A Comprehensive Approach for Real-Time Character Recognition System for Image Video and Handwritten Documents

[1]Dr. Jayshree R. Pansare, [2]Aditi Gaikwad, [3]Vaishnavi Ankam, [4]Shikha Sharma, [5]Priyanka Karne

MES College of Engineering,
SPPU, Pune, India

***Abstract*: Text categorization is a challenging task when it comes to categorizing text from different sources such as images, videos, and handwritten text. Handwritten text may vary as per the diversified user. Hence, it is difficult to find the best technique to categorize such kind of texts due to the unavailability of standard dataset and evaluation measures. Our system presents a standard method for recognition of text from all kinds of aforementioned input sources using the Support Vector Machine (SVM) classifier. Additionally, it classifies and places the words into predefined classes of parts of speech for English language using Deep Learning algorithms.**

***Index Terms*: Real-Time Character Recognition, SVM, Optical Character Recognition (OCR), Deep Learning algorithms.**

I. INTRODUCTION

Text categorization is a powerful educational tool especially for different language learners like English, Chinese, Kannada, French, and German. Images, videos and handwritten texts are the sources for text recognition. Extraction of text from the running slide of video by using detection algorithm and localization method the 90.8% accuracy is achieved [1]. Ground truth generation by Semi-automatic method for identification of parallel and nonparallel text. annotation time is reduced by 22% using segmentation for multimedia documents [2]. For text recognition we need a huge database. Some of the databases are insufficient for variety of reasons so database of city name, state name, ZIP code for recognition on handwritten addresses is assumed. There are about 3000 classes for city names and 42000 for ZIP codes. This database overcomes the problem of previous databases which were scanned in the laboratory environment [3]. A handwritten text recognition using Multiwriter task and Writer independent task have the accuracy 49.1%. After addition of SCFG based syntax analysis the accuracy improves from 49.1% to 54.4% [4]. Recognition rate is achieved at 100% for English text recognition based on features combination. Interference of external noise may be reduced by using proposed method of English character recognition[5].

For recognizing Kannada and English character accuracy achieved are 73.33% for Kannada and 96.13% for English lowercase alphabets using SVM classifier [6]. Through text recognition like humans, machines can also read the text, comprehend its meaning and program itself according to the command. It involves text acquisition, text identification, image to text transformation, character recognition. This study is considered as a step towards bringing the machine world more closer to the humans world [7]. A text recognition from various types of document is a difficult task so by using image processing it is possible to recognize the patterns. Segmentation is used for detecting word, line or character from the pre-processed image using convolutional neural network and genetic algorithm for feature mining and feed-forward network as a classifier. After classification we get desired output[8]. Using segmentation and tokenization method obtain accuracy of 70% for Name entities recognition and classification [9]. Text document classification by fusion of sentence vector space model and bi-gram model takes thousands of text documents and create output as different folders containing the same types of documents [10].

The paper is organized into following Sections: Section I emphasizes several existing techniques in Text recognition for Image, Video and Handwritten documents. Section II describes the overall structure of the proposed method. In Section III, we present the Comparative Performance of Real-Time Character Recognition System based on the performance of specific systems. We conclude with the conclusion in Section IV.

II. RELATED WORK

Image text recognition is specially divided into four phases namely: detection, localization, extraction and recognition. The detection phase classifies the regions into non-text regions and text regions. The localization phase finds the boundaries of words and strings. The extraction phase filters the background image. Fig. depicts how image is converted into binary image using detection, localization and extraction method. The recognition phase is done by OCR software.

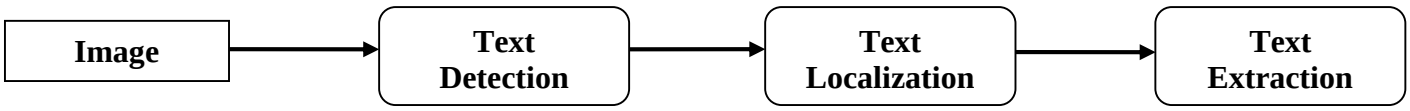

FIG: CONVERSION OF INPUT IMAGE INTO BINARY IMAGE

Text Detection:
The primary phase in text recognition techniques is text detection. This method is classified into two different classes. In the first texture class the whole image is divided into blocks. For this purpose various approaches can be used, e.g. wavelet transforms, spatial variance, or Gabor filter. In second class the text block and non-text blocks are classified. It may be done by using the neural network or support vector machine. To classify text and non-text block the background of an image should be clear. For background elimination of image background-complexity-adaptive thresholding algorithm is used.

Text Localization:
Localization is secondary phase in text recognition. In localization phase the accurate boundaries of text are determines. There are two methods for text localization first is bottom-up paradigm and another one is top-down paradigm. In bottom-up paradigm groups a small text regions into entire text strings based on some rules. In top-down approach is based on splitting the image into regions. The first region may be whole image, and then dividing this whole image in horizontal and vertical directions based on color or edge distribution.

Text Extraction:
The text extraction method filters the background image. Extraction may be done by two groups. One is color-based method and other is stroke-based method. The color-based method is based on assumption that the text pixels color is different than background pixel color. On the other stroke-based method apply some filters only on specific content to filter the background image.

III. PROPOSED SYSTEM

In this section we focus on System Architecture and Work Flow.

(A) System Architecture

Specially, the process adopted in Text Recognition has been shown in Fig. is fragmented in five consequent phases applied to captured image/handwritten text namely; Image acquisition, Pre-Processing, Segmentation, Word Extraction and Feature Extraction.

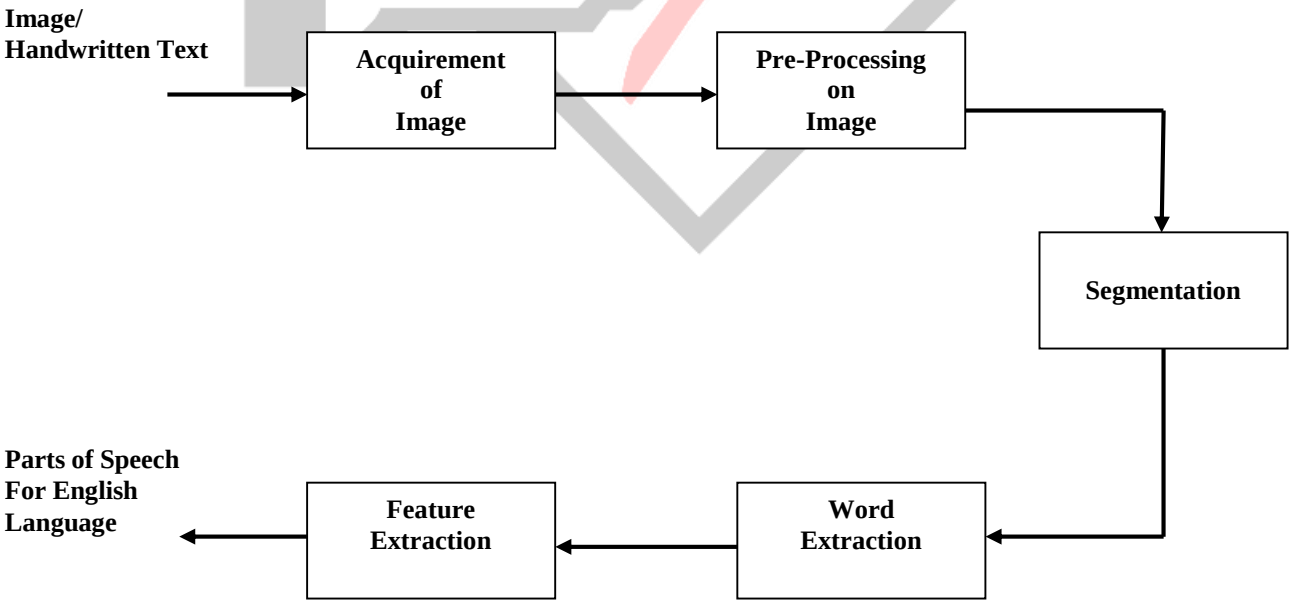

Fig: System Architecture of Real-Time Character Recognizer

1. Image Acquisition:

The input is taken as scanned image .The image is required in .JPEG, .JPG and .PNG format and all images should be in machine editable arrangement. Input device such as scanner and digital camera are used to capture the image.

2. Pre-processing:

Pre-processing especially done in order to improve the quality of image. There are some basic methods for image pre-processing and some of them are noise diminution, edge identification, dilating and filing of image.

Noise Diminution

Noise reduction has variety of techniques to remove the noise from an image or video and makes it easy to compress. There are some algorithms which try to blur the unwanted part of image. Filtering and morphological operations are used for noise reduction.

Binarization

Binarization means conversion of gray scale image to binary image. White and black are two colors used for binary image. Thresholding method is used for binarization. In thresholding the pixels which are similar in gray scale are put in same group.

Edge Detection

Edge detection is a technique used for detection the boundaries of object in images. Discontinuities in image are detected by this method. Canny technique, Sobel method, Prewitt method, Roberts's system and fuzzy logic all this methods are used to perform edge detection.

Dilation and filing

In dilation the pixels are added to the borders in the image, on another side erosion is the practice of removing pixels on item's margins of image. The process in which holes present in images are filled is called as filing.

3. Segmentation:

Segmentation means dividing an image into multiple parts. There are different ways of segmentation; thresholding method such as Otsu's method, Color-based Segmentation such as k-means clustering, transforms methods such as watershed segmentation. In segmentation each character in the word is individually segmented.

4. Word extraction and Feature Extraction:

This approach is useful when image size is large and we require reduced feature representation. Convolutional Neural Network (CNN) is a powerful technique based on deep learning. In diagonal based feature extraction technique every single image is resized to 100*100 pixels and then it is subdivided into 10*10 pixels.

5. Classification and Recognition of characters:

Classifier is used for reaching to accurate result. It is very essential to choose a good classifier and best feature mining method. Feed forward network is used for categorization and identification. The classifiers are categorize as supervised and unsupervised classifiers. Decision tree, Artificial Neural Network (ANN), Bayes' classifier, Nearest Neighbor classifier are some classification methods.

(B) Work Flow

Work flow of our system is presented in Fig. It comprises the correlated methodologies required to implement for all the aforesaid phases of system architecture.

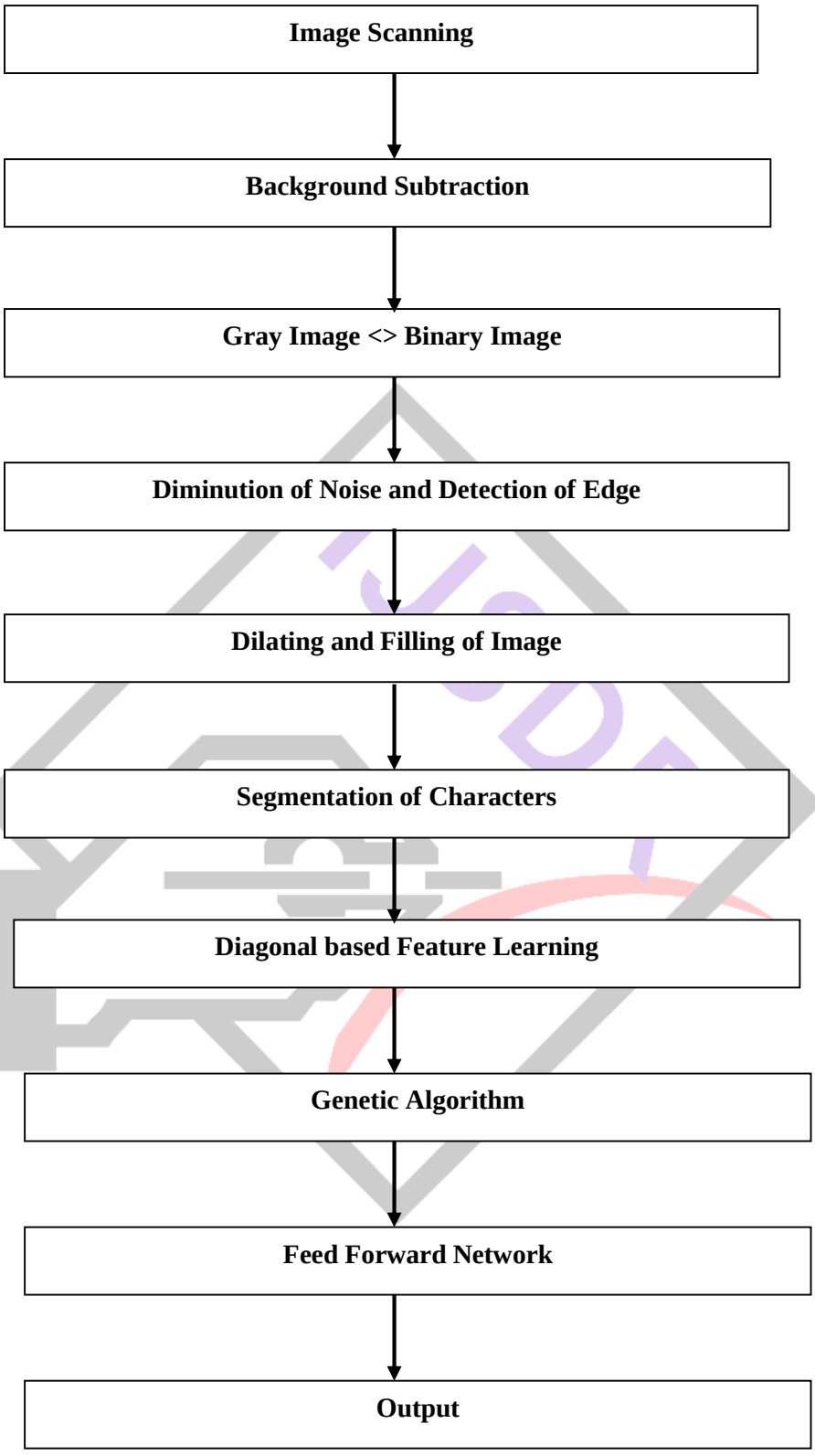

Fig: Work Flow of Real-Time Character Recognizer

IV. Comparative Performance of Real-Time Character Recognition System

This section emphasizes on comparative performance of diversified systems applied for recognition of various characters as shown in Table It is observed that Character Recognizer System detects English lowercase alphabets using SVM classifier achieves the highest performance accuracy of 96.13%. Moreover, Writer Independent Text Recognizer identifies handwritten text by employing Stochastic Context-Free Grammar achieves the lowest recognition rate of 49.1%.

Table: Comparative Study for Text Recognition

Sr. No.	System	Methodology	Input	Accuracy
1	OCR system	Text discovery and localization by OCR method.	Text in current slide of video	67.3%
2	Kannada SVM classifier	SVM classifier character recognition	Kannada consonants	73.33%
3	English SVM classifier	SVM classifier character recognition	English lowercase alphabets	96.13%
4	Kannada and English SVM classifier	SVM classifier character recognition	Mixed English and Kannada characters	83.02%
5	DNN for parts of speech	Deep neural network Cosine similarity for Name entity recognizer	English sentences	50% - 60%
6	BP Network System	Recognition of character using BP network	English text	100% Recognition Rate
7	Recognition System	Writer Independent task for text recognition	Handwritten text	49.1%
8	Text Recognition System	SCFG-based syntax analysis for text recognition	Handwritten text	51.09%

V. Conclusion

We have studied many techniques used to recognize texts in detail. In this paper, we developed an approach to recognize texts in detail. A new version for classifying parts of speech of words has been presented. We have combined methods like SVM classifier and Deep Learning. The system mentioned in this paper works only for the English language. In future similar works can be carried out for other Languages like Chinese, Kannada etc.

REFERENCES

[1] M. R. Lyu, et.al. , "A Comprehensive Method For Multilingual Video Text Detection , Localization , and Extraction " , IEEE Transaction On Circuits And Systems For Video Technology, vol.15, no.2, pp.243-255, 2005. doi:10.1109/TCSVT.2004.841653

[2] T. Q. Phan, et.al. , "Semiautomatic Ground Truth Generation for Text Detection and Recognition in Video Images", IEEE Transaction On Circuits And Systems For Video Technology, , vol.24, no.8, pp.1277-1287,2014. doi:10.1109/TCSVT.2014.2305515

[3] J. J. Hull, "A Database for handwritten Text Recognition Research", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.16, no.5. pp:550-554,2017.doi:0162-8828/94

[4] M. Zimmermann, et.al. , "Offline Grammar-Based Recognition of Handwritten Sentences", IEEE Transaction On Pattern Analysis And Machine Intelligence, Vol. 28, No. 5. pp. 818 – 821, 2006.

[5] Y. Yang, et.al. , "English Character Recognition Based on Feature Combination" , 2011 International Conference on Advances in Engineering , pp.159-164,2011. doi:10.1016/j.proeng.2011.11.2619

[6] G. Mukarambi, et.al. , "A Zone Based Character Recognition Engine For Kannada And English Scripts", pp 3292 – 3299, 2012.

[7] S. L. Wasankar, et.al. , "Machine Learning with Text Recognition", doi:10.1109/ICCIC.2010.5705811

[8] U. Yadav, et.al. , "A Deep Learning Based Character Recognition System From Multimedia Document", International Conference on Innovations on Power and Advanced Computing Technologies. pp:-1-7,2017

[9] S. P. Singh, et.al. , "Deep Neural Based Name Entity Recognizer and Classifier for English Language", Proceeding of Second International conference on Circuits, Controls and Communications, pp. 241 – 246, 2017.

[10] A. D. Amensisa, et.al. , "A Survey on Text Document Categorization Using Enhanced Sentence Vector Space Model and Bi-Gram Text Representation Model based on Novel Fusion Techniques", IEEE Xplore Proceedings of the Second International Conference on Inventive Systems and Control(ICISC 2018), pp.218225,2018.

[11] J. R. Pansare and M. Ingle, "Comprehensive Performance Study of Existing Techniques in Hand Gesture Recognition System for Sign Languages," Int. J. Computer. Sci. Inf. Technol., vol. 7, no. 3, pp. 1343–1347, 2016