Handwritten Character Recognition Using CNN, KVM, and SVM

Mr. Yash Kotkar[1] Mr. Chirag Sankhe[2] Ms. Bhumika Nair[3] Mr. Rishon Yaragal[4] Mr. Kashif Shaikh[5]
[5]Lecturer
[1,2,3,4,5]Department of Computer Engineering
[1,2,3,4,5]Thakur Polytechnic, Kandivali, India

Abstract — One of the most active and difficult research areas in the field of image processing and pattern recognition has been handwritten character identification. It can be used for many tasks, such as creating organized text out of any written content, bank checks, and blind reading assistance. This research helps others with the understanding of handwritten characters for English alphabets without the need for feature extraction using only a multilayer feed-forward neural network. The character data sets each contains 26 alphabets. The neural network is trained using 50 different personality data sets. The trained network is used for classification as well as for recognition. Each character's size in the displayed system has been expanded to 30x20 pixels, followed by immediate training. To train the training variables for a character's scaled neural network. The findings show that the suggested approach matches feature extraction-based strategies in terms of handwritten character recognition rates.

Keywords: Handwritten Character Recognition, CNN, KVM, and SVM

I. INTRODUCTION

Identification, fragmentation, and recognition of characters in photographs are the job of character identification [1]. The ultimate goal of handwritten character recognition is to imitate human reading abilities so that a machine can scan, update, and interact with text in a manner that is almost identical to that of a human [2]. Handwriting recognition has emerged in recent years as one of the most fascinating and challenging research areas in the fields of image identification and pattern classification. In many applications, it enhances human-machine interaction and significantly contributes to the advancement of automated operations. Numerous studies have concentrated on creating novel approaches and techniques that would speed up processing while improving recognition accuracy [3][4]. Character recognition comes in two flavors: online and offline. Online character recognition uses a special pen to record writing activity on a digital surface. Data that has already been typed on a piece of paper is scanned via offline recognition. Offline Character Recognition: All typed or printed characters are categorized in the offline mode. The method of identifying characters in a document that has been scanned from a flat surface and recorded digitally in greyscale is known as offline character recognition. When processing scanned documents, it can be challenging or impossible to perform many processing tasks such as information retrieval, alteration, and management. Applications requiring handwritten text recognition systems include postal address recognition, document analysis, bank transaction processing, and mail processing. Because of this, offline handwritten text recognition is still a popular topic for research, with specialists looking into novel strategies to improve recognition precision.

II. PROPOSED SYSTEM

The proposed recognition method is explained in this section. Preprocessing, segmentation, classification, and post-processing are the four stages of a typical handwriting recognition system. Figure 1 is a schematic illustration of the recognition mechanism.

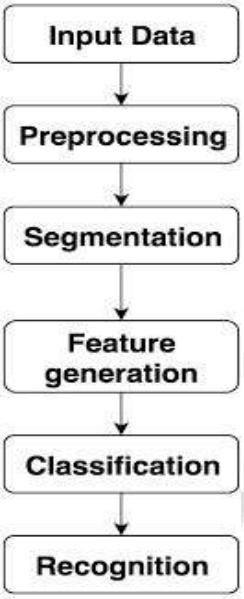

Fig. 1: Schematic Diagram of Proposed Handwritten Character Recognition System

A. Input Data / Image Acquisition

The recognition device takes a scanned image as an input image during image capture. It is advised that you use a specific picture format, such as JPEG or PNG.

B. Pre-processing

Preprocessing involves several operations on the scanned raw images. It significantly enhances the image and gets it ready for segmentation. Figure 2 shows the many operations that were carried out on the image during the pre-processing phase. The binarization technique employs the general thresholding technique to convert a binary image from a grayscale image. In the following two phases, the image is additionally dilated and filled with holes to form a preprocessed image suitable for segmentation. The Sobel algorithmic method is used to lengthen the edges in the binarized image.

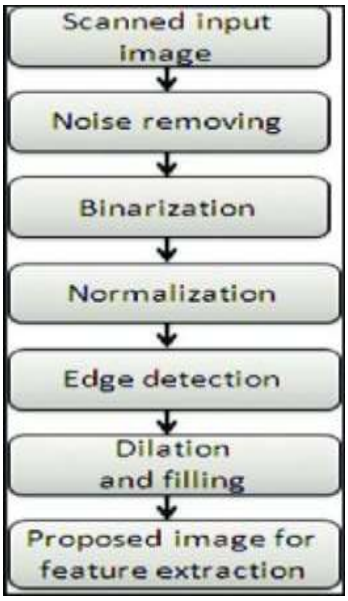

Fig. 2: Pre-processing stage of image.

C. Segmentation

A picture of a sequence of characters is separated into sub-photos of individual characters during the segmentation stage. During the segmentation stage, a photograph of a series of characters is divided into sub-pictures of individual characters. This symbol represents the number of characters in the image. Every character is scaled down to a size of 30X20 pixels.

D. Classification and Recognition

Decisions are made during the categorization stage of the recognition system. This study uses a feed-forward back propagation neural network to recognize and recognize handwritten characters. 600 pixels from the segmentation stage's compressed character are fed into the classifier as input. The neural classifier consists of two hidden layers: an input layer and an output layer. Because one of the characters should always be recognized, the output layer is a competitor layer in the hidden levels that apply the log sigmoid activation function. Since the suggested system is built to recognize the English alphabet, the output layer includes 26 neurons.

E. Post-Processing

The post-processing stage is the top level of the suggested recognition system. The test samples' recognition value is used to calculate the equivalent ASCII value, and the relevant recognized characters are then published in an ordered textual style.

III. Pre-processing method

The raw data undergoes several initial processing steps to make it useful for the descriptive phases of character analysis. Pre-processing aims to create data that is simple for OCR systems to accurately process. The following are the preprocessing's main objectives:

- Noise Reduction
- Binarization
- Edge detection
- Dilation and filling
- Segmentation of processed image

A. Noise Reduction

When a paper is scanned, extra noise can distort the scanned images, which will hinder the subsequent stage of document processing. Therefore, before sending pictures to the following document stage, to enhance their quality before processing, they must go through a preprocessing step. It is crucial to get rid of all of these flaws because noise can lead to disconnected line segments, enormous gaps between lines, and other issues. This way, the information can be retrieved as quickly as possible. There are many distinct kinds of noise in photos. It is possible to see "Salt and Pepper Noise," a single add-on noise that is composed of black and white spots scattered throughout a photo and resembles salt and pepper. The two primary categories of noise reduction techniques are filtering and morphological processes.

B. Filtering

Its objective is to lessen noise and false points, which are typically brought on by an uneven writing surface and/or a data collection tool with a low sample rate. A variety of spatial and frequency domain filters can be used for this.

C. Morphological Operations

When processing images, morphological methods are widely employed to extract important data for representing and describing how a region is formed. Due to low-quality ink and paper, as well as erratic hand movement, noise on document pictures can be effectively eliminated utilizing morphological techniques.

D. Binarization

Binarization of the grayscale input image is a critical step in the process of offline character recognition. Good binarization facilitates character segmentation and recognition. A grayscale image is transformed into a binary image using the binarization process. This work presents new techniques for binarizing noisy grayscale character images captured in an industrial setting. Because characters are frequently composed of minuscule lines of constant width, our methods were specifically developed to binarize grey-scale character images more successfully. The results of trials show that these methods yield the best binarization results.

E. Edge Detection

For object segmentation, registration, and identification, edges serve as the definition of object boundaries. Edge detection preserves the image's essential structural characteristics while reducing the amount of data in the image and filtering out extraneous information. There are many different ways to identify edges. Contrarily, gradient, and Laplacian are the two main types of gradients used in most approaches. To find edges, the gradient approach examines the first derivative of the image for upper and lower boundaries. The Laplacian method searches for zero crossings in the second derivative of the image for finding edges.

F. Thresholding:

Greyscale or color photographs are frequently encoded as binary images to reduce file size and hasten to process. To do

this, choose a value above which everything is set to 1, and below which everything is set to 0. There are two forms of thresholding: global and adaptive. Global thresholding selects a single threshold value for the entire scanned image, frequently based on an estimate of the background level derived from the intensity histogram of the image. Changing threshold settings for various sections of the image is necessary when using the image processing technique known as adaptive thresholding. A goal-directed evaluation criterion that contrasts the accuracy of various character recognition systems is used to compare several common thresholding approaches. The best outcomes were produced by Niblack's strategy.

G. Dilation and filling

1) Skew Detection

In a document scanning process, skewness is a possibility. There are numerous techniques for identifying skew in a page; some concentrate on discovering related parts and figuring out the mean angles between their centroids. Skewness should be removed because it reduces the document's accuracy. The skew angle is established, and with the aid of the skew angle, the skewed lines are made horizontal.

2) Slant Estimation and Normalization

Slanted characters are a common feature of handwritten text. For example, slanted characters slant from right to left or vice versa. Various deviations might even happen within a single word, not just within a paragraph. The slant adjustment does not affect the word's connectedness, and the words that arise are natural. To normalize all characters to a consistent form, slant normalization is utilized. The most common method for slant estimation is to calculate the average angle of near-vertical portions. The Wigner-Ville distribution and the vertical projection profile of word pictures are used in this study to propose a slant removal strategy. This method divides the image into vertical and horizontal windows to identify the slant. The tilt is calculated by averaging the center of gravity of each window's top and lower parts. In another study, multiple approaches for estimating and correcting average slants were offered. When a word's slant shifts from character to character, however, average slant assessment has the problem of overestimating or underestimating local slant. This work provides three local slant estimation approaches to tackle the problem: an 8-directional chain code approach, a simple iterative method, and a high-speed iterative method. The results of the experiments indicate that the offered strategies are effective and more accurate than average slant correction in estimating and correcting local tilt. Finally, a variation of the Hough transform is utilized, which involves scanning the image from left to right and calculating projections in 21 distinct slants. Any slant's top three projections are put together, and the slanting value is determined by the slant with the highest count.

H. Segmentation

The most crucial step in character recognition algorithms is segmentation. Segmentation is a technique for dividing distinct characters in a photograph. Unrestricted handwritten words are more challenging to divide into upper, middle, and bottom characters and zones than printouts. This can be attributed to several characteristics, including inter-character spacing, skew, tilt, size, and bent handwriting. The segmentation task gets substantially more challenging when parts of two succeeding characters touch or overlap. Due to altered upper-zone and lower-zone letters, such touching or overlapping is frequent in Indian languages. The ability to distinguish words, lines, or characters directly affects the script's recognition rate, making segmentation a crucial stage. Segmentation can be divided into two types:

I. External Segmentation

External segmentation is used to separate the website layout into logical sections. External segmentation is a method for dividing a written piece into sections, sentences, or words. It is a crucial component of document examination. The goal of DAR (Document Analysis and Recognition) is to retrieve data from printed documents that were originally meant for human interpretation. The segmentation of the document image into text and non-text components is necessary for OCR software. As a result, everyone involved in the CR industry needs to have a fundamental understanding of document analysis techniques. Page segmentation, one of the most important processes in layout analysis, is particularly difficult when dealing with complex layouts.

The two stages of page layout analysis are as follows:

The first phase is the structural analysis, which requires segmenting the image into various document-building units (section, line, word, etc.). The second method is functional analysis, which uses layout principles like size, location, and alignment to identify the functional content of document components.

The next step in the page segmentation process involves scanning grayscale or color images for textured areas. For instance, a technique for evaluating the effectiveness of document page segmentation algorithms is shown. They suggested utilizing an automated bitmap-level method to test page segmentation algorithms on the mixed text and half-tone data. A precise qualitative diagnosis provides a quantitative assessment of segmentation methods.

J. Internal segmentation

Internal segmentation is a technique for breaking down a picture of a series of characters into individual symbol sub-pictures. Despite tremendous progress in the previous decade and the emergence of several solutions, segmenting cursive writing into letters remains an unresolved task.

IV. Experimental results:

The IAM database had 13,353 images of the handwritten text created by 657 authors. The IAM database, which includes contributions from 657 authors and has 1,539 handwritten pages with 115,320 words, is regarded as a component of the current collection. The database is labeled at the word, line, and sentence levels. Each character underwent preprocessing using Fourier extraction, and a Backpropagation Neural Network was then used to recognize and identify each one. Ten nodes were present in the input layer, along with seventy, forty, and thirty nodes inside the hidden layer, and one node in the output layer. The network employed in the test set process accounted for 87 percent of the total. We train the system to find the centroid and membership function of the

prototype using the RCP training method. Finally, the recognition stage uses the FD of an unidentified character image. These observations allow the suggested system to operate with up to 91.5% accuracy.

V. CONCLUSION

This work uses a CNN-based handwritten character recognition system to categorize and identifies the 26 English alphabets and words. The scaled characters from the segmentation stage's pixel values were used to immediately train the convolutional neural network. The neural network architecture with two hidden layers, each with 100 neurons, was found to have the highest recognition accuracy of 89.01 percent out of all the neural network architectures tested to classify the characters.

The handwritten recognition approach shown in this research has potential applications in handwritten name recognition, document scanning, and conversion of any handwritten content into textual format.

ACKNOWLEDGMENT

I would want to express my gratitude to the writers who contributed to the study and made it possible. I also want to thank Thakur Polytechnic's Computer Engineering Department and our Guide for their assistance.

REFERENCES

[1] Mori, C.Y. Suen, and K. Kamamoto, "Historical review of OCR research and development," Proc. of IEEE, July 1992.
[2] N. Arica and F. Yarman-Vural, "An Overview of Character Recognition Focused on Off-line Handwriting", IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2001.
[3] V.K. Govindan and A.P. Shivaprasad, "Character Recognition – A review," Pattern Recognition, 1990.
[4] R. Plamondon and S. N. Srihari, "Online and off-line handwritten character recognition: A comprehensive survey," IEEE. Transactions on Pattern Analysis and Machine Intelligence, 2000.
[5] U. Bhattacharya, and B. B. Chaudhuri, "Handwritten numeral databases of Indian scripts and multistage recognition of mixed numerals," IEEE Transaction on Pattern Analysis and machine intelligence, 2009.
[6] Martin Rajnoha, Radim Burget and Malay Kishore Dutta, "Handwriting Comenia Script Recognition with Convolutional Neural Network", 2017 40th International Conference on Telecommunications and Signal Processing (TSP).
[7] Lyzandra D'souza and Maruska Mascarenhas, "Offline Handwritten Mathematical Expression Recognition using Convolutional Neural Network", 2018 International Conference on Information, Communication, Engineering, and Technology (ICICET).
[8] Chunpeng Wu, Wei Fan, Yuan He, Jun Sun and Satoshi Naoi, "Handwritten Character Recognition by Alternately Trained Relaxation Convolutional Neural Network", 2014 14th International Conference on Frontiers in Handwriting Recognition.
[9] Curtis Wigington, Seth Stewart, Brian Davis, and Bill Barrett, "Data Augmentation for Recognition of Handwritten Words and Lines using a CNN-LSTM Network", 2017 14th IAPR International Conference on Document Analysis and Recognition.
[10] Nagender Aneja and Sandhya Aneja, "Transfer Learning using CNN for Handwritten Devanagari Character Recognition", 2019 1st International Conference on Advances in Information Technology.

