Journal of Theoretical and Applied Information Technology
31st October 2019. Vol.97. No 20

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DEVELOPMENT OF AN ANDROID APP FOR TEXT DETECTION

[1]MUHAMMAD AHMED ZAKI, [2]SAMMER ZAI, [3]MUHAMMAD AHSAN, [4]UROOBA ZAKI
[1]ME Scholar, Mehran University of Engineering and Technology, Jamshoro, Department of Computer System Engineering, Pakistan
[2,3]Assistant Professor, Mehran University of Engineering and Technology, Jamshoro, Department of Computer System Engineering, Pakistan
[4]M.Phil Scholar, University of Sindh, Jamshoro, Institute of Information and Communication Technology, Pakistan

*Corresponding Author

E-mail: [1]ahmed_zaki64532@yahoo.com, [2*]sammer.zai@faculty.muet.edu.pk, [3]ahsan@faculty.muet.edu.pk, [4]zurooba6@yahoo.com

ABSTRACT

The field of text recognition is an important area as it finds various useful applications in computer vision such as document analysis, image search, robot navigation etc. Even though, a bunch of research has been carried out in this area but still there is a room for the improvement as none of the presented methods are error-free due to the challenges that may include the font of the characters, multi-oriented text, and the quality of pictures that contain text. Text detection is an important step towards the various content-based image retrieval techniques (CBIR). This study aim is to provide a flexible method to extract texts from challenging images and show them line-by-line by employing basic image processing techniques. The proposed system is an image processing and data retrieval system for English text. The system tends to build up an application that can recognize the English content present in an image captured by using a smart phone and then display the recognized text into editable format on the mobile screen. In order to show the effectiveness of the proposed technique extensive experiments have been performed on smart phone and the results are compared with various top-ranked commercial mobile apps.

Keywords: *CBIR , OCR, Android App Development, image processing, Tesseract, Artificial Intelligence.*

1. INTRODUCTION

Text recognition is the process of extracting text data from given images which is also considered as Optical Character Recognition (OCR) [1]. The problem of text recognition from images belongs to a class of computer vision, which may involve OCR or text detection or handwritten text detection. Simple text detection refers to find out the printed text on the images. Along with all the contents in images, text information is of immense interests, as it can be effortlessly understood by human and computer, hence finds wide applications in multimedia systems, digital libraries, and geographical information systems and many more [2]. The habitual recognition of text in natural images, is a significant challenge for visual understanding. Text, as the physical life of language, is one of the basic tools for preserving and communicating in pattern. Contemporary world is considered to be interpreted through the use of textual cues in addition to other labels, so text finds itself spread during many videos and images.

Habitually, text recognition has been paying attention on document images, where OCR techniques are appropriate to digitize planar, manuscript based documents. Still, when applied to natural scene images, these document OCR techniques fail as they are tuned to the largely black along with white, line based background of written documents. The text that occurs in natural scene images is vastly changeable in exterior as well as layout, being drawn from a huge number of fonts moreover styles, distress from incompatible illumination, occlusions, orientations, noise in addition, and the presence of background objects causes unauthentic false-positive detections. Text recognition involves generating candidate character or word region detections, while word identification takes these proposals furthermore infers the words depicted. Besides, the use of our framework for retrieval is further established in a real world application being used to immediately look for

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

through thousands of hours of archived information recording for a user specified text query at very high accuracy.

Despite the fact that investigating OCR blunder modification, we came diagonally a range of examples of OCR text which was also tough for diverse online spell checkers to confirm plus correct. The suggestions generated were awfully distant from the truth for the majority of the words. The chief reasons for this are:

1) Limited glossary is majorly imperfect.
2) Limited realization of language precise set of laws for authenticate appropriately spelled vocabulary.
3) The propose for such spell checkers is based on typing blunder, not on regular OCR n-gram confusions.
4) Ambiguities in suggestions from glossary due to huge number of basic word forms. A number of examples of mistaken OCR output for which the online spelling improvement systems were unsuccessful. We provide accurate suggestions for these words all the way through our proposed method.

2. LITERATURE OVERVIEW

With rising use of digital devices like digital cameras, mobile phones, PDAs, content based image analysis methods have caught serious concentration in the current past years. An image text information mining system is divided into four stages: text detection, text identification, text localization, text mining as well as text enhancement. Among these stages, text detection and text localization are significant. Numerous methods are projected for text detection with text localization problems [3], a number of them achieved high-quality results for specific relevance. The presented methods of text detection with text localization can be generally categorized into two groups: Section based plus connected component (CC) based text detection.

Various issues in addition to challenges specifically for Urdu text based on OCR have been discussed by Zaki et al. [4]. Sharma et al. [5], works on OCR procedure based on Convolutional Neural Network (CNN), their method uses this OCR method to mine information of a student filled in a particular form. This form contains 175 cells. A quantity of these cells are filled with capital alphabets, others are few numbers. It also explains how to identify each cell using the feature and CNN. Their method has achieved a good accuracy for both numerical as well as alphabetical data. Prakash et al. [6], discussed with the OCR challenge for the Telugu, They have created a database for Telugu, a deep learning Technique, in addition to a consumer server solution for the online procedure of the algorithm. Moteelal and Murthy [7] have proposed a framework of continuous content locations that usually deal with different ideas and recognize content. This technique depends on the relevant section. This approach can deal with the content on dispose image plus not bright districts. They build up windows based applications that can track content as well as isolate content in this field. They use Windows telephone of 8.1 rendition, OCR, C#. This approach is a direct and open source. Mol et al. [8], provides a method for text detection as well as recognition in pictures. This method emphasizes Fractional Poisson enhancement for removing Laplacian noise of the input image. Next, the Maximally Stable Extremal Regions (MSER) emphasized on the edge of the preprocessed image. Local filtering is used to filter areas of non-texture in addition to be recognized by an OCR. The results of this method are better than previous methods in terms of Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM) calculations. This method has used standard ICDAR dataset that have nearly real time images. Purkaystha et al. [9], used Convolutional Neural Network (CNN) for Bengali handwritten character detection with satisfactory accuracy. Razik et al. [10], equipped a database of handwritten Bangla numeral named SUST:BNHD. He used deep learning particularly CNN for recognized work.

Besides, various researchers have developed different mobile apps for the purpose of text detection. In recent years, the investigation took into account the words and writings from the camera of cell phones has been organized to translate those words and texts into different languages. One such mobile based approach to convert Thai script into a Malay script has been presented by Aini et al. [11], works efficiently in offline mode. This system consists of three stages of translation, dictionary development, and display result. Yu and Wan [12] uses smart phones to detect Chinese text. The system sets 8 different sizes on 7 common Fonts and 3755 characters. This system generally claims to detect Chinese texts from natural scene. Velumurugan et al. [13], suggest a smart reader based on the problem using Raspberry Pie and use the simulation of OCR using Matlab. Captured images are sent for preprocessing, where functions such as noise removal plus skew correction are performed. The

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

image is light weight in addition to binarization is done. After this, image is approved in a segmentation phase, where the image breaks into characters. Verma et al. [14], introduced an OCR system to change photos in text. Synthetic analysis consists of three stages: Extraction of character, Recognition as well as post processing. In the recognition stage, the template largest correlation is referred to as a character in the image. Rao et al. [15], suggest a Neural Network (NN) technology for the handwritten OCR. Tamil text image are scanned, filtered and pre-processed. Preprocessing consists of character segmentation and picture binarization. The features are extracted and the picture is sent for post processing which converts the picture into text. Text is then fed to a speech engine which generates synthesize Tamil speech as output consists of a multipurpose response via pulsation motors, a new twin material case design in addition to a high resolution mini video cam. Kumari et al. [16], investigated the methods and strategies in achieving the role, in an efficient tracking algorithm, the traditional approaches divide the picture into associate pictures, which are then freely rated. The second class of segmentation parts clearly illustrated by the collected spatial features. The third class is between the primary two and is used with recombination. Farhat et al. [17], introduced a mixed feature extraction method. This article describes a set of zoning, vector crossing techniques. Zoning techniques divide the whole picture into a small picture. Then data attributes (for instance, role density) are calculated for each zone. Vector crossing technique is used to extract the features. The thought of this method is to calculate the number that crosses character pixels in the vector (horizontal and vertical) image. With crossing vector technology, it can be helpful with further technologies. Characters can't be recognized as a vector crossing alone. This method is also useful for other ways. Because it only collects some data features. The cons of these technique are explained by pleasing the arrangement of these two techniques. Rownak et al. [18], proposed a study for efficient segmentation of Bangali characters on paper documents by means of Curvy Scan. Ayyaz et al. [19], planned a hybrid feature mining process, multi-class SVM is used for recognition works, good accuracy for numbers and English characters has been achieved in their study. Singh et al. [20], used SVM to recognize the optical character, they removed the character restrictions using Moore Neighbor's tracing and then applied the chin code. Maitra et al. [21], extracted feature and multifunctional SVM as 5 layer CNN to perform recognition tasks. They achieved good accuracy. Wu et al. [22], proposed CNN's view where they train more than one CNN-based model and use their models for final recognition. Kamble et al. [23], uses a rectangular Histogram Oriented Gradient (HOG) based representation to extract the feature. Then they applied the SVM and the feed Forward ANN Technology. Shekar and Smitha [24] suggested a new perspective of a multi-scale Weber's Local Descriptor (MWLD) present in a text video and visual image. Based on the spatio-temporal association, the shots are recognized in addition to the key frames were mine from the input video. From every frame, local area information was detected using WLD through dissimilar radius and neighborhood association of pixel values therefore obtained intensity improved key frames at numerous scales. These are merging MWLD edge together and then parallel matching numbers were counted using action process. The acquired results were then based on mutual as well as false positive emotions. Finally, an integrated module analysis and aerial procedure are carried out to find a text area for manuscript localization. Naik et al. [25], focused in detecting and extracting textual images. Recognition of region images of natural scenes twisted into a significant area because of its several applications. Text information mining system includes areas in a specified picture, performing the text localization, mining of text component and text recognition by developing OCR. In this research, picture of test will be preprocessed by making the most of alteration of RGB to Gray scale, binarization, technique of boundary detection and technique of noise elimination which is numerical based. The extracted components are used by the SVM classifier which has been trained for detecting areas of manuscript. Nagaraja et al. [26], presented a text detection scheme by means of Raspberry pi this paper presents a complete optical recognition system. The pictures are captured using Raspberry Pi camera module or webcam. Picture preprocessing procedures are applied somewhere the unnecessary noise is detached by applying threshold morphological alteration, discrete cosine, dilation, black hat. The drawing area is pulled out by dragging the box about the desired text and the slant is approved. The ratio between horizontal in addition with vertical is adjusted to remove unnecessary high frequency components. Later than threshold, contours of the picture created using a special openCV function. Image to text conversion has been passed in the speech engine that can change the speech in the text by a dedicated library. This paper uses TTS Festival for this purpose.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Manuel et al. [27], presented a way to find bend features of Malayalam font. The technique includes scanning, degree of curvature etc. The curvelet transform is an enhanced technique for finding curve features than wavelet transform. The character detection is done using a multilayer perceptron. By this method, they can figure vector feature by using small number of coefficients. Still, they face difficulties in the process that there is no built in toolbox to analyze conversion changes. Therefore, using these methods reduces the accuracy. Achanta and Hastie [28] proposed the Telugu using CNN is also motivating. They used 50 fonts to train each size 48-48 images in 4 styles. However, they didn't consider all possible results (for 457 classes) of CNN. Zagoris et al. [29], proposed a process contains manuscript block segmentation, block descriptor extraction in addition to classification by two binary combination SVMs for machine printed documents plus handwritten. Suryawanshi et al. [30], offered a paper based on the nature prosody making in the text to voice for English by the phonetic integration. This method is very complex to trace with collect the word but is used the fewer memory space. Chandran et al. [31], worked on a method that changes text format in verbal communication. This paper incorporate recognized text transformed proper echo signal. OCR technology is executed to achieve system production.

Sharma et al. [32], worked on text to speech alteration for applications in the Android surroundings. The main purpose of their study is to increase Android platform of offline OCR technology. They proposed a system that has been converted into multiple resolutions and uses a language translation program to detect text converted in different languages. Joshi et al. [33], gave the basic idea of the interface between speech to personal computer. Automatic speech recognition (ASR) has been showing an important role in providing people verbal communication technology. This method is classified in additional feature mining plus recognition. Its aim was to study and research how to implement a Neural Network to identify the description of a certain word as option to usual methods. Mangale et al. [34], proposed a system designed for blind users who are interested in the academic field. It makes them audio content of their own choice of autonomous printed matter by NGO's or the Braille book. The entire method mainly includes OCR in which various operations such as thresholding, filtering, segmentation were carried out. Arora et al. [35], offered detection of Gurmukhi text from signboard pictures which were taken through mobile phone camera. The pictures were first binarized and made noiseless. Extracted text was segmented into words by means of projection profiles. Features were extracted using zoning method whereas the characters were classified using SVM method. Geetha et. al. [36] planned the method of extracting and recognizing image manuscript by using the hybrid approach area based and connected component techniques. Mining of picture concerns about extracting association of picture data, implicit information or a few examples which are not stored clearly in the pictures. If the event of this type of text could be robotically recognized, detected as well as segmented, they would be important source of semantics at high point for recovery and indexing. This planned method depends on OCR associated component and the effective edge detector. Images are required to analyze photos based on detecting and extracting substance. The planned plan has developed an efficient way to extract and identify text using morphological operations utilizing MATLAB. Wu et al. [22], projected a handwritten detection technique using Relaxation Convolutional Neural Networks (R-CNN) moreover Alternately Trained Relaxation Convolutional Neural Network (ATR CNN). R-CNN enlarge in parameters in addition to ATR-CNN regularizes the limitation during training phase. They attain an error rate of 0.255% on MNIST arithmetical database by this approach. Kumar et al. [37], projected a Deep Belief Network (DBN) for recognizing handwritten Bangla numbers. Stefano et al. [38], planned a scheme in which feature extracted by Genetic method to recognize handwritten characters. Ranjini et al. [39], uses the various strategies to discuss English text as of the image of blob comic. Extract text from a picture between complex areas in digital image processing. This was an awful procedure of detecting and identifying manuscript from different image styles, complicated backgrounds, unlike sizes and ideals of gray scale. The procedure of extracting text from the comic picture helps in preserving text and presents the text that is printed. Due to designed application in recovery of picture. This work was accessible previously, Japanese content was mined from picture of Manga Comic by means of Blob Extraction technique. In the signify time, extraction of manuscript is completed from dissimilar constraints by utilizing OCR as well as from Japanese language translation of Manga to few multi languages in a quantity of conventional process for allotment the delight of comprehension Manga via Internet. Kasar and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Ramakrishnan [40] planned system for extraction of text lines of arbitrary curvature plus carry into line them parallel. Misra et. al., [41] suggested a technique using Neural Network (NN). Manuscript extraction in addition to recognition were an important step in creating a system with well-organized search plus indexing for a multimedia database. The mandatory aim was to make an unconstrained scheme of picture recovery along with indexing by developing NN. An HSV based approach was adopted for color reduction. Features for ROI specific color plane are removed and then these features can be used in a feature-based classifier that whether there are blocks of non-text or text. The marked block is supply as input to OCR. OCR Outputs in the form of letters of the ASCII frame stored in the database as a future search database. Poignant et al. [42], used flat dilation and erosion to attach the characters of identical string. Certainly, the size of the morphological operator essentially characterizes the dimension of uniform segmented regions. Therefore, huge text areas were prone to over segmentation, while little text regions might be neglected. Permanent size structure elements are just valid for partial spatial resolutions along with a limited range of font range. However, the common fixed size structure elements can't include images of different resolutions and sizes. Saha et al. [43], worked on detecting text using a NN. In their study, the proposed method is divided into two stages. In the first phase, the Neural Network was used to detect English-written text, whereas second step involves the calculation of Elevation Distance Matrix.

3. PROPOSED METHOD

Text detection and recognition has brought many challenges to the research community. Challenges in scenes such as complex backgrounds, small font sizes, inappropriate font styles, stable photos, emerging photos, bright lights have made the task difficult. This study provides a flexible method to extract texts from challenging images by employing basic image processing techniques. The system builds an application that can recognize the English content present in an image captured by using a smart phone and then display the recognized text into editable format on the mobile screen. In addition, the results of proposed app are compared with other Commercial mobile apps. The planned system has used the characters from figure with the assist of ANN for recognition. Following implementation steps are involved in general.

Database gathering: In favor of the detection of text from a hidden image, initially we require to build up our dataset, which contains the description of the character recognized from the analysis to demonstrate the text enclosed by the picture.

Pre Processing: For the period of this step preprocessing schemes such as quantization moreover resizing is used. As the major complexity in the image processing field is the range of the image which depends on the image resolution means number of pixel limited by the picture.

Segmentation with fundamental filtering: After the preprocessing step the subsequent step is to mine the thought from the image that somewhere manuscript is near the image this can be finished by means of Edge Canny detector filter.

3.1 OCR FOR SMART PHONES

First we preprocess the Printed document image. In preprocess we should apply Gray-scale conversion, Binarization, Segmentation. The overall procedure of the proposed method for text recognition and detection is illustrated in Figure 1 which clearly shows that the proposed method comprises of following steps: Gray-scale conversion, Binarization, Segmentation, Text recognition, Display output which are defined in the following sections. Also, the overall block diagram of the proposed method is shown in Figure 1.

3.1.1 Gray-scale conversion

RGB color values are represented in three dimensions x, y, z respectively. The renovation of a color image into a gray image is converting the RGB values 24 bits into gray scale value that consist of 8 bits only. In this way the character image is changed to gray image. Because the images are taken at the space of 1m from the camera, the manuscripts are tiny and in low resolution. Consequently, it is sensible to use gray scale, because it involves in binarization method, Or else, the text may join together or deform. The OCR engine may wrongly identify the characters if it can't differentiate between the characters with the noise. Gray scale images, match up to binary images, and provide the OCR software an improved probability of differentiating between texts with noise which may lead to improved accuracy. For this conversion, we have to calculate the Red,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Green and Blue values then add these values to get the gray scale image (GS) using equation(1) where R, G, and B symbolizes red, green and blue components of the input image.

$$GS = 0.299R + 0.587G + 0.114B \quad (1)$$

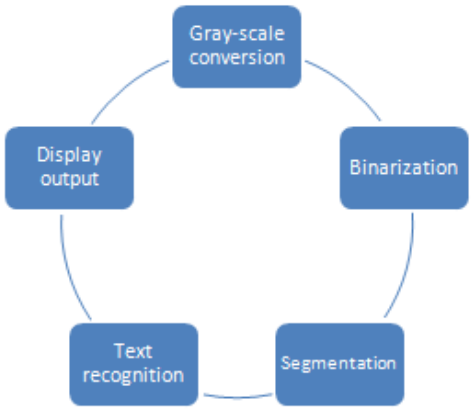

Figure 1: Block Diagram of the Proposed Method

3.1.2 Binarization

Gray image undergoes thresholding and background deduction. After this stage, only the letters of the stroke are left as foreground represented by 1 and 0 for the remaining image background. A character accumulates the Boolean matrix that is used to store either 1's or 0's. The dark pixel is stored as 1's and dark pixel is stored as 0's with the help of the image zoning method. The pseudocode for binarization is given below,

Pseudo Code for Binarization:

1 Take input image for the noise removal process
2 Examine the intensity of a pixel inside the image.
3 IF intensity of a pixel is less than threshold value, Substitute the intensity with 0
  Else
  Substitute the intensity with 1
  End IF
4 Repeat the process for every pixel in the image.

3.1.3 Segmentation

Segmentation is the process of separating characters from an image. When we get characters separately we can perform further processing on it. Segmentation is comprising of three groups: character, word and line segmentation. In this approach, character and line segmentation is applied. After binarization it is assumed that within each restricted section, gray level values of foreground pixels are lower or higher than the standard intensity, connected components with 255 value are extracted as the candidate text components 100 values are not considered.

3.1.4 Text recognition

This project focuses on mobile phones text recognition. For simplicity, we execute this step on a server with Google's open source OCR i.e. Tesseract (http://code.google.com/p/tesseract-ocr/) is used as an optical text recognition engine. Character recognition deals with cropping a particular region from an image, perform the correlation and write in the file. Before you perform the correlation, you need to load an existing saved model template so that it identifies the recognized character that has been found in the template. After achieving the character by character segmentation, save the image of the character. An important field of digital image processing and Artificial Intelligence. It is widely used in data entry from books, bank documents, research papers, office papers or written all types of data. It is also used to save the old and moldy written materials like hand written books, documents etc. Early on versions are mandatory to be programmed with images of every character and worked on one or else multiple fonts at a single time. Therefore, smart systems based on Artificial Neural Network are considered which adjust themselves to the changeable fonts of characters. Characters are in fact the typical alphabets of any language. Character detection systems are chiefly classified into online and offline character detection system. In online character detection system, characters are detected at real time once it is written. Offline character detection can be additional divided into printed and handwritten characters detection. In offline character detection, handwritten characters are scanned from manual paper and transformed into the form of gray image to relate in detection algorithm.

3.1.5 Display output

The recognized text from the above step is shown to the mobile device (Android based) and displayed in the window on the mobile screen.

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

To show the effectiveness of the proposed method, number of experiments have been performed on different challenging images. This study has used almost hundred test images for experimentations which are captured by using Samsung Galaxy J5. All of these images are captured with different angles, different lighting zones, including blurriness with different background colors. The proposed method is estimated with the performance rate of success recognition. This study has calculated the performance rate for each of the steps described in Section III. The results for each of these steps are shown in Table 1, where the accuracy of correctly recognized text is shown for 100 sample images. From Table 1, it can be observed easily that the projected method attains quite satisfactory percentage for accuracy especially for recognition step.

Table 1 Accuracy Computation

S. No.	Steps	No. of Samples Tested	No. of Correctly Tested	Accuracy (%)
1	Gray Scaling	100	98	98%
2	Binarization	100	94	94%
3	Segmentation	100	92	92%
4	Recognition	100	97	97%

Besides the quantitative evaluation, Figure 2 represents the qualitative results for randomly selected eleven test images. To show the efficiency of the proposed android application, this paper compares the proposed app with that of three different Commercial apps, which are named as C1, C2, and C3 in Figure 2.

In Figure 2, the first column from left represents the test images, whereas, the second, third, and fourth column represents the results of the previous commercial apps, C1, C2, and C3, respectively. Finally, the last fifth column exhibits the results of the proposed text detection application. Each of the red box in Figure 2 indicates a false identification of the text. Among all these three commercial apps C1 gives comparatively good results but the only limitation of this app is that it does not work in offline mode. Every time it requires an internet connectivity to carry out the text detection task. However, the proposed method does not require internet connection to perform text recognition in an offline mode along with high accuracy. Although C1 performs well as compare to the other two apps C2 and C3, but fails to give good results when compared with the proposed app. Hence, it can be observed clearly by looking at the results given in Figure.2, that the proposed app provides better text identification as compared to previous commercial apps, C1, C2, and C3.

Though, there is a situation where the proposed method somehow fails to give accurate results as compared to the other apps (C1, C2, and C3) as depicted by the last row of Figure.2. However, it happens due to the reason that the background color of the input image is not homogenous and also the it contains multi-font which becomes the reason of the failure of the proposed app in this particular case which is quite challenging. Observe the result of the proposed app for an input image in second-last row of Figure 2 (tenth image), despite inhomogeneous distribution of the background color, the planned method properly identifies the manuscript as compared to the all three previous apps. To represents the robustness of the proposed app in comparison to the previous three commercial apps specifically in a situation where the input image comprises of very light background along with blurry effect, Figure 3 shows the image where unfortunately all of the three apps fail to recognize the text correctly. Whereas, the proposed app successfully identifies the text mentioned in the input image despite low image quality. The statistical evaluation for this particular case shown in Figure 3 is computed by using the following performance measure of recognition rate.

$$\text{Recognition rate} = \frac{\text{Accurately Recognized Characters}}{\text{Total Characters}} \times 100 \quad (2)$$

In Table 2, the recognition rate of the projected app with that of three commercial apps is shown in terms of correctly recognized characters/alphabets and total number of characters present in an input image used in Figure 3. It can be seen easily in Table 2, proposed app obtains an accuracy of 99% specifically for a challenging case mentioned in Figure 3.

Table 2 Comparative Analysis

S. No.	Apps	Total Characters	Correctly Recognized Characters	Accuracy (%)
1	C1	140	117	83
2	C2	140	124	88
3	C3	140	120	85.7
4	Proposed	140	139	99

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

To the authors' knowledge there is no any system available in the literature that extracts complete text. Furthermore, it has been observed that no solo algorithm is robust for recognition of an unrestrained range of text appearing in the wild. Various methods have been residential to mine text from composite color background pictures. Text eminence is low due to the existence of noise in addition to image decoding and encoding methods. Resolution of manuscript on the display is low, is also a problem as OCR's allow 300dpi resolution manuscript for getting high-quality text recognition results. Mining of text from wild defy over OCR of paper images. Challenges comprises of: worse resolution, indefinite text color, mysterious text sizes, location, orientation, and layout, unrestrained background, Low contrast can effect into impossible difference between background and character strokes.

Recognition of sample in the intersections and distances to describe the rule that was one of the significant feature of our method. The distance outcome filtered with intersection outcome get better precision of detection of characters. In view of the fact that, the method doesn't have segmentation complication, the time in use to identify a character is extremely small. The method mines the same quantity of aspect from all the letters. Consequently, overall accurateness is enhanced. For precision of 91% for a testing dataset of 140 characters confirm that the projected algorithm doesn't show any over appropriate anomalies. The method has only considered the minimum and maximum distances. The intersected distances in between the highest and lowest were unseen. Consequently, the method can be enhanced by calculating these to enhance its accuracy of detection. This method can be more enhanced by increasing the number of directions measured and introducing additional composite characters to the method.

In addition, the method can be bringing in with a different parameter since the summation of intersections of a letter in every direction to raise the precision of the detection.

5. CONCLUSION

This article has presented an android app based on the basic image processing procedure for the recognition of text from the challenging images. The results are evaluated quantitatively as well as qualitatively for various images containing multi font and different backgrounds. The obtained results are also compared with the previous three commercial apps and it is concluded that the proposed app attains a good accuracy by outperforming the previous apps in terms of correctly identifying the text from challenging images.

ACKNOWLEDGEMENT

The authors would like to thanks to Mehran University of Engineering & Technology, Jamshoro, Pakistan, for providing us the resources necessary to conduct this research.

REFRENCES

[1] U. Zaki, M. Memon, D. N. Hakro, K. U. R Khoumbati, M. Hameed, M. A. Zaki and G. Nabi, "Implementation Challenges in Information Retrieval System," SindhUniv. Res. Jour. (Sci. Ser.), vol. 51, no. 2, pp. 339–344, 2019.

[2] H. Li, D. Doermann, and O. Kia, "Automatic Text Detection and Tracking in Digital Video," no. December, 1998.

[3] K. Jung, K. In, and A. K. Jain, "Text information extraction in images and video : a survey," vol. 37, pp. 977–997, 2004.

[4] U. Zaki, D. N. Hakro, K. Khoumbati, M. A. Zaki, and M. Hameed, "Issues & Challenges in Urdu OCR," vol. 3, no. 1, 2019.

[5] A. Sen Sharma, M. A. Mridul, Marium-E-Janna, and M. S. Islam, "A Deep CNN Model For Student Learning Pedagogy Detection Data Collection Using OCR," no. October, 2018.

[6] K. C. Prakash, Y. M. Srikar, G. Trishal, S. Mandal, and S. S. Channappayya, "Optical Character Recognition (OCR) for Telugu : Database , Algorithm and Application," no. January 2018, 2017.

[7] T. Moteelal and V. S. Murthy, "Text detection from a video using frame extraction and text tracking," in International Conference on Intelligent Sustainable Systems (ICISS), 2017, pp. 457–461.

[8] J. Mol, A. Mohammed, and B. S. Mahesh, "Text recognition using poisson filtering and edge enhanced maximally stable extremal regions," in International

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), 2017, pp. 302–306.

[9] B. Purkaystha, T. Datta, and M. S. Islam, "Bengali handwritten character recognition using deep convolutional neural network," in 2017 20th International Conference of Computer and Information Technology (ICCIT), 2017, pp. 1–5.

[10] S. Razik, E. Hossain, S. Ismail, and M. S. Islam, "Sust-bhnd: A database of bangla handwritten numerals," in 2017 IEEE International Conference on Imaging, Vision & Pattern Recognition (icIVPR), 2017, pp. 1–6.

[11] M. Pu, N. Aini, A. Majid, and B. Idrus, "Framework based on Mobile Augmented Reality for Translating Food Menu in Framework based on Mobile Augmented Reality for Translating Food Menu in Thai Language to Malay Language," Int. J. Adv. Sci. Eng. Inf. technolgy, vol. 7, no. February, 2017.

[12] B. Yu and H. Wan, "Chinese Text Detection and Recognition in Natural Scene Using HOG and SVM," no. Itms, pp. 1–5, 2016.

[13] D. Velumurugan, "Hardware Implementation of Smart Reader for Visually Impaired using Raspberry Pi," IJAREEIE, vol. 5, no. 3, pp. 2055–2063, 2016.

[14] A. Verma, S. Arora, and P. Verma, "OCR-OPTICAL CHARACTER RECOGNITION," Int. J. od Adv. Res. Sci. Eng., vol. 5, no. 9, pp. 181–191, 2016.

[15] N. V. Rao, "Optical Character Technique Recognition Algorithms," J. Theor. Appl. Inf. Technol., vol. 83, no. 2, 2016.

[16] K. N. Kumari and M. Reddy, "Image text to Speech conversion using OCR technique in Raspberry pi," IJAREEIE, vol. 5, no. 5, 2016.

[17] A. Farhat et al., "OCR based feature extraction and template matching algorithms for Qatari number plate," in 2016 International Conference on Industrial Informatics and Computer Systems (CIICS), 2016, pp. 1–5.

[18] A. F. Rownak, M. F. Rabby, S. Ismail, and M. S. Islam, "An efficient way for segmentation of Bangla characters in printed document using curved scanning," in 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV), 2016, pp. 938–943.

[19] M. N. Ayyaz, I. Javed, and W. Mahmood, "Handwritten Character Recognition Using Classification with Hybrid Feature Extraction," Pak. J. Engg. Appl. Sci., vol. 10, pp. 57–67, 2012.

[20] D. Singh, M. A. Khan, A. Bansal, and N. Bansal, "An application of SVM in character recognition with chain code," in 2015 Communication, Control and Intelligent Systems (CCIS), 2015, pp. 167–171.

[21] D. Sen Maitra, U. Bhattacharya, and S. K. Parui, "CNN Based Common Approach to Handwritten Character Recognition of Multiple Scripts," 2015.

[22] C. Wu, W. Fan, Y. He, J. Sun, and S. Naoi, "Handwritten Character Recognition by Alternately Trained Relaxation Convolutional Neural Network Handwritten Character Recognition by Alternately Trained Relaxation Convolutional Neural Network," no. September, 2014.

[23] P. M. Kamble and R. S. Hegadi, "Handwritten Marathi Character Recognition Using R-HOG Feature ScienceDirect Handwritten Marathi character recognition using R-HOG Feature," no. December, 2015.

[24] B. H. Shekar and M. L. Smitha, "Text Localization in Video Using Multiscale Weber ' s Local Descriptor," no. April 2015, 2017.

[25] S. Naik and S. Nayak, "Text detection and character extraction in natural scene images," Int. J. Emerg. Technol. Adv. Eng., vol. 5, no. 2, 2015.

[26] L. Nagaraja, D. Nithin, R. S. Nagarjun, V. S. Murthy, and N. M. Anand, "Vision based text recognition using raspberry PI," Int. J. Comput. Appl., vol. 975, p. 8887, 2015.

[27] M. Manuel and M. T. Scholar, "Handwritten Malayalam Character Recognition using Curvelet Transform and ANN," Int. J. Comput. Appl., vol. 121, no. 6, pp. 24–31, 2015.

[28] B. R. Achanta and T. Hastie, "Telugu ocr framework using deep learning," pp. 1–34, 2015.

[29] K. Zagoris, I. Pratikakis, A. Antonacopoulos, B. G. C, and N.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Papamarkos, "Distinction between handwritten and machine-printed text based on the bag of visual words model," Pattern Recognit., pp. 1–12, 2013.

[30] S. D. Suryawanshi, R. R. Itkarkar, and D. T. Mane, "High Quality Text to Speech Synthesizer using Phonetic Integration," no. February 2014, 2018.

[31] P. Chandran, S. Aravind, J. Gopinath, and S. S. Saranya, "Design and Implementation of Speech Generation System using MATLAB," Int. J. Eng. Innov. Technol., vol. 4, no. 6, 2015.

[32] D. Sharma and R. Kanwar, "Text To Speech Conversion With Language Translator Under Android Environment," Publ. Int. J. Emerg. Res. Manag. &Technology Issn 2278–9359, vol. 4, no. 6.

[33] S. C. Joshi and D. A.N.Cheeran, "MATLAB Based Back-Propagation Neural," pp. 10498–10504, 2014.

[34] K. Mangale, H. Mhaske, H. Mhaske, and H. Mhaske, "Printed Text to Audio Converter using OCR," Int. J. Comput. Appl., pp. 27–30, 2015.

[35] S. Arora, D. Sharma, and S. Arora, "Recognition of Gurmukhi Text from Sign Board Images Captured from Mobile Camera," Int. J. Inf. Comput. Technol., vol. 4, no. 17, pp. 1839–1845, 2014.

[36] M. N. Geetha and E. S. Samundeeswari, "Image text extraction and recognition using hybrid approach of region based and connected component methods," Int. J. Eng. Res. Technol., vol. 3, no. 6, pp. 1195–1200, 2014.

[37] M. Kumar, M. K. Jindal, and R. K. Sharma, "Segmentation of Isolated and Touching Characters in Offline Handwritten Gurmukhi Script Recognition," I.J. Inf. Technol. Comput. Sci., no. January 2013, 2014.

[38] C. De Stefano, F. Fontanella, C. Marrocco, and A. S. Freca, "A GA-based feature selection approach with an application to handwritten character recognition," PATTERN Recognit. Lett., 2013.

[39] S. Ranjini and M. Sundaresan, "Extraction and Recognition of Text From Digital English Comic Image Using Median Filter," Int. J. Comput. Sci. Eng., vol. 5, no. 04, pp. 238–244, 2013.

[40] T. Kasar and A. G. Ramakrishnan, "Alignment of Curved Text Strings for Enhanced OCR Readability," Int. J. Comput. Vis. Signal Process., no. January 2013, 2014.

[41] C. Misra, P. . Swain, and J. . Mantri, "Text Extraction and Recognition from Image using Neural Network," Int. J. Comput. Appl., vol. 40, no. 2, 2012.

[42] J. Poignant, L. Besacier, G. Qu, and F. Thollard, "FROM TEXT DETECTION IN VIDEOS TO PERSON IDENTIFICATION."

[43] S. Saha and T. SOM, "Handwritten character recognition by using Neural-network and Euclidean distance metric," Int. J. Comput. Sci. Intell. Comput., vol. 2, no. 1, pp. 1–5, 2010.

Journal of Theoretical and Applied Information Technology
31st October 2019. Vol.97. No 20

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 2: Comparison of the proposed app with three commercial apps

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

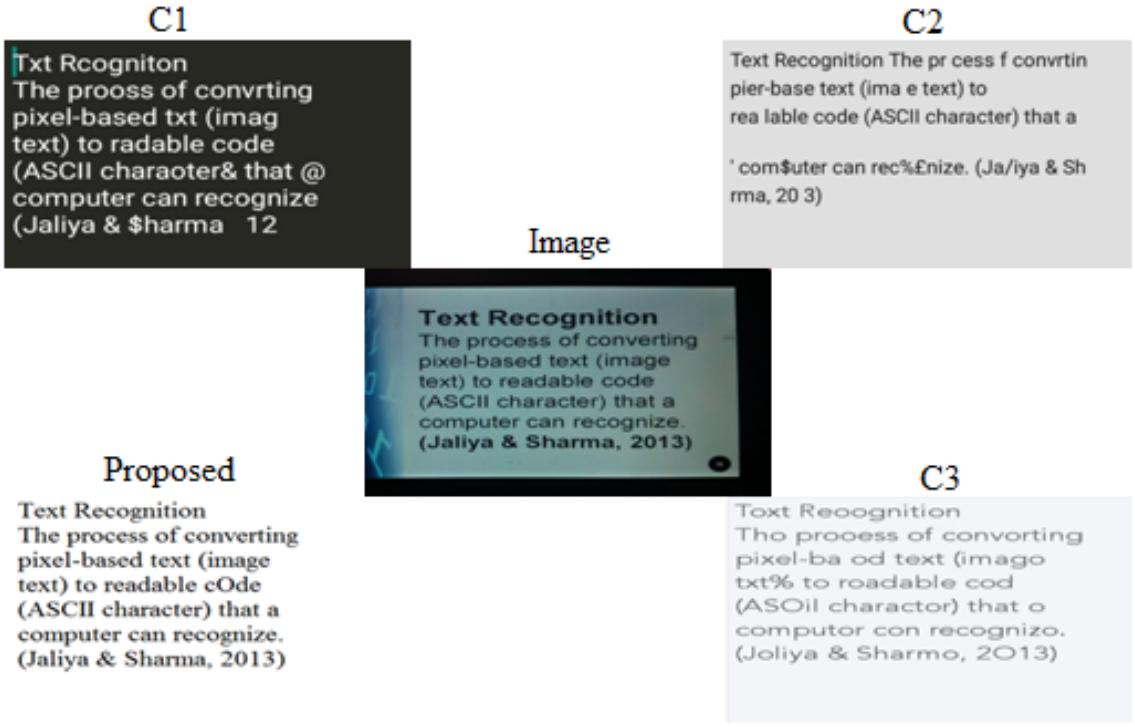

Figure 3: Comparison of the proposed method on a challenging image