

# Hand Written English Character Recognition Using Column\_Wise Segmentation Of Image Matrix (CSIM)

Alfiya L. Sheikh<sup>1</sup>, Abhilasha R. Mahure<sup>2</sup>, Mayuri D. Kokate<sup>3</sup>, Prof.P.M.Bihade<sup>4</sup>

1(IT, SRPCE, Nagpur [Alfiyasheikh18@gmail.com](mailto:Alfiyasheikh18@gmail.com))

2(IT, SRPCE, Nagpur [abhi.mahure@gmail.com](mailto:abhi.mahure@gmail.com))

3(IT, SRPCE, Nagpur [Mayurikokate567@gmail.com](mailto:Mayurikokate567@gmail.com))

4(Asst.prof, IT, S.R.P.C.E., Nagpur, [pmBihade29@gmail.com](mailto:pmBihade29@gmail.com))

## Abstract:

Research is going ahead to create both equipment and programming to perceive written by hand characters effectively and precisely. Fake Neural Network (ANN) is an extremely proficient strategy for perceiving transcribed character Endeavors have just been made to perceive English letter sets utilizing comparative kind of strategies. New technique has been attempted, in this paper, to enhance the execution of the beforehand connected strategies. Info picture lattice is compacted into a lower measurement grid with a specific end goal to diminish non presence of a specific character.

**Keywords:** ANN, CSIM, Compression, Perceptron, Segmentation, Learning Rule

## I.INTRODUCTION

Master frameworks can be produced utilizing Artificial Neural Network (ANN) which can perceive hand composed English letters in order effectively and precisely. The penmanship styles of various people change interminably which makes the advancement of master frameworks to perceive written by hand characters exceptionally troublesome.

Regular highlights removed out from distinctive penmanship styles have turned out to be an extremely effective strategy. Summed up issues can be illuminated utilizing Multiscale Training Technique

(MST), [1, 2, 3]. One such technique has been connected in Devnagri Content where some component extraction procedures like crossing point, shadow highlights, chain code histogram what's more, straight line fitting were utilized, [4]. One component extraction technique is discovering twelve directional component inputs which rely on the angles. The highlights of the characters are bearings of the pixels w.r.t. neighbouring pixels,

[5]. Written by hand English character acknowledgment utilizing Line astute Segmentation Technique (RST) is a way to deal with discover normal highlights of same characters written in various hand

composing styles by dividing the information design grid into discrete columns and endeavoring to discover normal lines among distinctive hand composing styles, [6]. With a specific end goal to discover normal highlights among the characters composed by various people a few strategies are added to the effectively created ANNs like perceptron learning strategy, [7, and 8]. In this paper an endeavor has been made through the packed Column-wise Segmentation of Image Framework (CSIM) to perceive manually written characters utilizing Neural Network.

The general program is partitioned into three sections, pressure of the picture lattice, fragmenting the packed framework section shrewd and preparing the net lastly testing the net by giving characters taken from various people

## II. CLASSIFICATION

Grouping in penmanship acknowledgment alludes to one of the accompanying procedures:

- (1) grouping of characters;
- (2) Grouping of words; and
- (3) grouping of highlights

Various order methods has been produced and examined for the. Arrangement of characters, words, and highlights. The order systems have utilized different factual and keen classifiers, including k-NN, SVMs, HMMs, and neural systems. For the grouping of numerals/characters, a plentiful number of systems have been investigated in the writing. Numerous factual methods have been utilized for characterization, for example, k-Nearest Neighbour. Nonetheless, some measurable techniques have been observed to be illogical in genuine applications, as they

require that all preparation tests be put away and looked at for the grouping procedure (Liu and Fujisawa, 2005). As of late, probably the most prominent, intense, and effective strategies have utilized neural system classifiers (Cho, 1997; Verma et al., 2004) and HMM-based methods (Arica 1999), acquiring acknowledgment rates over 99% for disconnected written by hand, confined numerals. As of late, bolster vector machines have been utilized for numeral/character grouping, likewise getting great outcomes over 99% (Liu and Fujisawa, 2005). It has likewise been discovered that the utilization of multistage and consolidated classifiers has been exceptionally fruitful for numeral/character grouping.

## III. PROPOSED SEGMENTATION APPROACH

Division is the most critical piece of the cursive manually written acknowledgment issue, when an expository approach is utilized for the acknowledgment. Little adjustments on the division calculations result in genuine changes in the acknowledgment rates. The division strategy proposed in this investigation, is roused from crafted by Lee et al., where the character division issue is characterized as the issue of finding the briefest way in a chart characterized over a division locale, limiting the gathered power. In, the division areas are recognized from the pinnacles of the even projection profile in the dark level picture expecting machine printed characters. Then again, our point is to build up a framework for the manually written characters, which might be both inclined and skewed. Therefore, the proposed technique plays out the division by

consolidating the attributes of dim scale and parallel pictures. In the first place, the division districts are resolved in the parallel word picture by utilizing the form of the written work. At that point, an enhanced hunt process is connected to the division districts on the dark scale word picture for deciding the division limit. In this area division calculation and pre-processing steps are exhibited. Counterfeit neural system is prepared physically for the right and erroneous division focuses got from the proposed division strategy. MATLAB 7.0 is utilized for all trials performed on arrangement of 1.6 GHz processor and 1 GB DDR RAM. 2.1 Pre-processing and proposed division calculation The first dim scaled picture is pairs utilizing Otsu calculation by choosing naturally a limit an incentive for a given picture If required, after linearization, incline remedy is performed. At last, picture is changed over to skeleton organization to permit clients verity of composing gadget, pen tilt and to stifle additional information.

**The proposed segmentation algorithm is:**

Division is the most critical piece of the cursive manually written acknowledgment issue, when an expository approach is utilized for the acknowledgment. Little adjustments on the division calculations result in genuine changes in the acknowledgment rates. The division strategy proposed in this investigation, is roused from crafted by Lee et al., where the character division issue is characterized as the issue of finding the briefest way in a chart characterized over a division locale, limiting the gathered power. In, the division areas are recognized from the pinnacles of the even

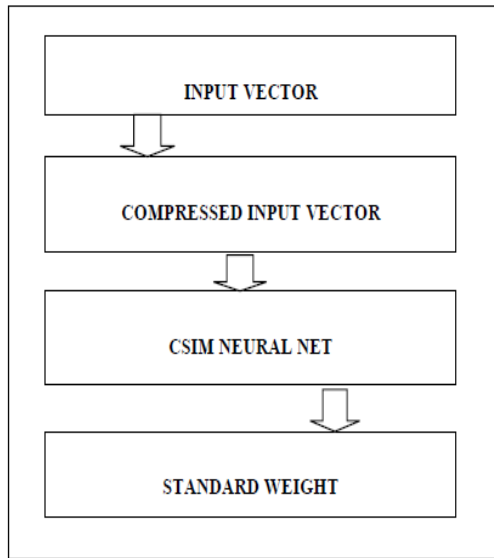
projection profile in the dark level picture expecting machine printed characters. Then again, our point is to build up a framework for the manually written characters, which might be both inclined and skewed. Therefore, the proposed technique plays out the division by consolidating the attributes of dim scale and parallel pictures. In the first place, the division districts are resolved in the parallel word picture by utilizing the form of the written work. At that point, an enhanced hunt process is connected to the division districts on the dark scale word picture for deciding the division limit. In this area division calculation and pre-processing steps are exhibited. Counterfeit neural system is prepared physically for the right and erroneous division focuses got from the proposed division strategy. MATLAB 7.0 is utilized for all trials performed on arrangement of 1.6 GHz processor and 1 GB DDR RAM. 2.1 Pre-processing and proposed division calculation The first dim scaled picture is pairs utilizing Otsu calculation by choosing naturally a limit an incentive for a given picture If required, after linearization, incline remedy is performed. At last, picture is changed over to skeleton organization to permit clients verity of composing gadget, pen tilt and to stifle additional information.

**EXISTING SYSTEM:**

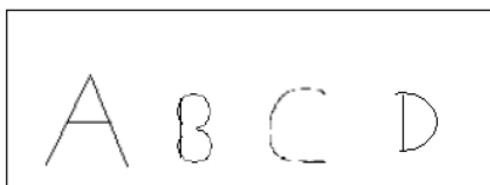
- In Current Scenario, Handwritten Character Detection is a riotous work.
- to perceive the verifiable archives and notes, current approach is very not evaluative and impressive.
- Most of the Current Recognition Techniques does not take a shot at Raw Images and Line Extraction.

□ Segmentation is the most critical piece of the cursive manually written acknowledgment issue, which is utilized for the acknowledgment

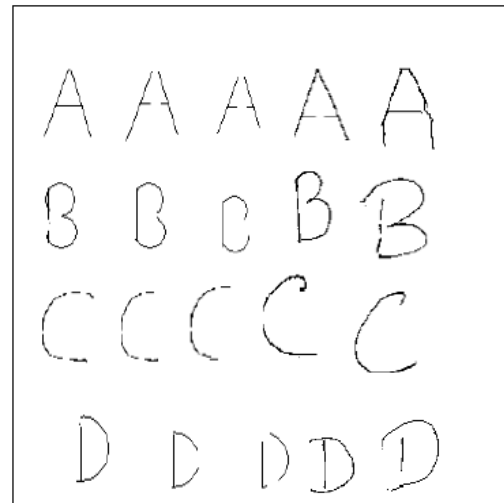
#### IV. System Architecture:



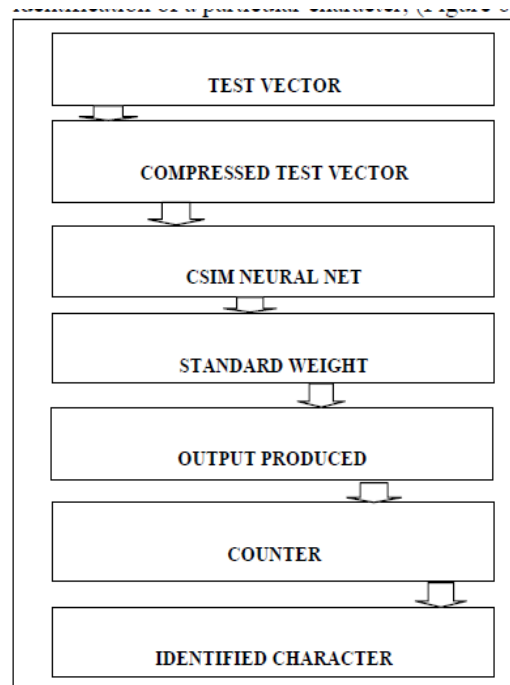
**Fig.1. Conceptual Diagram Of CSIM Training**



**Fig. 2 ALPHABET COSIDERED FOR CSIM TRANING**



**Fig.3.Alphabet Considered For CSIM Testing**



**Fig. 4. Conceptual Diagram of CSIM Training**

## V. PROPOSED SYSTEM:

- Here we are proposing the new System for Handwritten Character Detection.
- In our Project; we are using Histogram Technique for Scanned Hand written Images.
- This project works in many Levels.
- For this we need to create the database which includes the set of different styles of handwritten characters i.e. alphabets.
- Handwritten character detection, partition and Recognition implemented by using the Histogram technique.
- This Technique is User Friendly to Use and Effective in Manner

## VI. MODULES OF PROJECT:

- Reversal of Image Mode
- Creating A Data- Base
- Line Extraction and Separation
- Word Extraction and Separation
- Character Extraction
- Word Detection and Output

### MODULES DESCRIPTION:

#### 1) REVERSAL OF IMAGE MODE

In this take Scanned Image of the Document. By using the Gray Scale Technique; we will reverse the Input Format of Image.

For Ex: if it is a Colourful Image, then we will convert it into Black and White Image. Then we will reverse the Mode of Image i.e. if Background White and Fore Colour

Black then after reversing;

Background = Black and Fore Colour = White

Now this is our Proper Input of Image for Further Line and Word Extraction.

#### 2) CREATING DATABASE

In this Module; we will create a set of Data base. Data base of hand written characters i.e. alphabets a-z in various different Styles and manners like the handwriting of different persons. Every Alphabets which is in database will having an ASCII value. For the development of unconstrained large vocabulary handwriting recognition systems, another limitation is the amount of data required for training and testing. Statistical methods provide an automatic procedure to 'learn' the regularities in the handwriting data directly. The need of a large set of good training data is thus more critical than ever. Furthermore, with the tendency of incorporating contextual information in the recognition, more data are needed for the training and validation of systems.

3) LINE EXTRACTION AND SEPARATION  
For This Module; we will be using histogram Graph Technique. This technique will be applied as a whole of the Image. This will form Peak valley type Graph of each and every line of Scanned Image. The peak Valleys will indicate number of Lines in Scanned Image.

#### 4) WORD EXTRACTION AND SEPARATION

After Line Extraction, we will perform the Word Extraction. Here also we will use histogram Graph Technique. It will be having 1's and 0's values.

The peak Valleys will indicate number of Words in Scanned Image.

#### 5) CHARACTER EXTRACTION

In this Module; we will perform the Character Detection. After Word Extraction; we will perform the Character Extraction. By comparing the Characters of the Word with the own constructed Hand written Data base.

It will give us the recognized Extracted Character of the Input Word.

#### 6) WORD DETECTION AND OUTPUT

After performing Character Extraction and Detection; we will perform this in this module; we will get the corrected proper format or our output as a complete Word. Proper Format Input will give us Proper Out

#### VII. APPLICATIONS OF PROJECT

- This will be applicable where ever there is a need of any kind of Hand written or Cursive character Detection.
- It will be used In Doctors Prescriptions or Scientists hand written Thesis, Journals and Research Papers.
- It will be very much useful and convenient for Historical Research and Files...

#### VIII. CONCLUSION

The issue of the cursive penmanship is made complex by the way that the composition is innately uncertain letters in a word are for the most part connected together, ineffectively composed, and may even be missing. As a result, cursive content acknowledgment requires modern strategies, which utilizes a lot of shape data and which adjusts for the equivocalness by the utilization of relevant data. The acknowledgment techniques vigorously rely upon the idea of the information to be perceived. Little alterations and touchy choice of the parameters in the research centre condition give high acknowledgment rates to a given informational collection. Be that as it may, an effective framework detailed in the writing may flop, in actuality, issues

because of unpredicted change in information and infringement of the underlying suppositions, in actuality, condition. In this investigation, we endeavor to amplify the measure of data to be held in the information to a specific degree. The proposed technique stays away from the greater part of the pre preparing tasks, which causes loss of critical data. Notwithstanding, the data, extricated in the worldwide parameter estimation arrange, is utilized over the first picture as required. We utilized the inclination and skew point data, however we didn't make incline or skew edge rectification. We utilized the paired data, however we didn't depend just on the parallel data. One of the significant commitments of this investigation is the improvement of an effective division calculation. Usage of the character limits, nearby maxima and minima, incline point, upper and lower baselines, stroke stature and width enhances the pursuit calculation of the ideal division way, connected on a dark scale picture. This approach

#### REFERENCES

- [1] Survey of the State of the Art in Human Language Technology.  
<http://cslu.cse.ogi.edu/HLTsurvey/HLTsurvey.html>, November 1995.
- [2] T. M. Cover and P. E. Hart. Closest neighbour design classification. IEEE Transactions on Information Theory, 13(1):21–27, Jan. 1967.
- [3] C. Frankish, R. Body, and P. Morgan. Acknowledgment precision and client acknowledgment of pen interfaces. In Proceedings ACM CHI'95 Conference on Human Factors in Computing Systems, 1995.
- [4] T. Kohen. Self-Organizing Maps, volume 30 of Springer Series in Information Sciences. Springer-Verilog, 1997. Sec-on Extended Edition.
- [5] J. Laaksonen, J. Hurry, E. Ojai, and J. Kanga. Correlation of versatile techniques for on-line character acknowledgment. In Proceedings of International Conference on Artificial Neural Networks, pages 245–250, 1998.
- [6] M. J. Paloma. Client acknowledgment of written by hand acknowledgment precision. In Proceeding of ACM CHI94 Human Factors in Computing Systems Conference, page 107, April 1994.

