

Grid Infrastructure Based Intelligent Character Recognition: A Novel Algorithm for Extraction of Handwritten and Typewritten Characters Using Neural Networks

Pasikanti Susruthi Divya Sruthi*

M.Tech (CN&IS), Department of IT,
SNIST, Yamnampet, Hyderabad, AP, India

S. Durga Devi

Assistant Professor, Department of IT,
SNIST, Yamnampet, Hyderabad, AP, India

Abstract-

Image Processing is a field under Digital Signal Processing. It is considered to be a favorite topic in the IT department nowadays. Intelligent Character Recognition is one of its major applications. The purpose of Intelligent Character Recognition (ICR) system based on a grid infrastructure is to effectively and efficiently perform Analysis of Document Image, processing of document in electronic formats converted from paper formats. Compared to various existing available character recognition systems it improves the accuracy of recognizing the characters during document processing. Since our character recognition is based on a grid infrastructure, heterogeneous characters of different universal languages with different font properties and alignments are recognized easily. We proposed a novel algorithm in this paper to extract text/characters from a scanned image using neural networks by using a method for combining independently trained networks to achieve higher performance. Proposed system consists of the following steps 1) Image Processing 2) Image Training 3) Image Recognition 4) Image Editing and 5) Image Searching. In our paper it is shown that, the proposed system is better than the existing systems and try to improve the efficiency and accuracy of recognizing the characters from a scanned image.

Keywords- Intelligent Character Recognition, Grid infrastructure, Document Image Analysis, Feed Forward Neural network, Kohonen Neural network, Unsupervised Training, Supervised and Image Recognition.

I. INTRODUCTION

The term Digital Image Processing generally refers to the processing of a two-dimensional picture by a digital computer i.e. altering an existing image in the desired manner. Digital Image Processing is a swiftly developing field with the growing applications in science and engineering. Image Processing has the capability of developing systems which can perform all the visual functions of living things. The first step in image processing is to get an image. Basically an image an arrangement of picture elements (pixels) thus a digital image is an array of real & complex numbers represented by finite number of bits. In the present world, there is demand growing for the software systems to recognize characters in computer system when information is scanned through paper documents as we know that we have number of newspapers and books which are in printed format related to different subjects. Manual data entry from hand-printed papers is very time consuming - more so in offices that have to deal with very high volumes of scanned papers. Now a days there is a lot of demand for storing and searching the printed information available in papers as images in computers but the problem is that the font characteristics in the paper documents are different from font in the system, which makes the system unable to recognize the characters. The process of storing the documents in computer and then processing the data is called Document Processing. In order to overcome this problem a software called as Document Image Analysis (DIA)/ Character Recognition is used. It is applicable to many languages.

To perform Document Image Analysis, different kinds of Character recognition techniques are available in the world. In this paper, we have chosen Grid Infrastructure based Intelligent Character Recognition to recognize characters. Neural networks like Feed Forward network and Kohonen network are used for training and recognition purpose. The classification performance is good provided the handwritten character entry in the document is accurately extracted and appropriately presented to the neural network classifiers. Good preprocessing techniques preceding the classification process can greatly enhance recognition accuracy. Grid infrastructure is used in Intelligent Character Recognition for character recognition in-order to perform Document Image Analysis (DIA). Thus the system is useful and effective in Virtual Digital Library's design and construction.

It is always fascinating to be able to find ways of enabling a computer to mimic human functions, like the ability to read, to write, to see things, and so on. ICR enables the computer or machine to visualize a image and extract text from it such that it can edit the text, store it, display or print it. The conversion of paper documents in to editable text is an on-going task in many of the organizations particularly in Research and Development (R&D) area, in government institutions, in large business enterprises, so on. Intelligent Character Recognition in electronic devices such as digital cameras, cell phones can be used to obtain images and recognize them as a part of face recognition and validation.

II. RELATED WORK

Automatic character recognition improves the interaction between man and machine in many applications like mail sorting, cheques verification, office automation and a large variety of banking. Character recognition methods include statistical, semantic, neural network, pattern recognition, etc.

A. Statistical Approach

A statistical approach looks for a typical spatial distribution of the pixel values that are represented by a spatial configuration of “0”s and “1”s which characterize each character. In this approach it searches for the statistical characteristics of various characters. These characteristics could be very simple, like histogram, the ratio of width and height ratio, black pixels to white pixels, etc. or complex, like higher order statistical parameters such as the third moments of the image. Some of handwritten characters indeed consist of pixels. Statistical methods ignore is that the pixels also form lines and contours. Statistical based character recognition consists of following steps.

- 1) Calculate the statistics of the character image and
- 2) Compare them with predefined statistics of the characters.

Most of the handwritten characters vary from ideal or predefined characters. Thus, there is a need for additional features to this approach to overcome this problem.

B. Semantic Approach

A semantic approach recognizes the way in which the contours of the characters are reflected in the pixels that represent them and try to find out typical characteristics for each character. Semantic data: open, loop, black stroke count, end point, etc. The following steps are involved in Semantic approach for character recognition.

1. Identify contour's starting point,
2. Trace the contour,
3. Identify the characteristics of the contour,
4. Compare the obtained characteristics with the similar descriptions in the database.

The main problem in this method is that it depends on the correct extraction of character contours correctly and this approach may fail when a character image is broken, but a statistical approach could still reflect broken character image statistics with sufficient accuracy to enable correct identification.

C. Hybrid Approach

It is clear that statistical and semantic approaches to character recognition have specific advantages and disadvantages. hybrid approach overcomes the problems associated with the statistical and semantic methods when utilized independently. The idea of Hybrid method in order to compensate their individual weakness and to prevent their individual strength has been widely used in character recognition field. In Hybrid approach training time is very quick receiving the high accuracy rate.

III. PROPOSED SYSTEM

Our proposed system is grid infrastructure based Intelligent Character Recognition which is a system that supports recognition of the characters of multiple languages. The multiple functionalities include editing and searching too whereas the existing system supports only editing of the document. In this context, Grid infrastructure means the infrastructure that supports group of specific set of languages. Thus ICR on a grid infrastructure is multi-lingual. This feature is what we call grid infrastructure which eliminates the problem of heterogeneous character recognition and supports multiple functionalities to be performed on the document. The benefit of proposed system that overcomes the drawback of the existing system is that it provides heterogeneous characters recognition and supports multiple functionalities such as editing and searching.

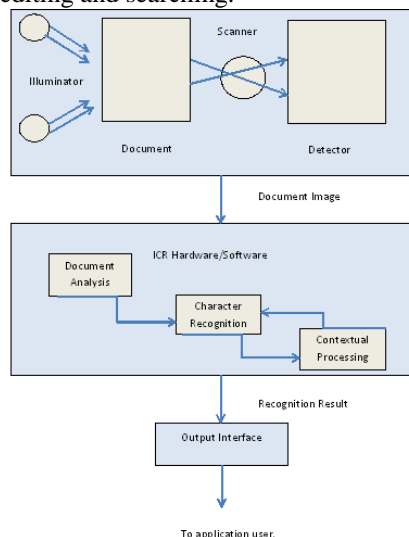


Fig.1 Architecture of the proposed system.

A. Architecture of Proposed System

The three main components in the architecture of the grid infrastructure based Intelligent Character are 1) Scanner 2) ICR Hardware or Software and 3) Output Interface.

IV. WORKING OF PROPOSED SYSTEM

The basic steps involved in Grid Infrastructure based ICR are:

- 1) Image Processing
- 2) Image Training
- 3) Image Recognition
- 4) Image Editing and
- 5) Image Searching.

A. Image Processing

The activities involved in this step are 1) Scanning an Image, 2) Saving the image, 3) Pre-processing the image, 4) Recognizing the characters in the image and 5) Generating grid infrastructure data structure. In the recognition process, it uses the ICR method in support of grid infrastructure datastructure.

1) *Preprocessing Techniques:* The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions or enhances some image features important for further processing. Any image needs some Pre-processing, before being fed to the recognition system. Conversion of any kind of image into a Binary image is the first step, and then follows preprocessing techniques like Skew Detection and correction, cleaning, smoothing, enhancing etc.

a. *Skew Detection and Correction:* Skew detection is one of the first operations to be applied to scanned documents when converting data to a digital format. It aligns an image before processing because text segmentation and recognition methods require properly aligned next lines. Different algorithms of skew detection are implemented. In Scan line based skew detection method, image is projected at several angles and the variance in the number of black pixels per projected scan line is determined.



Fig. 2 Images before and after skew and shift correction

b. *Smoothing:* Smoothing is used for noise reduction and blurring in an image. Removal of small details from an image is done using blurring. Smoothing operations are used to reduce the noise or to straighten the edges of the characters. binary (black and white) document images, filtering is used for smoothing and noise removal. There are two types of filtering approaches: linear and nonlinear. Filtering is a process, in which the value of any given pixel in the output image is determined by applying some algorithm to the values of the pixels in the neighborhood of the corresponding input pixel. For smoothing purpose masks are passed over the entire image, and this process is repeated until there is no change in the image. The centre pixel of the mask is called as the target. Pixels are ignored which are overlaid by any square marked “X”. The target pixel is forced to match them to have the same value if the pixels overlaid by the squares marked “=” all have the same value, that is, all zeros, or all ones, otherwise it is not hanged. Single bumps or single pixel indentation in all the edges, can be filed or removed using masks.

B. Image Training

The next step after Image processing is Image Training. Before converting the printed images in to editable and searchable documents, the mandatory step is providing training to the system. Training is to make system understand the font followed in the scanned document. All the characters that are required for recognition from the scanned document should be typed and saved as an image file. Then the image file should be provided as an input during this process and training button is clicked. Now training gets completed and the system gets familiar with the new font. It supports both Training the system with the pre-defined fonts and with the new fonts that are not present in the system and that cannot be identified by the system.

Two types of training that can be employed with a neural network are 1) Un-Supervised Training and 2) Supervised Training. In unsupervised training, training sets are present in neural network, but anticipated output is not provided whereas in Supervised training it provides the neural network with training sets and the anticipated output. Unsupervised training is a very common training technique for Kohonen neural networks. The main difference between supervised and unsupervised training is that the expected outputs are provided in supervised training. It allows the supervised training algorithm to adjust the weight matrix based on the difference between the actual output and the anticipated output of the neural network.

1) *Kohonen Neural Network*: The Kohonen neural network contains only an input and output layer of neurons, but there is no hidden layer in a Kohonen neural network. The input patterns presented to the Kohonen neural network are the dot image of the character that was hand written. The Kohonen neural network should classify the input pattern into one of the 26 input patterns. The 26 output neurons correspond to the 26 English alphabets. As the Kohonen neural network is trained the weights should be adjusted so that the input patterns are classified into the 26 output neurons. This method results in effective character recognition. The Kohonen neural network does not use any sort of activation function or any sort of a bias weight. When a pattern is presented to a Kohonen network one of the output neurons is selected as a "winner". Often these "winning" neurons represent groups in the data that is presented to the Kohonen network. Input neurons are used to give input to a Kohonen neural network. Floating point numbers are given to these input neurons that make up the input pattern to the network. These inputs must be normalized to the range between -1 and 1 in Kohonen neural network. Presenting an input pattern to the network will cause a reaction from the output neurons. The training process for the Kohonen neural network is competitive. For each training set one neuron will "win". This winning neuron will have its weight adjusted so that it will react even more strongly to the input the next time. The ability to recognize that particular pattern will be increased as different neurons win for different patterns,.

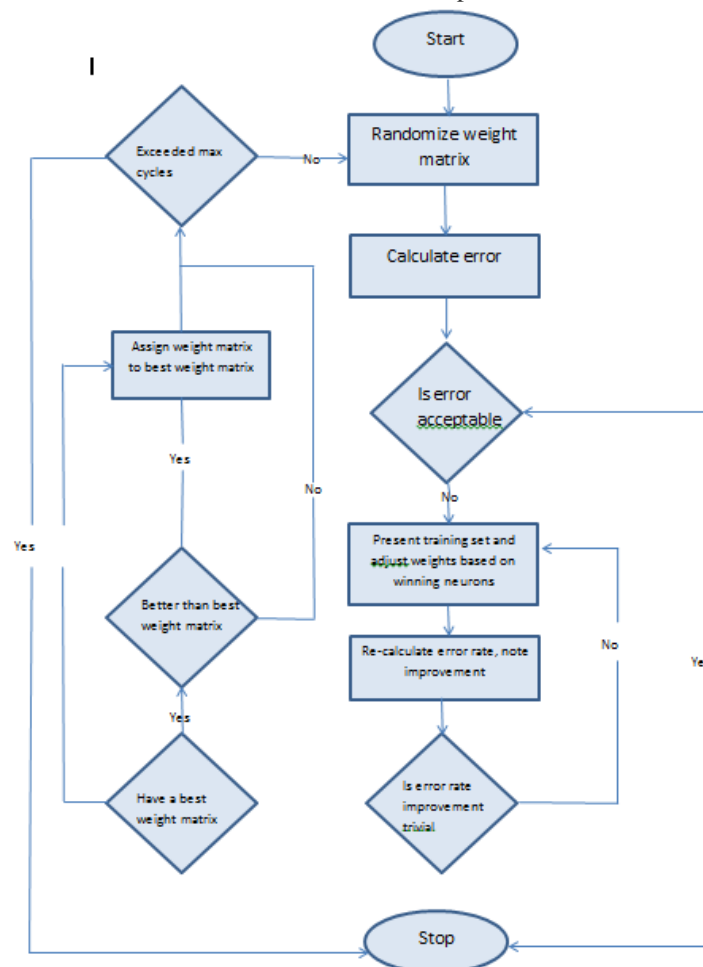


Fig. 3 Working of Kohonen neural network

2) *Feed Forward Neural Network*: A "feed forward" neural network is similar to the other neural network. The feed forward neural network begins with an input layer just like many other neural network types. This input layer is connected to a hidden layer and this hidden layer can then be connected to another hidden layer or directly to the output layer. At least one hidden layer is provided but there can be any number of hidden layers. Generally only one hidden layer is used in neural networks. Very rarely two or more hidden layers are used in a neural network. The input layer to the neural network is the conduct through which the external environment presents a pattern to the neural network. Pattern to the external environment are presented using the output layer of the neural network.

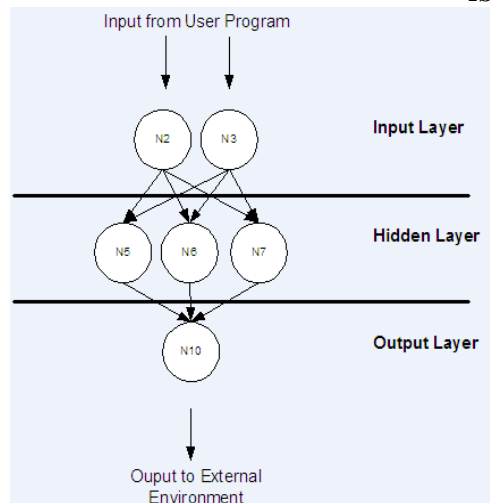


Fig. 4 Feed Forward Neural network architecture

C. Image Recognition

In this step two types of recognitions are possible. They are Hand written and Scanned document. In Hand written recognition the characters of any language can be recognised due to the adaptation of grid infrastructure. In handwritten recognition, the handwriting of the user in any language is trained to the system only for the first time. The system recognizes the characters or words there onwards. Whereas in scanned document recognition, font characters training need to be given first to the system. Then, image document is uploaded, text zone is selected and characters are extracted allowing the user to edit and search the text. It is possible to recognize Heterogenous characters in this system.

- 1) *Grid Infrastructure Formation*: Each character consists of certain geometrical similarities which are local to a particular area. For example consider alphabet 'O' and 'Q'. These two characters have similarities in particular areas. However when handwritten characters are considered these properties might change slightly. By using grid infrastructure information about these properties & their relative positions are collected. Standard array is divided into nine squares or grids in grid formation. Each grid will have a set of lines & or singular points. The grids are analyzed and the number of lines & SPs are found. If characteristics are not present for a particular grid, variable is set as 0 for that particular grid. Thus its lines & SPs will characterize each grid.

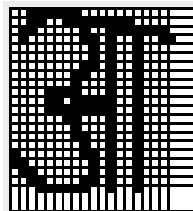


Fig. 5 Grid Infrastructure

Intelligent Character Recognition system using Grid infrastructure can be used to speed up the translation of image based documents into structured documents that are easy to process.

D. Image Editing

In this step, the document is first converted into a form such that it is editable. It allows user to add specific content to the documents, deletion of certain content from documents and any other modification. The desired form of the document may be MS-Word, Text, ... as specified by the user.

E. Image Searching

In this step, the edited and saved document can be searched. The user requests the system to search for a particular document. The main objective of this step is to facilitate the user to search the documents. The system finds the documents based on ICR methodology and returns the result of the search to the user.

V. DATA FLOW DIAGRAM

The dataflow is a graphical representation of the "flow" of data through an information system. The flow of data in our system can be described in the form of dataflow diagram as follows:-

Steps involved in data flow diagram are

1. Image processing.
2. Image Recognition
3. Image search
4. Image editing

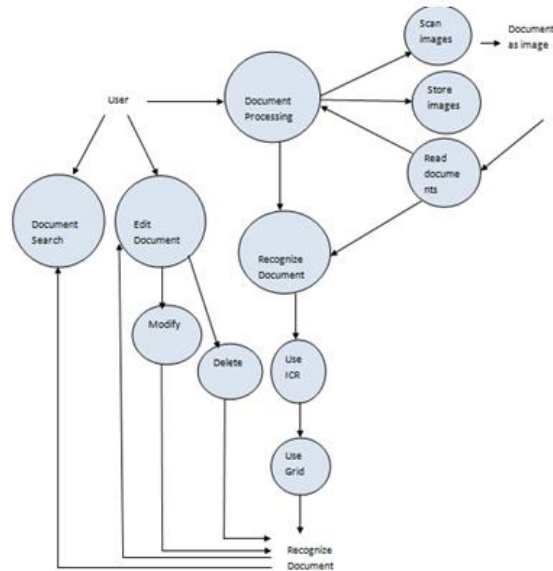


Fig. 6 Data Flow Diagram of the proposed system.

VI. CONCLUSION

To conclude, a robust algorithm has been described for extraction of text from scanned images, measuring and correcting the skew and shift values that are present in a scanned image. Subsequently, three techniques, viz. (i) pre-processing techniques, (ii) Recognition using Kohonen Neural Networks and (iii) Recognition using Feed Forward Neural Networks, that together comprise the handwritten character and scanned image extraction process have been presented. The impact and necessity of these methods on the overall performance of the ICR system has been illustrated by examples. It is observed from all the above explanations, that the character recognition is a very small part of a very vast field of Digital Signal Processing, it is considered to be a boon for many institutions for different purposes. The other applications like automated entry of data by OCR is one of the most attractive, labor reducing technology, pattern recognition of the scanned digital images from the satellite and comparing them with the previous images is also one of the application. The prediction about the climate can be made effectively by doing this. It can be implemented in many branches like banking systems, satellite communications where there is a necessity for extraction of handwritten and scanned images.

REFERENCES

- [1] Chirag I Patel, Ripal Patel, Palak Patel, "Handwritten Character Recognition Using Neural Networks", *International Journal of Scientific & Engineering Research Volume 2, Issue 5*, May-2011.
- [2] Alexander J. Faaborg, "Using Neural Networks to Create an Adaptive Character Recognition System", *Cornell University, Ithaca NY*, (May 14, 2002)
- [3] M. D. Garris, C. L. Wilson, and J. L. Blue. Neural networkbased systems for handprint ocr applications. *IEEE Transactions on Image Processing*, 7(8):1097–1110, August 1998.
- [4] P. J. Grother. Karhunenlo`eve feature extraction for neural handwritten character recognition. *Applications of Artificial Neural Networks III, SPIE, Orlando*, 1709:155–166, April 1992.
- [5] Z. Shi and V. Govindraj. Character image enhancement by selective region-growing. *Pattern Recognition Letters*, 17:523–527, 1996.
- [6] G. Nagy. Twenty years of document image analysis in pami. *IEEE Trans. Pattern Anal. And Mach. Intell.*, 22(1):38–62, January 2000.
- [7] Neves, L.A.P., J.M. De Carvalho, J. Facon and F. Bortolozzi, 2006. A new table extraction and recovery methodology with little use of previous knowledge. Proceedings of the International Workshop on Frontiers in Handwriting Recognition, (FHR'06), La Baule. France.
- [8] Chen, J.L. and H.J. Lee, 1996. Field data extraction for form document processing using a gravitationbased algorithm, *Pattern Recognit.*, 34: 1741-1750. DOI: 10.1016/S0031-3203(00)00115-1
- [9] Tseng, L.Y. and C.T. Chuang, 1992. An Efficient Knowledge-based stroke extraction method for multifont Chinese character. *Pattern Recognition*, 25: 1455-1458.
- [10] Liolios, N., N. Fakotakis and G. Kokkinakis, 2002. On the generalization of the form identification and skew detection problem. *Pattern Recognit.*, 35: 253-264. DOI: 10.1016/S0031-3203(01)00030-9
- [11] Pizano, A., 1992. Extracting line features from images of business forms and tables. Proceedings of the 11th IPAR International Conference on Pattern Recognition, Aug. 30- Sep. 3, IEEE Xplore Press, The Hague, Netherlands, pp: 399-403. DOI:10.1109/ICPR.1992.202008
- [12] Mandal, S., S.P. Chowdhury, A.K. Das and B. Chanda 2005. A hierarchical method for automated identification and segmentation of forms. Proceedings of 8th International Conference on Document Analysis and Recognition, Aug. 29-Sep. 1, IEEE Xplore Press, India, pp: 705-709. DOI: 10.1109/ICDAR.2005.17

- [13] Guillevic, D. and C.Y. Suen, 1993. Cursive script recognition: A fast reader scheme. Proceeding of 2ndInternational Conference on Document Analysis and Recognition, Oct. 20-22, IEEE Xplore Press, Tsukuba Science City, Japan, pp: 311-314. DOI: 10.1109/ICDAR.1993.395725
- [14] Mandal, S., S.P. Chowdhury, A.K. Das and B. Chanda, 2006. Fully automated identification and segmentation of form document form processing, springer, Comput. Graphics, 32: 953-961. DOI: 10.1007/1-4020-4179-9_139
- [15] Boatto, L., V. Consorti, M. De Buono, S. Di Zenzo and V. Eramoet *al.*, 1992a. An interpretation system for land register maps. Computer, 25: 25-33. DOI:10.1109/2.144437
- [16] Wang, D. and S.N. Srihari, 1994. Analysis of form images.Int. J. Pattern Recognit., 8: 1031-1031.
- [17] Casey, R.G. and D.R. Ferguson, 1990. Intelligent forms processing. IBM Syst. J., 29: 435-450. DOI: 10.1147/sj.293.0435
- [18] Casey, R., D. Ferguson, K. Mohiuddin and E. Walach, 1992. Intelligent forms processing system. Machine Vision Appl., 5: 143-155. DOI:10.1007/BF02626994.
- [19] M.F.Zafar, D.Mohamad and R.M.Othman, "Writer Independent Online Handwritten Character Recognition Using a Simple Approach", *Information Technology Journal* 5(3):476-484, 2006 ISSN 1812-5638, © 2006 Asian Network for Scientific Information.
- [20] Jackel, L.D., Graf, H.P., Boser, B., Denker, J.S., Le Cun, Y.,Howard, R.E., Matan, O., and Baird, H.S. (1990).Hardware considerations for Neural-net character Recognition Systems. In Simon, J., editor from pixels to Feature II, (Bonas, France, 1990).