



OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

HANDWRITTEN CHARACTER RECOGNITION USING NEURAL NETWORK

Pradip Kushal Dantrao(BECA75)¹, *Shweta Shahaji Shinde*(BECA77)², *Priti Milind Kamble*(BECA86)³ *Dr Kavita Joshi*⁴
G. H .Raisoni Collage of Engineering & Management Pune, India^{1,2,3,4}

Abstract: *Handwriting recognition is the ability of a machine to receive and interpret handwritten input from multiple sources like paper documents, photographs, touch screen devices etc. Recognition of handwritten and machine characters is an emerging area of research and finds extensive applications in banks, offices and industries. The main aim of this project is to design expert system for , “HCR using neural Network” that can effectively recognize a particular character of type format using the Artificial Neural Network approach.*

Keywords— *Component, Formatting, Style, Styling, Insert*

I INTRODUCTION

This project, ‘Handwritten Character Recognition’ is a software algorithm project to recognize any hand written character efficiently on computer with input is either an old optical image or currently provided through touch input, mouse or pen.

Character recognition, usually abbreviated to optical character recognition or shortened OCR, is the mechanical or electronic translation of images of handwritten, typewritten or printed text (usually captured by a scanner) into machine-editable text. It is a field of research in pattern recognition, artificial intelligence and machine vision. Though academic research in the field continues, the focus on character recognition has shifted to implementation of proven techniques. Optical character recognition is a scheme which enables a computer to learn, understand, improvise and interpret the written or printed character in their own language, but present correspondingly as specified by the user. Optical Character Recognition uses the image processing technique to identify any character computer/typewriter printed or hand written. A lot of work has been done in this field. But a continuous improvisation of OCR techniques is being done based on the fact that algorithm must have higher accuracy of recognition, higher persistency in number of times of correct prediction and increased execution time.

II.

OBJECTIVES

The objective of this project is to identify handwritten characters with the use of neural networks. We have to construct suitable neural network and train it properly. The program should be able to extract the characters one by one and map the target output for training purpose. After automatic processing of the image, the training dataset has to be used to train “classification engine” for recognition purpose.

III.

PROBLEM STATMENT

To arrange or manage the hard copies of a written document has been a type of deadly so far. Further more to keep those at a secure place even necessitate a lot of awareness. And to change them in softcopy requires a lot of time, money and people. So supervision the written hardcopies become more painful and too much difficult. The summit most fundamental source for every association is its credentials. It can in be in soft (computerized or digital) or in hard (printed on a paper) form. They all want to keep whole documents or information that is easy to manage anytime, anywhere. In this fast-moving world where the markets are getting larger and world is getting smaller, fast access to their information is must. With the swiftly shifting world and the extremely competitive and versatile nature of industry, all the information in any system of its network is important to the organization.

One of the features that should be under control is to save the

file in different formats. Optical Character Recognition is an application, which will be used to make a soft copy with the hardcopy.

IV. SYSTEM ARCHITECTURE

A. Algorithm :

To solve the defined handwritten character recognition problem of classification we used Pycharm computation software with Neural Network Toolbox and Image Processing Toolbox add-on. The computation code is divided into the next categories:

- Pre-processing of the images
- Feature Extraction
- Creating CNN Model
- Training and Testing of Network
- Recognition

A. Block Diagram:

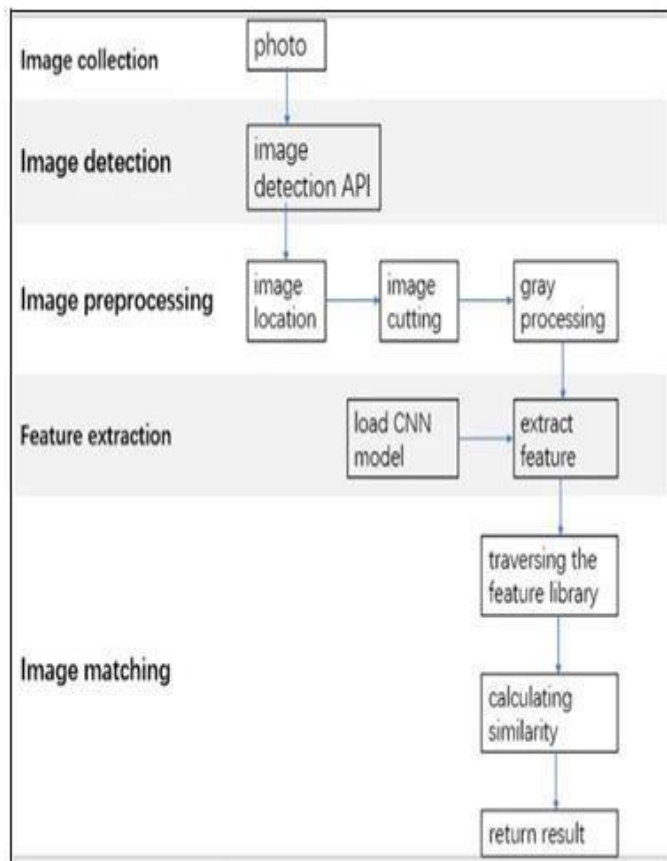


Fig 1. Block Diagram

Given a 32-by-32 pixel window, the detector decides whether the window contains a centered character. Similarly, the recognizer decides which of 62 characters (26 upper case, 26 lowercase letters, and 10 digits) is in the window. As described at length in Section 3, we scan a full scene image to identify candidate lines of text, on which we perform word-

level segmentation and recognition to obtain the end-to-end results.

B. Flowchart:

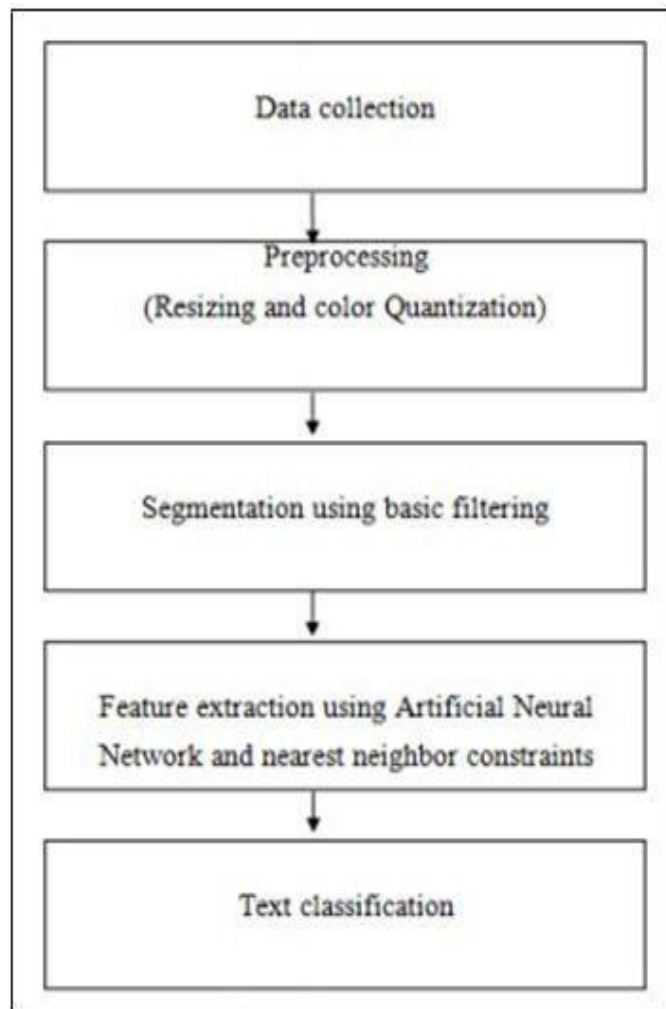


Fig 2. Flowchart

D. Pre-Processing :

We proposed a system here for recognition of characters from images with the help of Artificial Neural Network and nearest neighbour approach. The main parts of the consideration are the data collection, pre processing step, then the extraction of symbol of interest from the input scanned image, after that conversion of this identified or located symbol into a pixel matrix of specific or standard size. Then the next part is the classification of the symbol under consideration which can be done by assigning the different loads to the layer of neural network. In case of data collection, the hidden layer neurons have to assign weights for the classification of characters according to the specific font style. Nearest Neighbour Approach: It is used for classification purposes. Here, for every component, two consecutive nearest neighbours are found by the application of Euclidean distance. The different

parameters like dimensions, distance and the alignment of component and their neighbours are compared.

Algorithm for classification of characters :

Form network according to the specified topology parameters

1. Fix the number of layers of the network and also fix the number of neurons in each layer.
2. Assign weights to the different neurons. Also set the biased value for each output neuron.
3. Initialize the hidden layer by providing desirable data from training set file.
4. Input the image and identify and place different character symbols in a liner array.
5. Formation of output is done at output layer with the help of Unicodes.
6. For every character symbol feed forward output is collected for further use.
7. Now the Error computation can be performed using error function.
8. Adjust weights according to the error using the concept of back propagation.
9. Repeat steps 6 to 8 for other characters too.
10. Averaging of error can be computed for all characters.
11. Up to fix number of epoch repeat step 6 to step 10.

E. Recognition:

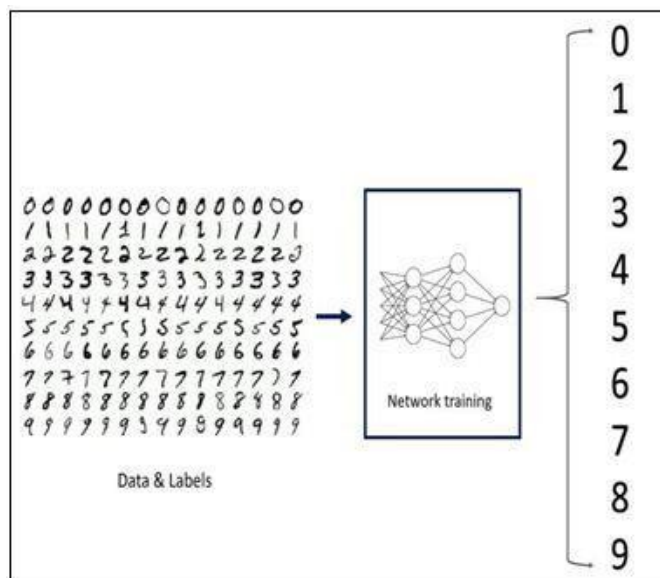


Fig. Data Recognition

V CONCLUSION

Classification of characters and learning of image processing techniques is done in this project. Also the scheme through which project is achieved is Artificial Neural Network scheme. The result which was got was correct up to more

than 90% of the cases, but it would be improved at the end. This work was basically focused on envisaging methods that can efficiently extract feature vectors from each individual character. The method I came up with gave efficient and effective result both for feature extraction as well as recognition. There are also different methods through which ‘handwritten character recognition’ is achieved.

REFERENCES

1. G.Vamvakas, B.Gatos, N. Stamatopoulos, and S.J.Perantonis, “A Complete Optical Character Recognition Methodology for Historical Documents”, the Eighth IAPR Workshop on Document Analysis Systems.
2. Debasish Basa ,Sukadev Meher , “Handwritten Odia Character Recognition”, Presented in the National Conference on Recent in Microwave Tubes,Jagannath Gupta Institute of Engineering and Technology, Jaipur, March 4-5,2011
3. Rajean Plamondon, Fellow IEEE and Sargur N. Srihari, Fellow IEEE , “On-Line and Off-Line Handwriting character Recognition: A Comprehensive Survey”, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. VOL. 22, NO. 1. JANUARY 2000
4. Sanghamitra Mohanty, Hemanta Kumar Behera , “A Complete OCR Development System For Oriya Script”
5. B.B. Chaudhuri U. Pal M. Mitra, “Automatic Recognition of Printed Oriya Script”
6. Fumitaka KIMURA, Tetsushi WAKABAYASHI and Yasuji MIYAKE, “On Feature Extraction for Limited Class Problem”
7. Supriya Deshmukh, Leena Ragma, “Analysis of Directional Features - Stroke and Contour for Handwritten Character Recognition”, 2009 IEEE International Advance Computing Conference (IACC 2009) Patiala, India, 6-7 March 2009
8. M. Blumenstein, B. K. Verma and H. Basli, “A Novel Feature Extraction Technique for the Recognition of Segmented Handwritten”
- 9.Tapan K Bhowmik Swapan K Parui Ujjwal Bhattacharya Bikash Shaw, “ An HMM Based Recognition Scheme for Handwritten Oriya Numerals”.
- 10.Sung-Bae Cho, “Neural-Network Classifiers for Recognizing Totally Unconstrained Handwritten Numerals”,IEEE Transactions on Neural Networks, Vol. 8, No. 1, January 1997