

A BRIEF INTRODUCTION TO HANDWRITTEN CHARACTER RECOGNITION

Priyank Patel *1, Roshan Shinde *2, Siddhesh Raut*3, Sheetal Mahadik*4

*1,2,3Student, Department of Electronics Engineering, Shree L R Tiwari College of Engineering, Mumbai, Maharashtra, India.

*4Asst. Professor, Department of Electronics Engineering, Shree L R Tiwari College of Engineering, Mumbai, Maharashtra, India

ABSTRACT

There are numerous things we people share for all intents and purpose. Yet, there are different things that are exceptional to each person - DNA, fingerprints, and so forth Handwriting is one other such thing that is extraordinary to each person, which the ongoing examinations on Handwriting investigation have just demonstrated. This paper presents an understanding into the condition of craftsmanship in Handwriting Recognition frameworks and portrays the development and progress in the field. The quickly developing computational force prepares the execution of the current HCR approaches and furthermore drives an expanding interest on various arising application areas, which require progressed systems. This material aides as a guide lining archive and update for the readers, working in the Character Recognition zone. An outline of HCR frameworks and their development over the long haul is introduced trailed by the accessible HCR procedures with their superiorities and impediments are investigated.

Keywords: Handwritten Character Recognition, Optical Character Recognition, Neural Network, Convolutional, Segmentation, Recognition.

I. INTRODUCTION

The word pattern recognition covers a wide variety of functional problems in the processing of information, from speech recognition and handwriting character classification, to machinery fault identification and medical diagnosis. Recognition of patterns-the act of taking raw data and taking action based on the "category" of the pattern has been critical for our survival, and we have developed highly sophisticated neural and cognitive processes over the past decades of millions of years.

Template matching, mathematical grouping, structural or syntactic matching and neural networks are the four best recognized methods to pattern recognition. Model matching is one of the fastest and earliest approaches to pattern recognition, where a 2-dimensional form or a version of the pattern to be identified is usually available. Although taking into account all permissible pose and scale changes [1], the pattern to be recognized is compared against the stored prototype.

II. LITERATURE SURVEY

An early notable attempt in the area of character recognition research is by Grimsdale in 1959. The root of much research work in the early sixties was based on a technique proposed by Eden in 1968, known as the method of analysis-by-synthesis. The great significance of the work of Eden was that he formally proved that a finite number of schematic features are created by all handwritten characters, an argument that was loosely used in previous works. This notion was later used in syntactic (structural) character recognition approaches in all methods [6].

Writing, which over the centuries has been the most natural means of gathering, storing and distributing information, now serves not only for communication between humans, but also for communication between humans and machines. The intense research effort in the field of CR was not only due to its difficulty in human reading simulation, but also because it offers successful applications such as the automated processing of the bulk number of papers, transferring data into machines and web interface to paper documents. Historically, HCR systems have evolved in three ages: 1900-1980 Early ages-- The history of character recognition can be traced as early as 1900, when the Russian Scientist Trying attempted to develop an aid for visually handicapped. With the advent of modern computers, the first character recognizers emerged in the mid-1940s [5]. Early work on the automated identification of characters was based either on machine printed text or on a limited collection of handwritten text or symbols that were well differentiated. In this time, machine-printed

HCR systems typically used prototype matching in which an image is compared to an image library. Low-level image recognition approaches have been used with handwritten text.

III. IMPLEMENTATION

A. Image Acquisition

Digitized/Computerized Picture is at first taken as info. The most well-known of these gadgets is the electronic tablet or digitizer. These gadgets utilize a pen that is computerized in nature. Information pictures for transcribed characters can likewise be taken by utilizing different strategies, for example, scanners, photos or by straightforwardly writing in the PC by utilizing a pointer.

B. Pre-processing

Pre-processing is the essential period of character Recognition and it's significant for acceptable Recognition rate. The fundamental goal of pre-handling steps is to standardize strokes and eliminate varieties that would somehow or another confuse Recognition and decrease the Recognition rate. These varieties or contortions incorporate the unpredictable size of text, missing focuses during pen development assortments, jitter present in content, left or right twist in handwriting and lopsided distances of focuses from neighboring positions. Pre-preparing incorporates five basic advances, to be specific, size standardization and focusing, interjecting missing focuses, smoothing, incline amendment and resampling of focuses.

C. Segmentation

Segmentation is achieved by the isolation of an image's individual characters. In general, paper is stored in a hierarchical fashion. Lines are segmented at the first stage using a row histogram. Words are derived from each row using the column histogram, and characters are eventually extracted from words.

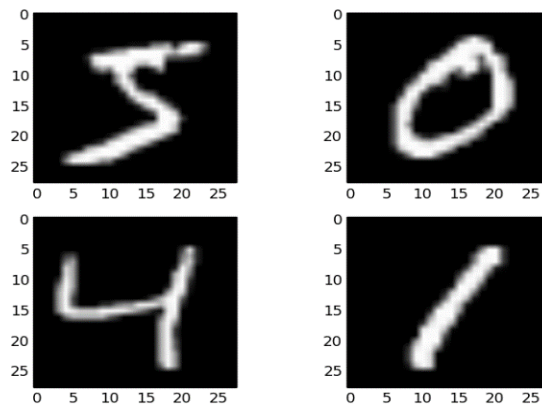


Fig.-1: Segmentation on Handwritten digits

D. Feature Extraction

The primary goal of the extraction process of features is to isolate the pattern that is most applicable to classification. Feature extraction techniques such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Chain Code (CC), Scale Invariant Feature Extraction (SIFT), zoning, gradient-based attributes, Histogram may be used to extract individual character features. To train the machine, these characteristics are used.

E. Classification

At the point when information picture is introduced to HCR framework, its highlights are removed and given as a contribution to the prepared classifier like fake neural organization or backing vector machine. Classifiers contrast the info include and put away example and discover the best coordinating class for input.

F. Post Processing

Post-processing alludes on the methodology of amending misclassified results by applying semantic information. Post-preparing is handling of the yield from shape Recognition. Language data can build the

precision got by unadulterated shape Recognition. For handwriting input, some shape recognizers yield a solitary series of characters, while others yield various choices for each character, regularly with a proportion of certainty for every other option.

IV. HISTORY

In the last part of the 1950s, SRI's Slash Crane started exploring different avenues regarding electronic pens. He knew about the character Recognition rationale of ERMA, SRI's bank check readers, and got inspired by strategies for perceiving typical handwriting continuously—rather than the Recognition of pre-printed data. This spearheading work established the framework for three significant applications Crane's initial work finished in a first patent in 1964 for an extraordinary pen gadget. These underlying endeavors took a further developed structure in the mid-1970s with a framework that pre-owned what was commonly alluded to as the "SRI pen" as a strategy for contributing characters into a PC. A fundamental strategy for contributing manually written Chinese characters was distributed later in the exact year [2]. The main patent in the territory of mark check was given in 1974. In 1991, SRI picked up a patent for a form of the pen that takes dynamic data in five measurements.

V. RECENT TRENDS AND MOVEMENTS

As the long periods of serious innovative work passed by, PCs turned out to be substantially more remarkable than previously. Individuals could compose the manner in which they typically did, and characters need not need to be composed like determined models, and the subject of unconstrained handwriting Recognition picked up significant energy and developed rapidly. Starting at now, numerous new calculations and procedures in pre-handling highlight extraction, and ground-breaking grouping strategies have been created.

There is huge exhibit of arising innovations in the field of character Recognition, these are as per the following: I. Numerous specialist designs for characterization of manually written content. ii. Displaying as a direction following issue. iii. Tree search/quick quest methods for streamlining. iv. Model sets for on-line Recognition of disconnected characters [5]. v. Joining word references [4] and vi. Hypothetical change approach [3]. Practically all these arising techniques are new giving promising structures, as they offer unique properties. Issue related highlights and calculations are popular. These strategies speak to a move from mathematical coordinating towards algorithmic hunt. They give a method for quick, normalized advancement and parallelism. A few strategies take into consideration creative blends of classifier yields which are symmetrical and corresponding to existing methodologies.

VI. CONCLUSION

In the wake of checking on the papers, it was seen that a few strategies like course include extraction and corner to corner highlight extraction procedures were end up being better in producing higher precision results contrasted with the customary level and vertical techniques. Likewise, neural organizations give an more component of having higher resistance to commotion.

Feed forward model of the Neural Organization is prepared utilizing Back Engendering to characterize and perceive characters. The blend of highlight extraction with standardization has demonstrated to yield higher exactness paces of character Recognition.

It was also observed that bigger the training data set, helps in achieving a higher accuracy rate when features are extracted from similar looking characters. This is highly beneficial; as handwritten characters appear similar, so good feature extraction techniques need to be used to avoid such anomalies.

VII. REFERENCES

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

- [3] Yusuf Perwej, Ashish Chaturvedi, "Neural networks for handwritten English alphabet recognition", arXiv preprint arXiv: 1205.3966, 2012.
- [4] Vijay Patil, Sanjay Shimpi, "Handwritten English Character Recognition Using Neural Network", Elixir Computer Science Eng. 41, pp. 5587-5591, 2011.
- [5] Wojciech Kacalak & Maciej Majewski, "Handwriting Recognition Methods using Artificial Neural Networks", pp 27 April 2015.
- [6] Megha Agarwal, Shalika, Vinam Tomar, Priyanka Gupta, "Handwritten Character Recognition using Neural Network and Tensor Flow", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue- 6S4, April 2019.