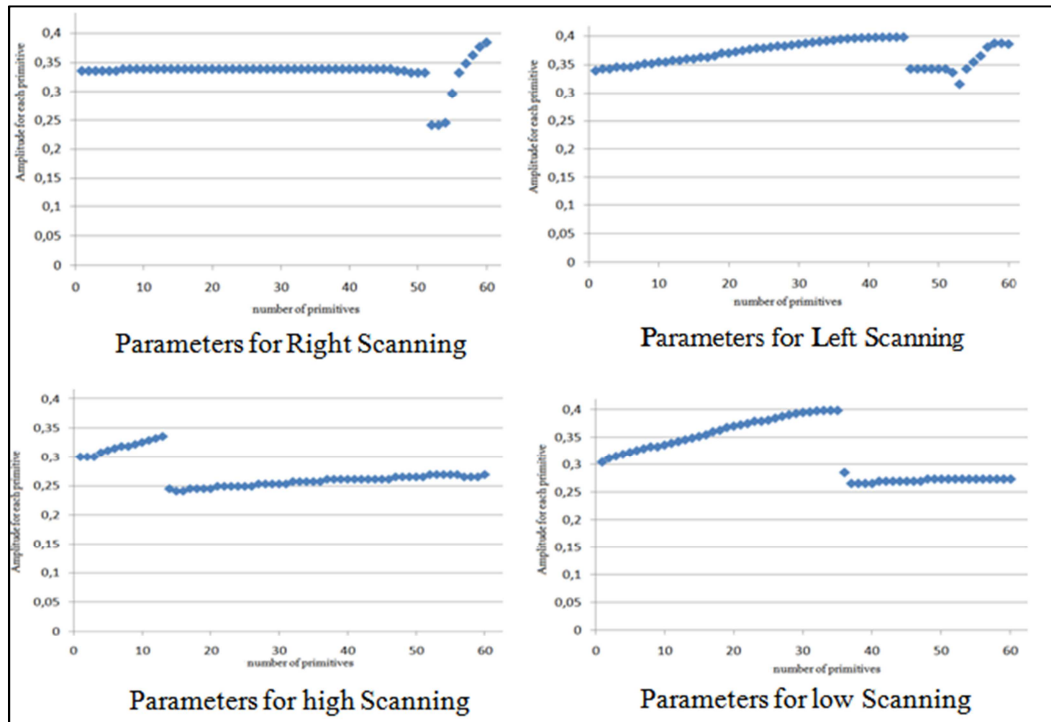


Fig.13: The vector of primitives for number “1” by The Descriptors Method of the Normal Gaussian Law

5.3. The Used Recognition Primitives Vectors

In our work, we have used five primitives” vectors in the learning and recognition steps for the two systems. These vectors are presented in table 3.

Tab.3: The used vector primitives

Primitives vectors	Input Layer	Hidden Layer	Output Layer
V1	100	80	10
V2	236	150	10
V3	240	170	10
V4	120	80	10
V5	240	150	10

6. The Recognition Step

The recognition step is divided into two parts, the first part is the learning phase, and the second part is the test phase. Our database contains 3000 images, we have chosen 2000 images for learning, while in the test phase we have used all database.

In the recognition step, we have trained the two systems by the five primitives vectors (mentioned previously), so we constructed for each vector of primitives two recognition systems, one by the neural network and the other by the Markov model. We have chosen a value of 0.01 for the two systems error rate.

6.1. System Combination

The combinations were made between the two recognition systems in the test phase for each vector of primitives, in that case the HMM helps the neural network for the manuscripts digits recognition. In the first step, the system tests all the database for the neural network, and in the second step, all the non- recognized manuscript digits are tested by the HMM.

6.2. Experimental Results

The following tables shows the recognition rate results for the three systems, the system of MLP, the system of HMM, and the hybrid MLP/HMM system.

Tab.4:Results of MLP systems.

Primitives vectors	R_R(%)	NR_R(%)
V1	93.13%	6.87%
V2	88.93%	11.07%
V3	90.30%	9.7%
V4	90.96%	9.04%
V5	96.7%	3.3 %

Tab.5: Results of the HMM systems

Primitives vectors	R_R(%)	NR_R(%)
V1	96.33%	3.67%
V2	95.86%	4.14%
V3	92.46%	7.54%
V4	95.96%	4.04%
V5	97.60%	2.4%

Tab.6: Results of the hybrid MLP/HMM System

Primitives vectors	R_R(%)	NR_R(%)
V1	56.3%	43.7%
V2	58.7%	41.3%
V3	31.7%	68.3%
V4	55.8%	44.2%
V5	36.07%	64 %

With:

- R-R: Recognition Rate
- NR-R: Non-Recognition Rate

This approach can also be used to combine the two systems for different primitive vectors, i.e. a combination between an MLP that uses a primitive vector, and an HMM that uses another primitive vector. The following table presents the best combination results.

Tab.7: Results of the hybrid MLP/HMM System using different vectors

PLM System	HMM System	R_R(%)	NR_R(%)
V5	V2	98.76%	1.24%
V4	V2	97.33%	2.67%
V1	V4	96.03%	9.97%

According to the results presented in tables (5 - 7), we observed in the first table, that the neural network is the best for the recognition compared to the Markov model, but in spite of the weak recognition of the Markov model, it helps the neural network to increase the recognition rate. It was also observed in table 8 that the system increases the recognition rate to 98.8% when the two systems are using different primitive's vectors.

7. Conclusion

This paper presents an hybrid MLP/HMM system for the recognition of numerical numbers with a set of primitive extraction methods. In the first step, each system is used for the recognition where the recognition rate is later on calculated.

In the second step, the two systems were combined to construct a hybrid recognition system (MLP/HMM) for each primitive's vector where their recognition rate was also calculated. At the end, a hybrid approach (MLP/HMM) was applied between the MLP using the primitives vector and the HMM using another primitives vector, in which the recognition rate was raised up to 98.8%.

Consequently, it has been observed that, in order to increase the recognition rate, it is necessary to increase the database size by taking into account a larger number of writers and writing instruments. Moreover, it is advised to use combinations between other recognition systems, and to use other extraction methods of primitives.

8. References

- [1] Ouchtati. S, Mouldi. B, and Lachouri. A. (2008) „Segmentation and Recognition of Handwritten Numeric Chains”, *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, Vol. 2 No 12, pp. 4139-4146
- [2] Boukharouba. A and Bennia. A. (2011) „Recognition of Handwritten Arabic Literal Amounts Using a Hybrid Approach”, *Cogn Comput*, Vol. 3, pp. 382– 393
- [3] Mantas. J. (1986) „An overview of character recognition methodologies, *Pattern Recognition*, Vol. 19 No. 6, pp.425–430.
- [4] Mori. S, Suen. C. Y and Yamamoto, K. (1992) „historical review of OCR research and development”, *proceedings of the IEEE*, Vol. 80 No. 7, pp. 1029-1058.
- [5] Koerich. A.L, Sabourin. R, Suen. C.Y and El-Yacoubi, A. (2000) „A Syntax Directed Level Building Algorithm for Large Vocabulary Handwritten Word Recognition”. Paper Presented at the *4th International Workshop on Document Analysis Systems*

- [6] Oliveira. L.S, Sabourin. R, Bortolozzi. F and Suen. C.Y. (2001) „A Modular System to Recognize Numerical Amounts on Brazilian Bank checks”. Paper Presented at the *6th International Conference on Document Analysis and Recognition*
- [7] Filatov. A, Nikitin. N, Volgunin. A and Zelinsky. P. (1998) „The Address Script TM recognition system for handwritten envelopes”, Paper Presented at the *In International Association for Pattern Recognition Workshop on Document Analysis Systems*
- [8] El-Yacoubi. A. (1996) Markov Modeling of Handwriting Application to Recognition of Postal Addresses. Unpublished PhD thesis, Rennes University, Rennes, France.
- [9] Hu. J, Brown, M. K and Turin. W. (1996) „HMM based on-line handwriting recognition”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, pp. 1039-1045
- [10] Buse. R, Liu . Z-Q and Caelli. T. (1997) „A structural and relational approach to handwritten word recognition”, *IEEE Trans. Systems, Man and Cybernetics*, Vol. 27, pp. 847-861
- [11] Cai. J and Liu. Z-Q. (1999) „Integration of structural and statistical information for unconstrained handwritten numeral recognition”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, pp. 263-270
- [12] Amin. A. (1998) „Off-line Arabic character recognition – the state of the art”, *Pattern Recognition*, Vol. 31, No. 5, pp. 517-530
- [13] Touzet. C. (1992) *The Networks of Artificial Neurons Introduction to Connections. Courses, exercises and practical work*
- [14] Sodanil. M, Nitsuwat. S and Haruechaiyasak. C. (2010) „Thai Word Recognition Using Hybrid MLP-HMM”, *International Journal of Computer Science and Network Security(IJCSNS)*, Vol. 10 No. 3
- [15] Conéujol. A, Miclet. L and Kodrattof. Y. (Eds), (2003) *Artificial Learning, Edition EYROLLES*.
- [16] Russel. M. Moore, J, and R. K. (1985), „Explicit Modeling of State Occupancy in Hidden Markov Models for Speech Recognition, Proceeding of ICASSP”, *International Conference on Acoustic, Speech and Signal Processing*, pp.5-8
- [17] Levinson, S.E. (1986) „Continuously Variable Duration Hidden Markov Models for Automatic Speech Recognition”, *Computer, Speech & Language*, Vol. 1 No. 1, pp: 29-45
- [18] Kadir. A and Rahman. M. (2016) „Bangla speech sentence recognition using hidden Markov models”, *International Journal of Multidisciplinary Research and Development*, Vol. 3 No. 7, pp. 122-127
- [19] Grandidier. F. (2003) *A new algorithm for selecting application characteristics for automatic reading of handwriting*. Unpublished PhD thesis, School of Advanced Technology, University of Quèbec Canad