

A Survey on Handwritten Character Recognition

Rakesh. S¹ Shanmuga Priya. N²

¹MCA Student ²Associate Professor

^{1,2}Department of Computer Applications

^{1,2}Dr. SNS Rajalakshmi College of Arts and Science, Coimbatore, Tamil Nadu, India - 641049

Abstract— At current years Handwritten Character Recognition is main significant and admired research sector in the part of Image processing. In Handwritten edition there is no constraint on the writing style. Handwritten letters are not easy to recognize due to diverse human handwriting style, size and shape of letters. In a Handwritten character recognition, the set of geographies plays as foremost issues, as method in choosing the related feature that profits minimum classification fault. Handwriting recognition is most challenging area if image and pattern recognition. Handwriting recognition is very useful in real world. Text recognition in the handwritten documents has been studied as one of the projecting research areas by different researchers during the last few decades.

Key words: Handwritten Character Recognition, Features Extraction, Classifiers, Pre-Processing

I. INTRODUCTION

Object Character Recognition is an overall performance of handwritten texts or digitizing pictures of printed so that they could be electronically amended, stored and searched more efficiently and correctly. The human can do various tasks that are still terrible for machine to do by their own. One of such tasks is handwritten text recognition. As we know every individual has their different writing style, so it is very difficult to recognize the correct handwritten characters and digits. Handwriting recognition system is advanced to accomplish the exactness and dependable performance. There are several real problems where handwriting recognition system is selfsame useful like documentation analysis, mailing lecture interpretation, bank check processing, signature verification, postal addresses. Feature extraction in handwriting character recognition is a very main field of image processing and objects recognition. Fundamental component of characters are called features. Offline character recognition consist of recognition of machine printed, hand printed and handwritten characters. The most difficult problem in the field of Object character recognition is the recognition of unhindered cursive handwriting.

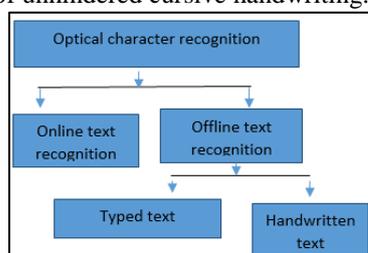


Fig. 1: Optical character recognition

II. STAGES OF HANDWRITTEN CHARACTER RECOGNITION

A. Preprocessing

Digitization of content through perusing the document, binarization of characters mined Noise elimination, retreating

the strokes, normalizing and filtering [6]. A series of procedures are performed on the input image during the pre-processing. It helps in enhancing the image rendering and makes the image suitable for segmentation [2]. In pre-processing stage various operation are performed like on image like binarization, noise removing, and edge recognition [3]. The preprocessing stage aims to extract the appropriate textual parts and prepares them for segmentation and recognition. The main intentions of preprocessing are noise reduction, normalization of data and compression in the amount of information to be retained [4]. Pre-Processing can be defined as cleaning the document image and making it appropriate for input to the OCR device [5].

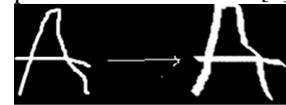


Fig. 2: Preprocessing segment results, including noise removal and skew correction

B. Removal of Noise

The preprocessing stage boosts the preeminence of the raw image and puts the data of significance. It is also admitted as pixel level or low level processing, which is arranged on the apprehended image to manage it for additional analysis. Skew correction of an image is executes in preprocessing segment to perfect the article text lines, dawn to switch a gray scale or color image into a binary image, and decrease of noise to decrease superfluous preprocessing phase results, including noise removal and skew correction perused documents can be soiled with dust, spots, dots, or lines, which are classified as noise that affect the recognition results to a large degree. Consecutively to construct it appropriate for additional processing, a scanned article image has to be freed from any existing noise. Image enrichment methods are applied to improve the image being regarded by the machine or human. Smoothing and nonlinear operations, such as morphological operations, are utilized for noise decrease.

C. Skew Correction

Skew correction is used to correct the text line in skimmed document images if they are not horizontally affiliated during the scanning process. The skew correction can be applied on the manuscript or line level. In the preprocessing step, a universal skew correction is applied on the manuscript level.

D. Feature Extraction

Geographies play an important role in handwriting recognition systems. Their main area is to increase the recognition rate by effectually representing the data. Feature extractions promises with extracting mainly with the necessary information from image pixels liable upon the difficulty to be resolved and the utilized data[1]. The bit map version realm the major features of input image in shorter planetary data length. Such that reduces the time intervened in NN Training without affecting the exactness of correct

character recognition[2]. Feature extraction can be distinct as extracting the furthest demonstrative information from the rare data, which decreases the within class outline variability while cultivating between class pattern variability[4].

E. Classification

For the recognition step a confidential classification outline is employed. Since characters with similar shapes i.e. ξ and ζ from the greek alphabet, are often commonly confused when using a low granularity feature representation, we proposition to combine the corresponding classes to the certain level of classification [7]. There can be a assortment of improvements on the work to investigate the not MNIST dataset further. There has been a wealth of algorithms executed for OCR, with fluctuating degrees of success, up to classification errors of less than 1%: neural networks, various other SVMs with different cores [8].

F. Recognition

Recognition of handwritten English alphabets is awfully tough problem. The English Alphabet might be scripted in various size, direction, width, arrangement and measurement. This might afford endless differences. The probable of neural network (NN) to simplify and insensitive to the misplaced information might be extremely helpful in categorizing handwritten English alphabets[1].

III. APPLICATIONS

Japanist Fujitsu's OLCR technology is castoff in several software products. Japanist 2003, which is a Japanese involvement software for Windows and for Pocket PC, which is a handwriting character recognition software for Pocket PCs. It also familiarize alternate OLCR software for PDAs that are emergent [13].

A. Japanist 2003 Software

Fujitsu Japanese input software that comprises a kana-kanji conversion (KKC) purpose and a dictionary search function. Japanist 2003 has two input methods in the handwriting input square. In the writing-box mode, written characters are predictable one by one and the text result is demonstrated in the result area. If there are forecast terms, they are presented in the prediction area above the result area. In the box-free mode, the written characters are ordinary all together later they have been input. The recognition outcome is displayed in the result area, as in the writing-box mode. However, the prediction task does not work in the box-free mode. Because both of these input modes approve hybrid adaptation, the recognition precision increases as the user continues to use the software [13].

B. Japanist for Pocket PC (Personal Computer)

Japanist for Pocket PC (Personal Computer) is an OLCR software for Pocket PCs. It has virtually the same routine as the writing-box mode of the Japanist 2003 handwriting input panel. It adapts to the text input framework condition for Software Input Panels (SIPs), it works as slice of the common text input scheme on Pocket PCs. In accumulation to the basic OLCR functions, it affords a customizable graphical user interface (GUI). Example, number of writing boxes, and button spots can be changed affording to the user's inclinations [13].

C. Real Time Box-Free Handwriting

Recognition GUI for PDAs Next, Writing boxes use a huge display area, they may make it challenging to input on a insignificant display device such as a PDA. To diminish the required quantity of display area, box-free OLCR software is developed, that does not practice writing box [13].

IV. RELATED WORK

Handwritten alphabets recognition system's accuracy of any image relay upon the sensitivity of the choice of features and category of classifier utilized. Therefore, so many feature deletion and categorization applies could be establish in the literature. Following research work execute handwritten alphabets recognition of English phrases. Later a very long time, human used to write their opinions in the form of letter, copies etc.; in order to bear them to others[1].

V. CONCLUSION

The chief approaches consumed in the domain of recognition of handwritten English alphabets through the previous period have been studied in this survey paper. This review paper epitomize the different technique are available for diagnose the hand written documents. This review paper also attentions on that in today's world hand writing reorganization is very difficult but very important. The recent movement on feature extraction and selection were discussed briefly. We also were investigating on metaheuristic algorithm which is coordination search algorithm as an optimization tool. Different pre-processing, segmentation, feature extraction, classification procedures are also discussed. Though, various systems for treating the problem of hand written English letters have advanced in last two decades, still a lot of study is needed so that a worthwhile software solution can be finished available.

REFERENCES

- [1] A Survey On Handwritten Character Recognition (HCR) Techniques For English Alphabets (ManojSonkusare and Narendra Sahu)-March 2016
- [2] Handwritten Text Recognition System Based on Neural Network IAhmed Mahdi Obaid, IIAhmed M. El Bakry, IIIM.A. Eldosuky, IVA.I. Shehab-March 2016
- [3] Handwriting Recognition System-A Review Pooja Yadav, Nidhika Yadav Muhammad, Arif Mohamad, Haswadi Hassan, DewiNasien and HabibollahHaro-March 2015
- [4] A Review on Feature Extraction and Feature Selection for Handwritten Character Recognition Muhammad Arif Mohamad, Dewi Nasien, Haswadi Hassan, Haiollah Haron-2015
- [5] Recognition for Handwritten English Letters: A Review Nisha Sharma, TusharPatnaik, Bhupendra Kumar-Jan 2013
- [6] Noise Reduction and Pre-processing techniques in Handwritten Character Recognition using Neural Networks
- [7] MageshKasthuri, V.Shanthi-Jan 2014 A Study on Preprocessing Techniques for the Character Recognition Poovizhi P-2014

- [8] Hierarchical Classification of Handwritten Characters based on Novel Structural Features G. Vamvakas, B. Gatos and S.J. Perantonis
- [9] Optical Character Recognition: Classification of Handwritten Digits and Computer Fonts George Margulis
- [10] Segmentation Methods for Hand Written Character Recognition Namrata Dave-2015
- [11] A New Segmentation Algorithm for Handwritten Word Recognition M. Blumenstein¹ and B. Verma^{1 2}
- [12] Study Of Various Character Segmentation Techniques for Handwritten Off-line Cursive Words: A Review Amandeep Kaur, SeemaBaghila, Sunil kumar-2015
- [13] Online Handwritten Recognition Technology and its Applications. Hiroshi Tanaka, Naomi Iwayama, Katusuhiko Akiyama-2003

