A REAL-TIME MULTIPLE-CAMERA OPTICAL CHARACTER RECOGNITION AND VERIFICATION SYSTEM

Tolga OVATMAN[1] **Fatih KAHRAMAN**[2] **Osman Kaan EROL**[3]

[1,3] *Computer Engineering Department*
Faculty of Electrical and Electronic Engineering
İstanbul Technical University, 34469, Maslak, İstanbul

[2] *Informatics Institute*
İstanbul Technical University, 34469, Maslak, İstanbul

[1] *e-mail: ovatman@cs.itu.edu.tr* [2] *e-mail: kahraman@be.itu.edu.tr*
[3] *e-mail: okerol@ cs.itu.edu.tr*

Keywords: Optical Character Verification(OCV) , Artificial Neural Networks , Real-Time Systems , Performance of Artificial Neural Networks

ABSTRACT

OCR systems are used in real world in many areas especially in document digitalizing. Industrial applications also use OCR to digitalize the printed data on any kind of media. Printing industry uses OCR for quality control of the printed products. In this paper, a real time automation system is examined, which is designed for verifying the numbers that are grouped and printed on certain areas of paper plates. The main subject of interest of the paper is artificial neural network(ANN) engine implemented in the system and the performance consideration of different ANNs. The dilemma of minimizing the character recognition time and maximizing the recognition rate is analyzed by implementing an ANN within ANN approaches in pattern recognition.

1 INTRODUCTION

Quality control and verification of the printed media is an important issue for printing industry. The synchronization problem of the automation system and the digital printing system sometimes yields to unwanted results that affects the quality of the work and decrease the throughput of the system.

OCR systems are used in printing industry to automatically and precisely control the printed media and help to eliminate the affected products of the printing system. This is often done real time where the production rate is important. The real-time OCR concept comes from the limited time where the printed media should be digitalized immedieately before automation system presents new inputs to OCR system.

This paper investigates ways to reduce the processing time of the recognition system by experimenting different ANN approaches in the implemented real time OCR system. The inputs and results are proven to be valid for practice rather then sticking to theory, since the system that is used for experimentation is real.

2 OPTICAL CHARACTER RECOGNITION

Optical character recognition(OCR) involves computer systems designed to translate images of typewritten text into machine-editable text to translate pictures of characters into a standard encoding scheme representing them (ASCII or Unicode). OCR began as a field of research in artificial intelligence and machine vision; though academic research in the field continues, the focus on OCR has shifted to implementation of proven techniques.

Optical character recognition includes two sub operations:

a) Segmentation: The process of obtaining and separating the printed characters.

b) Recognition: The process of acquiring digital match of the input from the image obtained by segmentation

3 REAL-TIME OCR SYSTEM

3.1 General Architecture of the Base Automation System

The system that has been implemented is designed to work synchronous with a printing automation system that prints 13 digit numbers on cards that are located in plates of papers. The system prints on 5 cards (prints five 13 digit numbers) at the same time.

The conveyor system that the cards flow can work in different automation levels. In each level the printing heads works synchronously with the system and responses to speed changes which is tuned by the users of the system.

Printing system is responsible of printing the numbers, that exist in a database, on the cards that are loaded to conveyor system. Unfortunately the printing system sometimes fail to cooperate with the conveyor system and prints wrong numbers form the database or do not print anything on a card.

The failure rate of the printing system increases as the speed of the conveyor band increases. In this prospect the real time OCR system is expected to detect the errors of the printing system by verifying the numbers in the database with the numbers that are read from the cards.

3.2 Limitations of the System

The limitations of the system to be designed can be handled in two prospects

a) Financial Limitations: In order to reach beyond the current systems, cameras with lower image qualities have to be selected.

b) Time Limitations : The target of the system for the maximum speed level is one plate per second. This means twenty five 13 digit numbers in a second. By the experimental results it is determined that 3.1 ms is required for each digit to be segmented and recognized.

3.3 Hardware Architecture of the System

There are three hardware units that are used to realize the system in Figure 1.

a) Cameras: Cameras with IEEE 1394 standard (FireWire) is used in the system. FireWire cameras are chosen because of two reasons. First reason is high data transfer rate up to 400mb/s . Being able to acquire multiple camera images at the same time is the second reason.

b) Synchronization unit: A unit that sends a CTS signal from the serial port for every column of a plate, with the help of the marks at the bottom of the plates.

c) Application computer: Chosen as a simple PentiumIV home PC.

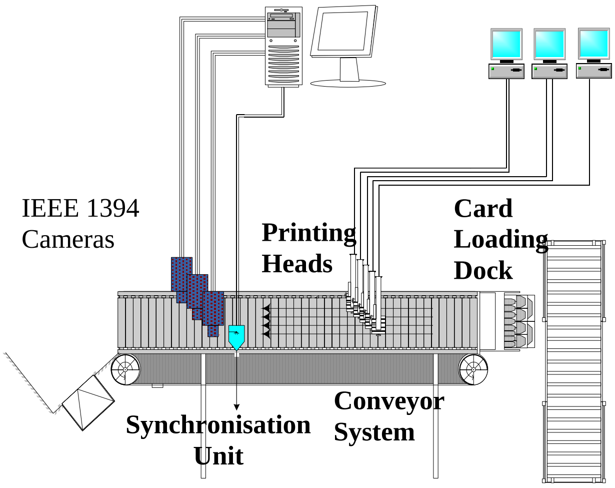

Figure 1. Overview of the designed system

3.4 Software Architecture of the System

Software system initially carries the camera images to the user of the system and lets the user define the database of numbers and plate properties to the system. Region of Interests(ROI) are one of these data which helps the software to apply the techniques in order to perform OCR. ROIs are ought to be the regions that contain the printed numbers on the plate on which the OCR techniques will be applied.

After all the data about the system has been taken, the system starts to run and the cameras take a snapshot for every column of a card plate that flows under the synchronisation unit. From every snapshot the ROIs are extracted and the OCR operation is applied to these regions. At the end of the OCR operation numbers on the cards are obtained and verified with the database.

The system doesn’t need %100 performance in OCR because on every card a non consecutive random number is printed, which means over a %80 match between the OCR and the real number will be acceptable for the system.

In order to implement the character recognition engine, a segmentation system is developed that simply uses vector quantization[1] to separate the background and the numbers in the ROIs. The results of the segmentation system of the developed software can be seen in Figure 2. As it can clearly be seen, the images acquired from the cameras are dramatically eroded.

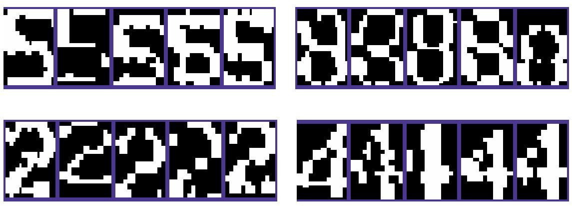

Figure 2. Some segmentation results

4 ARTIFICIAL INTELLIGENCE AND PERFORMANCE CONSIDERATIONS OF THE SYSTEM

Artificial neural networks are the most widely used pattern recognition systems in OCR. Their self-learning capabilities provide a utilization in a wide range of pattern recognition tasks. Therefore a number of ANNs are experimented in order to use in the real time OCR system. These ANNs are briefly explained and the experimental results are presented in the following sections.

The experiments on the ANN system of the real-time OCR are established with a set of near 700 training images and near 200 test images chosen from the real segmentation results that are seen in Figure 2.

4.1 Multilayer Perceptrons(MLP) and Backpropagation

Multilayer perceptrons have the ability to solve the problems which are not linearly separable. This is done by adding a group of neurons which is called a “layer” between the input and the neurons which decide the behaviour of the network, as it is seen in Figure 3. This layer is called a hidden layer and there may be many hidden layers in an MLP.

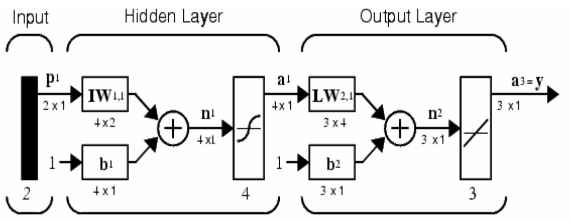

Figure 3. Structure of a Multi layer perceptron

Experimental results:

A single hidden layered neural network is experimented because of its high learning rates and having backpropagation as a powerful learning mechanism behind it.

But before all, the problem of having 24x36=864 pixels in the segmented images as neural network inputs yields to unwantedly large neural networks and low performance of the system. In order to overcome this problem, horizontal and vertical projections of the images in Figure 4 are calculated and the number of features for each image is reduced to 60. With the help of this feature reduction, 60 neurons are selected to be in the hidden layer of the system because of its utilization of the trade-off between performance and ANN size.

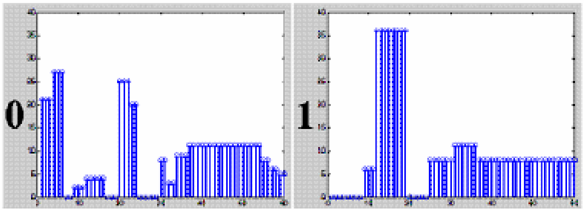

Figure 4. Horizontal and vertical projections of two numbers

With the projected inputs, the neural network has been trained with scaled conjugate gradient algorithm[4] and the results obtained with this technique is seen in Figure 5. The MLP performed %85 of successful recognition with the test set. Each disordered column is a wrong output.

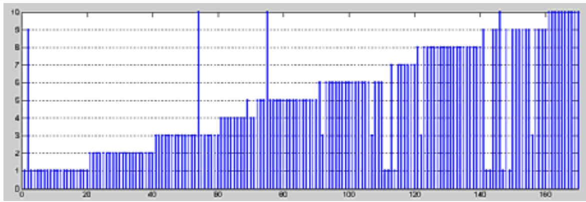
Figure 5. Experimental results for MLP. Every step of the ladder stands for a number starting from 0 ending with 9

Calculating the operation cost of ANN , it has 60 neurons in hidden layer and 60 inputs thus there are 60*60=3600 connections between these layers. There are 60*10=600 connections between hidden layer and output layer. As a result the network have **3600w+70n** operation cost where ‘w’ is cost for a floating point multiplication and ‘n’ is the cost of a neuron operation.

In order to overcome the feature reduction problem PCA[3] algorithm is applied to the input set. PCA method reduced the number of features in the input set to 116 features. This time running backpropagation on an ANN with 116 neurons, the MLP performed a %98.8 successful recognition rate that is seen in Figure 6.

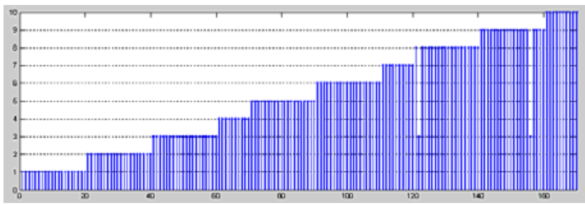
Figure 6. Experimental results for MLP with inputs from PCA

But this time the operation cost of the network becomes **14616 w + 70 n + P** where P is the PCA feature extraction cost from an ordinary input.

4.2 Adaptive Resonance Theory

It was the analysis of the Grossberg Network's[5] instability that led the researchers to develop ART system. ART model in Figure 7 uses shunting model neurons[6] and is developed in such a way that they are stable enough to preserve significant former data , yet remain adaptable enough to incorporate new information whenever it might appear.

The ART1 network[7] is designed as a real-time system. All the weight updates and other mechanisms are described as differential equations so that ART1 can be realized using analog circuits. By this property it is worth running experiments on ART1 to see its performance. If the performance is satisfactory enough, an unsupervised learning mechanism as a character recognition system can be embedded to system's hardware.

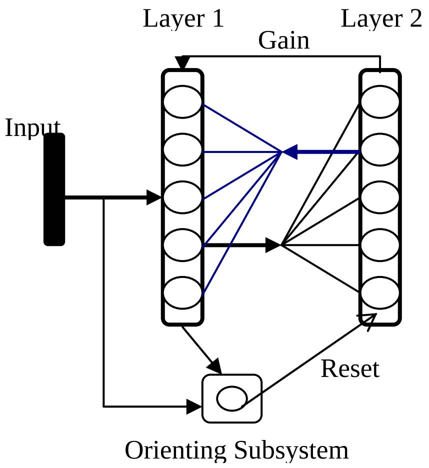

Figure 7. Structure of an ART network

Experimental results:

ART1 networks only accept binary input patterns, so feeding the network with the original input images without any feature extraction is more suitable for the network's nature, but on the other hand 864 dimensional input vectors are stil too big. In order to shrink the dimensions of the input image, the bits in the binary image are grouped in 4x3 sized windows and the window is carried to the shrinked image as a '1' if it contains more than 4 full pixels. By this way the size of the image is reduced to 6x12 and the input vector size is reduced from 864 to 72 as seen in Figure 8.

Using ART1, %95 performance on classification of inputs is obtained. But this was not enough since the ART1 clustered 27 groups while it ought to cluster 10.

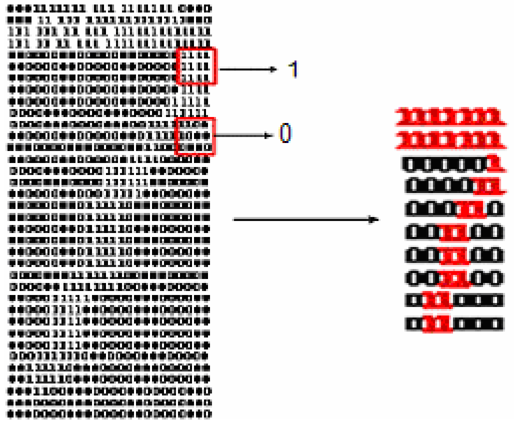

Figure 8. Algorithm that is used to shrink the image that will feed ART1

An example output is given next where the patterns between 0-70 represent '0's. As it is seen '0's are grouped into two classes where only two unwanted patterns remain in the classes.

Class	Pattern
1 -	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 27, 30, 32, 33, 35, 36, 37, 40, 41, 42, 43, 50, 68,
2 -	23, 26, 28, 29, 31, 34, 38, 39, 44, 45, 46, 47, 48, 49, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 69, 70, 71, 262, 265,

At this point, a new layer in front of the output layer of the ART1 is inevitable while an ANN model that performs this operation is presented in the next section. ART1's operational cost calculated as **10798 w + 145 n + C** where C is the cost for classification.

4.4 Learning Vector Quantization

LVQ Networks[8] are hybrid systems that both include unsupervised and supervised learning. This way, the pattern recognition power of the unsupervised learning networks can be joined with the classification power of the supervised learning networks.

In a LVQ network in Figure 9, there exists two layers. The first layer is the competitive layer which classifies inputs according to the similarity between them. This layer works unsupervised. The first layer of the LVQ is responsible for classifying regions from the input space, discarding the similarity between these regions.

With the acquisition of the inputs, the need for normalization is eliminated because the distance of the weight vectors are calculated automatically.

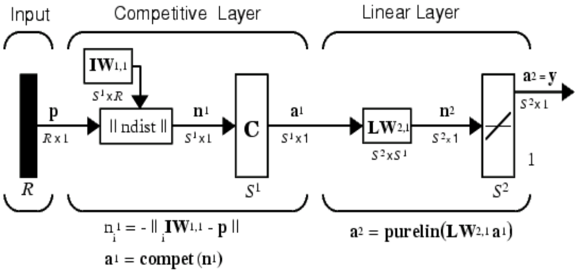

Figure 9. Structure of a LVQ network where R is the input vector size, S^1 is the number of competitive neurons and S^2 is the number of linear neurons

The first layer runs a competition among its inputs and the neuron that has the closest weight vector to the input will output a ‘1’ and the other neurons will output ‘0’.

Each winning neuron in the layer 1 indicates a subclass for the layer 2. Layer 2’s task is to combine these subclasses into the classes that are supplied by the supervised learning’s parameters.

The experiments with LVQ are carried out with the real input’s projected input vectors in Figure 4, which were also used in MLPs, that provided 60 dimensional input vectors.

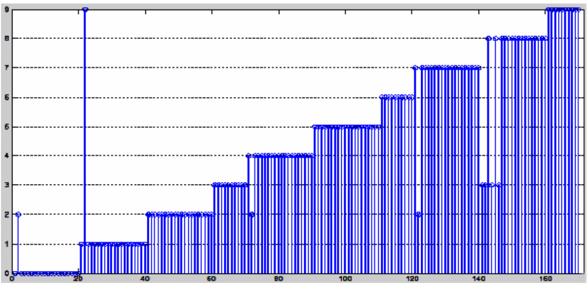

Figure 10. Experimental results for LVQ

At the end of experiments with LVQ network the same %95 success rate is obtained, like the unsupervised ART1 in the previous section. But this time as the result of the supervised layer the clustering is done precisely and the true %95 success rate is reached as can be seen in Figure 10.

Since there exists 20 neurons in the competitive layer of the LVQ, 1200 connection weights exists. But there is also the distance operation before feeding the first layer which is nearly equal to 9000 floating point multiplications. Using these values the operational cost of LVQ becomes **9200w + 30n** .

5 CONCLUSION

The trade-off between the success of pattern recognition and the performance of ANN is reached. This trade-off is also stated by the literature[2] as the main obstacle in the way of using ANNs in real world problems.

Method	Cost	~Success	~# of operations
MLP	4200w + 70n	%85	13k
MLP+PCA	14000w + 70n + P	%99	142k
ART1	11000w + 145n + C	%95	46k
LVQ	9200w + 30n	%95	30k

It can easily be seen from the table that the most successful method is the PCA+MLP but at the same time it requires the highest operation capacity for the system to accomplish. With the current system with an MLP as the recognizing engine, it takes 6.2 ms to segment and recognize one digit.

As the conveyor system running at full speed will require 3.1 ms of time to segment and recognize a digit, it is not possible to use any other methods than MLP in the system. Even when the MLP is being used, the conveyor system is able to run at half speed, so does the production.

In order to overcome the performance problem, updating the system with a more powerful hardware is the most robust solution.

Also ANN pruning techniques or evolutionary programming can be used to design a network with less neurons and less weights, but this solution seem to provide mutually less performance with respect to a hardware update.

6 REFERENCES

[1] R. M. Gray, ``Vector Quantization,'' *IEEE ASSP Magazine*, pp. 4--29, April 1984.

[2] R. P. Lippmann , “Pattern Classification Using Neural Networks”, IEEE Communications Magazine , November 1989

[3] I.T. Jolliffe, “Principle Component Analysis”, Spring-Verlag, New York, 1986.

[4] Martin F. Møller,“A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning”, November 1990

[5] S.Grossberg, E.Mingolla, D.Todorovic, “A neural network architecture for preattentive vision”, IEEE Transactions on Biomedical Engineering, vol.36 no.1 January 1989

[6] M.T. Hagan, H.B. Demuth, M. Beale, “Neural Network Design”, PWS Publishing Company, 1996

[7] G.A. Carpenter, S.Grossberg, “A massively parallel architecture for a self organizing neural pattern recognition machine”, Computer Vision, Graphics, and Image Processing, vol.37 1987

[8] T. Kohonen, “Self Organization and Associative Memory”, New York:Springer-Verlag, 1989