# Neural Network to Recognize Handwritten Character and Text Detection

**[1]Avishek Chakraborty, [2]M. Nalini**

[1]UG Scholar, [2]Assistant Professor, Department of Computer Science and Engineering
Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences
Chennai, India
[1]avishekchakraborty914@gmail.com, [2]nalini.tptwin@gmail.com

**Abstract**

The paper will define the quality technique to get extra than 91% accurateness in the arena of Neural network-based Handwritten Character and textual content reputation. There has been sufficient study done in the arena of HCR however silent it's miles an open complex as we are nonetheless out of place in accomplishment the finest accurateness. In this research, offline HCR may be attained the usage of neural community. An approach named Max Regression is used for conveying the chances to HC is one of the numerous fonts because it stretches the morals among zero and 1 summing up to at least one. The decision is to change the software program with a completely excessive exactness charge, nominal period, area problem and also finest.

*Keywords: Recognition of character and neural network.*

## 1. Introduction

The written characters or written papers is uncomplicated to kinsfolk as to study. The identical capability may be brought on to the Machines to boot with the help of the usage of Machine Learning and computer science. The sphere that offers with this problem is termed because the OCR or to boot brought up as Optical Character Recognition. It is the space or have a glance at amongst varied fields together with recognizing of pattern, conjointly Image ingenious and discerning and to boot AI. It is often used to make it easier to fluctuate electronic and film text into virtual person to browse through the technologies. The period employed in getting into the facts and conjointly the cupboard space needed by the papers is also notably condensed via usage of OCR or in several words, it may be retrieved apace. By utilization of OCR in finance subject, crime eventualities, and then forth. several important and sensitive files is also processed quickly while not human intervention.OCR earlier is also inferred in approaches supported kind of the matter content and report gaining. Moreover if we predict concerning the matter content kind, then these are more kinds, that is smart acknowledgement of the written text and Printed Character.we would like the good quality capability thanks to the assorted script of the folks. several instances the style of identical man or lady

is distinct at instances. More OCR is characterised into 2

work as Offline and on-line quality structures based on getting the papers. Offline system offers with recognizing the pre-written. Analysis within the region of word name, being earliest endeavor to understand the written man or lady. This mid-sixty studies exhibited utilization of the communicating with the help of the mixture technique being planned. He has established that each single written individual is restrained to a range of schematic highlights. This speculation was later applied at intervals to the locality of character name as part of the majority techniques for auxiliary methodologies.
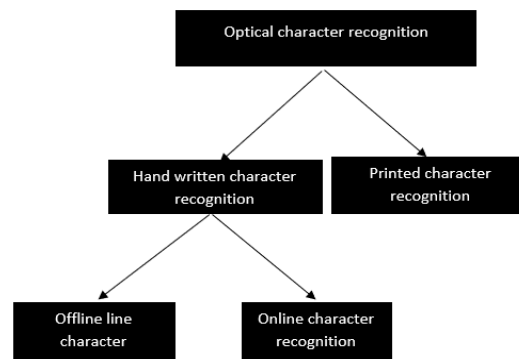


Figure 1: Character Recognition Hierarchy

## 2. Literature Review

K. Gaurav and Bhatia P. K, planned specific pre-

managing structures which were linked to character recognition. The methodology took an attempt from a primary photograph-based file to a hued and altered forces inclusive of foundation on the different kinds of photos. Totally different pre-handling and standardization methods, such as skew solution, distinguish evacuation, commotion removal, and various ways of updating were assisted. They reached the choice that a solitary manner can been supported. Those gentlemen had come to the point that a single or more precisely particular manner cannot be associated with preprocessing the photograph. When they came to that choice, to boot, there were some disparities utilising each kind of system likewise cannot deliver the standard truth comes concerning.

Salvador España-Boqueria, The analysts planned the employment of hybrid or 0.5 and 1/2 hid Markov show (HMM) to understand written content material in disconnected mode. The basic aspect of the optical model changed into organized by Markoff chain technique. Likewise, multilayer perceptron was used to gauge probabilities.

A modified quadratic classifier is applied to know the disconnected written numerals of six outstanding Local Indian languages. An identical paper in the same way copes up to understand English letters so as. For each of those, a multilayer perceptron became used. The following boundary and Fourier descriptors are used for the extraction of facets. The characters were known through inspecting the form and observing their highlights. Also, the number of hid layers was selected, once again the unfold device was applied.This terribly calculation accounted for a nomenclature of xciv available, with less preparatory time.R. Bajaj, S. Chaudhari, L. Dey, et al [5], had applied distinctive highlights such as clear half, thickness and minute highlights to group syllabary numerals. To boot, to increase popularity, the paper recommends unwavering multi-classifier fine for syllabary numerals written.

During this paper, SandhyaArora mainly presented four highlights of a kind of shadow, bar graph of the chain code crossing problem and horizontal line fitting highlights. Among those highlights the shadow was registered for the photographer all around, the remaining three were processed by partitioning the individual image into the one-of - a-kind parts.Within the one helpful execution mistreatment the 4900 instance dataset incontestible the ninety reality charge for syllabary characters written.

## 3. Proposed Recognition System

Normally HCR is categorised into six phases which might be the acquisition of images, pre-processing of getting into a photograph, segmentation, feature extraction, category and place up process. Within the first part, pre-processing offers the ways very important to construct the raw enter snapshots. The segmentation part deals with segmenting the raw pic into mateless characters so composition it into m*n pixels to traumatize it mathematically. The operate extraction approach is

probably of principal huge bother in achieving high recognition % that various techniques like Deformable templates, Gradient characteristic, Contour profiles, Fourier descriptors, Gabor functions, Graph description, Geometric moment invariants, templet matching, Unitary Image transforms, Projection Histograms, Zoning, Zernike Moments and Spline curve approximation ar used.

## Proposed Recognition System

### A. Image Acquisition

In the degree of image acquisition the enter image provides the device with popularity. Either the data may be in a picture presentation that has JPEG, BMT and then on. Or scanned image, virtual dig cam or the other applicable digital enter device otherwise you will draw at the canvas well-found at the computer programme.

### B. Pre-processing

The ordinal method referred to as Pre-processing is that access approach for character acceptance and very important when making the acknowledgement charge for a decision. Preprocessing works to normalize strokes and deduct variations to boot that can scale back the accuracy value.Preprocessing works especially at the various irregularities, like the size of the irregular content of matter, points ignored throughout the pen motion, jitters, left - right bend and uneven areas.
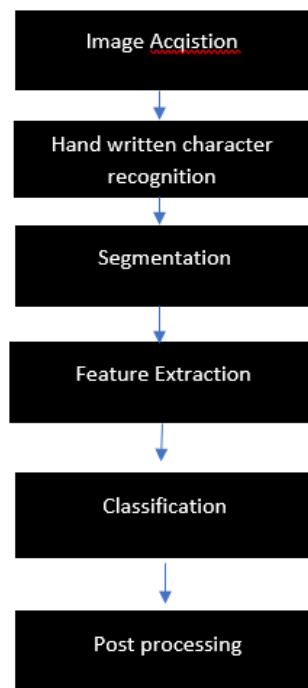


Figure 2: Proposed System Architecture

1.Noise Removing: to get rid of the undesirable or unsought designs, how referred to as noise doing away with is employed. There's a way like Uniform and non-uniform filtering this is often to be used.

2.Binarization: during this, all written characters ar

translated into grey-scale images. Each image of a person or lady is to be caught vertically when translating the grayscale pic into the binary matrix.

3.Normalization: it's the way of translating pic knowledge into identical recent needed form. principally filler in conjunction with skewing normalizations ar accomplished. Size modifies the pic photograph into the predefined constant size. Whereas skew is used at some stage during the scanning process, once the text is deviated from the lowest line, and for this skewing and detection, and again, results of propagation ar required.

### C. Segmentation

Segmentation is employed to remodel input image consisting {of many|of the persony} characters into the man or lady characters. The ways used ar phrase, line and individual segmentation. It's generally distributed by means of dividing single characters from the phrase image. Moreover, content is processed during a manner this is often tree-like. Within the preliminary state of affairs, a row bar graph is employed to phase the traces. Then after, each stage, characters ar retrieved by a way referred to as a bar graph so eventually obtaining it retrieved.

### D. Feature Extraction

Characteristic extraction intends to allow the extraction of a sample that is most essential for the category. a number of the Feature extraction ways like Principle element Analysis (PCA), Scale Invariant Feature Extraction (SIFT), Linear Discriminant Analysis (LDA), Histogram, Chain Code (CC), partitioning and Gradient based capabilities is also enforced to extract the capabilities of person characters. All of these capabilities arwont to teach the given convenience.

Each of the metameric photograph is taken of a number of pixels of measuring twenty eight * twenty eight. This could be performed to be a large array of numbers. There through demolition the array right into a vector of 28*28 = 784 numbers. Thus, the image currently converges to a minimal bunch of arrays {in a|duringa|in Associate in Nursing exceedingly|in a very} 784-mobile dimension of an implausibly inexperienced structure. The image currently turns into a tensor of n-dimensional array.

Denoted by summing over the reserved pixels at intervals the enter image x. We have a tendency to translate the proof of the tallies into the evaluated chances 'y' the usage of the "softmax" that is Associate in Nursing activation or link operate,

y = softmax(proof)

changing the chances of input

Softmax(x) = normalize(exp(x))

### E. Post-Processing

The last and last phase of the character name is that the Post-processing. It's the method for correcting the misclassified output by mistreatment language. It

approaches output by mistreatment obtaining it when the shape was known. If the shape is diagnosed then the accuracy may be advanced following the data of language. Form recognizers behave in in our own way for various handwriting enter. For the few, its effects during a personal character of string while conjointly that embody some vary of alternates within the second case, via consisting of the live of self-belief in every chance.

### 4. Result



Figure 3a: Upload Image



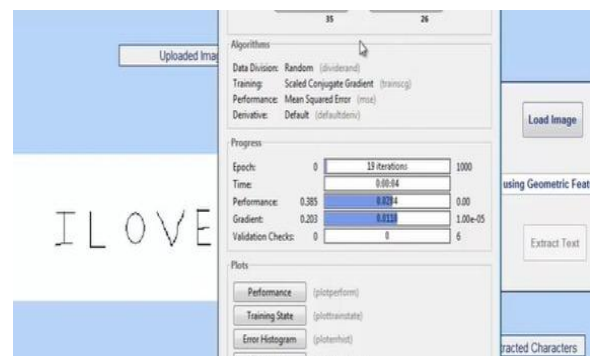Figure 3b: Image Uploaded Successfully



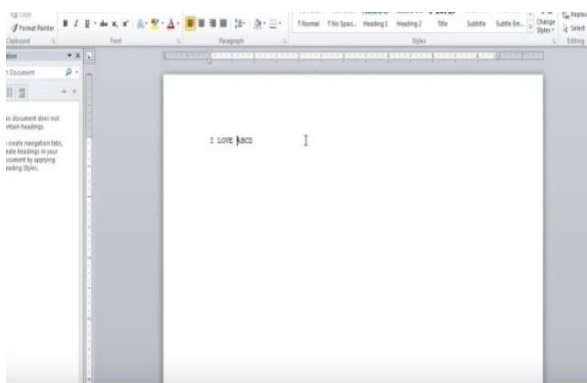Figure 3c: Train the Image to Select Relevant Algorithm

Figure 4: Image Successfully Converted into Text

## 5. Conclusion

It changed into decided that the function extraction approach like oblique and direction strategies are way better in producing high accuracy results in comparison to most of the out-of-date vertical and horizontal strategies. Also using a Neural community with pleasant tried layers offers the plus feature of having a better tolerance to noise hence giving accurate consequences. In the neural community, the version referred to as feed ahead is specially trained using the returned-propagation algorithm to classify and apprehend the characters as well as get skilled increasingly more. Apart from those, the use of normalization alongside function extraction yielded better and higher accuracy consequences in individual popularity. It is likewise discovered that the larger our education statistics set and better neural network layout, the better accuracy is the result.

## References

[1] Singh, Sameer, Mark Hewitt,"Cursive Digit And Character Recognition on Cedar Database", Pattern Recognition, 2000.Proceedings. 15th international conference on. Vol. 2. IEEE 2000.

[2] K. Gaurav and Bhatia P. K., "Analytical Review of Preprocessing Techniques for Offline Handwritten Character Recognition", 2nd International Conference on Emerging Trends in Engineering & Management, ICETEM, 2013.

[3] Salvador España-Boquera, Maria J. C. B., Jorge G. M. and Francisco Z. M., "Improving Offline Handwritten Text Recognition with Hybrid HMM/ANN Models", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, No. 4, April 2011.

[4] U. Pal, T. Wakabayashi and F. Kimura, "Handwritten numeral recognition of six popular scripts," Ninth International conference on Document Analysis and Recognition ICDAR 07, Vol.2, pp.749-753, 2007.

[5] Reena Bajaj, LipikaDey, and S. Chaudhury, "Devnagari numeral recognition by combining decision of multiple connectionist classifiers", Sadhana, Vol.27, part. 1, pp.-59-72, 2002.

[6] SandhyaArora, "Combining Multiple Feature Extraction Techniques for Handwritten Devnagari Character Recognition", IEEE Region 10 Colloquium and the Third ICIIS, Kharagpur, INDIA, December 2008.

[7] Nafiz Arica, and Fatos T. Yarman-Vural, ―Optical Character Recognition for Cursive Handwriting,‖ IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, no.6, pp. 801-113, June 2002.

[8] K. H. Aparna,, Vidhya Subramanian, M. Kasirajan, G. Vijay Prakash, V. S. Chakravarthy, SriganeshMadhvanath, "Online Handwriting Recognition for Tamil", IWFHR, 2004, Proceedings. Ninth International Workshop on Frontiers in Handwriting Recognition, Proceedings. Ninth International Workshop on Frontiers in Handwriting Recognition 2004, pp. 438-443, doi:10.1109/IWFHR.2004

[9] Surya Nath R S M, Afseena "Handwritten Recognition- A review", International Journal of Scientific and Research Publications, Volume 5, Issue 3, March 2015 1 ISSN 2250- 3153

[10] Anita Pal and DavashankarSingh,"Handwritten English Character Recognition Using Neural Network", International Journal of Computer Science and Communication, pp: 141- 144, 2011.

[11] J.Pradeep1 , E.Srinivasan2 and S.Himavathi3, "DIAGONAL BASED FEATURE EXTRACTION FOR HANDWRITTEN ALPHABETS RECOGNITION SYSTEM USING NEURAL NETWORK" International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 1, Feb 2011

[12] TheingiHtike and YadanaThein "Handwritten Character Recognition Using Competitive Neural Trees " IACSIT International Journal of Engineering and Technology, Vol. 5, No. 3, June 2013

[13] AyushPurohit #1, Shardul Singh Chauhan#2 "A Literature Survey on Handwritten Character Recognition" AyushPurohit et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (1) , 2016, 1-5

[14] ManojSonkusare and NarendraSahu "A SURVEY ON HANDWRITTEN CHARACTER RECOGNITION (HCR) TECHNIQUES FOR ENGLISH ALPHABETS" Advances in Vision Computing: An International Journal (AVC) Vol.3, No.1, March 2016

[15] S.V. Manikanthan, T.Padmapriya, "A Secured Multi-Level Key Management Technique for Intensified Wireless Sensor Network", International Journal of Recent Technology and Engineering, Vol. 7, issue6s2, 2019. https://www.ijrte.org/wp-content/uploads/papers/v7i6s2/F10720476S219.pdf

[16] Hallale, Sumedha B., and Geeta D. Salunke. "Twelve Directional Feature Extraction for Handwritten English Character Recognition." International Journal of Recent Technology and Engineering 2, no. 2 (2013)