



Implementation Challenges in Information Retrieval System

U. ZAKI, M. MEMON*, D. N. HAKRO, K.U.R. KHOUMBATI, M. HAMEED, M. A. ZAKI**, G.NABI

Institute of Information and Communication Technology, University of Sindh, Jamshoro

Received 18th May 2018 and Revised 30st January 2019

Abstract: Natural scenes text recognition has become increasingly attractive in recent years. Several research results have been developed on different languages for the sake of text recognition from natural scenes. However, there is no system for identifying Urdu text from natural scenes. This paper presents a framework steps involved in recognition of Urdu text from landscape images. Furthermore, the paper focus on issues and challenges faced during experimentation and implementation. Challenges includes multi-script text on a single image, different text font with variation in text sizes, light shadow, reflection, text color, text contrast with background, words or ligature missing, blur, low resolution, skew and neon signboard. In the end, we also give a summary of issues and challenges of Urdu Text retrieval system from natural scenes.

Keyword: Signboard, Urdu, Challenges of Urdu Signboard, Information Retrieval System

1. INTRODUCTION

Our world is full of signs and text in the form of images and visual graphic formats (Liao *et al.*, 2014). While, a sign is an object that provides the information of the surrounding in concise format like warning, traffic signal, commercial contents etc embedded in real world, (Park *et al.*, 2008). Text is any information that helps us to understand the objective more easily (Sharma and Prakash, 2012).

Text in images are of two types: scene text and artificial text (Science and Studies, 2015) (Sharma and Prakash, 2012). Artificial text: the text that is artificially place or subtitled on the videos and images; also called "Caption Text" shown in (Fig.1). Scene text: the text that is present within the natural scene images in form of number plates, warning signboard; also called "Graphic Text" as shown in (Fig. 2) (Kaur, 2015).



Fig. 1: Artificial Text



Fig. 2: Scene Text

Furthermore; new technologies enhancing the concept of image processing and retrieval system. The system for retrieving information (Sharma and Prakash, 2012) and text extraction (Sahare and Dhok, 2017) from natural scenes is a challenging task due to the text variation in size, color, resolution, layout, blurring problem, orientation, alignment, shape, texture, background, geometry of text, lighting problem, contrast with background etc.

This study gives a brief review of information retrieval system and presents a detailed of issues and challenges faced during retrieval system of Urdu text from signboard. The study is arranged into several sections. Section 2, Presents the related work. Section 3 discusses the framework of information retrieval system and Section 4 gives detailed about Issues and challenges. Conclusion is summarized in Section 5.

Text Detection and Recognition Reviews Analysis

Baran *et al.*, (2018) presents natural landscape images in a new and effective way to automatically detect text detection and characters. The text detection

¹zurooba6@yahoo.com, ¹dill.nawaz@gmail.com, khalil.khoumbati@gmail.com, MaryamGhori133@gmail.com,

²ahmed_zaki64532@yahoo.com

¹Institute of Business administration, University of Sindh, Jamshoro, Pakistan

²Department of Telecommunication, Mehran University of Engineering & Technology Jamshoro, Pakistan

method used the MSER function. Later, many filters applied to segment the image according to ROI. In last text is passed using the OCR system. Finally, briefly explain the IMCOP document's view.

Tian *et al.*, (2018) a binary-based framework is recommended for detecting and identifying track-based text from embedded subtitled video. This framework consists of three main components that track text, detection, and recognition. In this unified framework, the text is initially tracked. This tracking experience is improved through detection or recognition results. Text detection and recognition has improved by integration of multiple frames. In addition, the database has been released (USTB Vid TEXT). Various experiences on this database have confirmed that this method improves detection and recognition of web video text.

Kaushik and Verma, (2018) provides detailed review of various projects presented for the text recognition process of natural scenes pictures. It also offers different versions of different forms. The work of this review focuses mainly on the relevant ingredients and region-based methods. This paper explains the various benefits and limits of text recognition from nature to nature.

Guzel, (2017) proposed a system for text detection from natural scene by using two different approaches MSER and CSER. Using MSER algorithm, system first take input image to detect region of interest area then obtained text passed to OCR engine for further processing. Using CSER, the system first identifies the external region, classification, and then finds the region of interest and in the last for recognition OCR is applied. The result shows that CSER is more efficient to detect the text region from natural scenes.

Revathi and Balakrishnan developed software that enhance the ability of driver support system by detecting the Indian signboard and gives the message in text or in voice and also help to detect the right direction. The system consists of three phases in which the system first gets the image of Indian signboard then preprocessing. Next phase for the location of sign; done by using two different algorithms namely canny and sobel edge detector to detect edges of signboard. Recognition of traffic sign is the next phase that used template matching method to match the image with that image stored in database. Classification is the last phase. The system used neural network for image classification.

Huang *et al.*, (2014) proposed a new framework of text detection with the Convolutional Neural Network (CNN). To identify text through CNN network advanced features can be learned. This approach uses

both modes sliding windows and MSER. MSER operator reduces the number of windows scanning and increase detection of low quality text. The CNN sliding window applied properly to separate the connections of the characters in the component. It is estimated in the ICDAR 2011 benchmark data set, that the system performs more than 78% in Feature value.

Tam *et al.*, (2003) proposed a system that first detect the signboard and then extract the text information. The system consists of three phases; preprocessing is the first steps. The next phase used geometric transformation with four tie points for undistorted rectangle of signboard. Segmentation is the last phase. For character segmentation they used horizontal and vertical projection technique. An accuracy of 83.38% and 80.85% for uniform and non uniform illumination accuracy of character segmentation was reported.

Signboard Text Detection and Recognition

In the past few years, text tracking and identity has become a major issue. Video captions text detection and identity are also focused on web pages. Photos of natural scenes have been widely used in the field of detecting and extracting text. Different identification techniques can also be used (Kulkarni and Barbadkar, 2017).

The purpose of this proposed procedure is to promote a system that can detect and identify Urdu texts in natural scenarios. The first step in the proposed system is database creation then the next step is preprocessing. Then segmentation and feature extraction done and in the last recognition done. (Fig. 3) shows Flow chart for the proposed procedure.

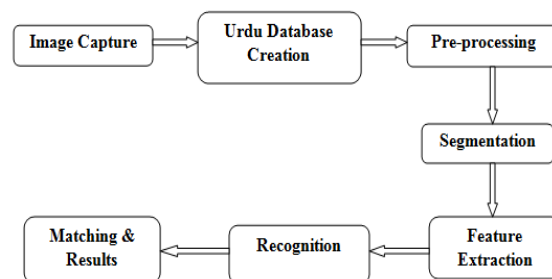


Fig. 3: Flow Chart of Proposed Methodology

Signboard Text Detection Issues and Challenges

Landscape images (natural images) are important for detecting and searching content analysis. The algorithm created for print text recognition does not always work for natural scene text recognition. It is due to problems like low resolution, background color difference, unusual light, and many more (Elakkiya and Safa, 2018). In addition, words may have different

font sizes, designs, and colors because of this features, it is difficult to get text and it is not handled by the current OCR algorithms (Banik *et al.*, 2012).

Text evaluation and its identity are useful for image processing but due to various factors such as light exposure, text on buildings and walls makes it difficult for identification (Ali, Pickering, and Shafi, 2018). New technology of smart devices will increase the value of image processing and retrieval system. The text is distributed into scenes, including the name of the city, street signs, shop name, posters etc. This text is an important indicator to understand the image content and provide useful and important information about the scene (Sharma and Prakash, 2012). This section is described in more detail about the challenges of text detection.

1.1 Textin Multi script

Natural scene images may have a variety of text. The images may have sometimes only Urdu text and sometimes have Urdu script with different type of other text and symbols that may includes digits, English letters or words, lines and different shapes . The text can be written in a multilingual script. Figure 4 shows variety of text with in single image.



Fig. 4: Multi Text

1.2 Varity in Text Size and Font Styles

Review of literature indicates that Urdu has a vast variety of text fonts. When the text is written on boards, banners and on wall the font style and size varied from place to place. (Fig. 5) shows variations in text size and font styles.



Fig. 5: Varity in Text Size and Fonts

1.3 Scene Complexity

Natural scenes full of natural and man-made objects. Sometimes these objects become an obstacle in text detection and recognition. (Fig. 6) shows complexity in scenes.



Fig. 6: Complexity in Scenes

1.4 Lighting problem

If take a picture of the signboard under the sunlight, it can be shown the reflective image of light and shadow as in (Fig 7 (a)) Similarly in (Fig 7 (b)) if the pictures are taken in the night, may be significantly reduced correction in detection and identification.



Fig. 7 (a): Light refelection (b): Darked Image

1.5 Skewness

The text image boundary is basically called skew. Therefore, pay attention to scanning especially for Urdu Text. When these images are segmented it may cause errors. So, we need to detect the correct location for detection. Figure 8 shows different angles of taking pictures.



Fig.8: Different Angles

1.6 Font Background Color and Contrast

Font color and background color are most important for text detection. Both are interlinked with each other. If font color is not proper according to the background so it may cause difficulty in determination of exact text detection and recognition. Hence, it is very much necessary to take a right color as font color according to the background. (Fig. 9) show complexity of text detection because of font and background color.



Fig. 9: Foreground and Background Color Contrast

1.7 Missing Word or Ligature

On the passage of time, signboard may older and their script writings become fading. Due to which sometimes a complete word, ligature and dots maybe vanished so this may also cause difficulty in perfect detection and recognition. (Fig. 10) shows a missing word image



Fig. 10: Missing Word

1.8 Neon Signboard

Neon signboard radiates electric signal to be light by a long luminous gas discharge tube containing dilute neon or other gas. These signboards are useful to see the text from far but a challenging task for text detection especially for Urdu Text. (Fig.11) shows a natural scene neon signboard image.



Fig. 11: Neon Signboard

1.9 Blur and less Resolution Images

Resolution means photo details and clarity. If the image looks fuzzy or blur and does not appear clearly, then it is a low resolution image. Motion during acquisition may result blur image. All other types of fog, such as misleading attention, can worsen image quality. Its impact on text detection is a challenge. (Fig. 12) show a blur image.



Fig.12: Blur and low resolution Image

According to various articles due to sizes, fonts, word options, fog problem, texture, background color, light problems natural landscapes are very difficult to find Urdu text. Due to the nature of Urdu scripts many

scientific challenges are faced during the promotion of information retrieval system. Table 1 presents the overall summary of issues and challenges.

Table 1: Summary of issues and challenges.

Issues and Challenges	Description
Textin Multi script	Images are sometimes available in single script, and sometimes written in more script including Urdu, number, English and other types of text and symbols.
Variety in Text Size and Font Styles	The script has a style to write diagonally and diversity in fonts style and sizes
Scene Complexity	Natural scenes full of object like building, trees, car, poles and many more. These objects cause hindrance in text detection and recognition.
Lighting problem	Lighting shadows and reflection also a barrier between detection and recognition of text .
Skewness	Skew makes identity more difficult.
Font Background Color and Contrast	Colors contrast also a big challenge it makes detection especially for Urdu a difficult task.
Missing Word or Ligature	Since there is no stroke between the characters, the word is not recognized or sometime the complete word is missed.
Neon Signboards	These types of signboards sometimes mixed the text so that text not recognized.
Blur and less Resolution	Blur images with less resolution also mixed the dots orientation for Urdu text.

2. CONCLUSION

Text extraction and its recognition from natural environment has become an active and attractive field in recent years. This paper gives the steps involved in Urdu recognition from graphic text images. In addition, this paper focuses on issues and challenges between experiences and processes. Challenges involved multi-script text, different text fonts, text size, light shadow, reflection, text color, contrast with text, missing words, fade, low resolution, neon signs. Finally, we also highlighted the issues and challenges; lot of caution required for the Urdu text retrieval system from natural text.

REFERENCES:

- Ali, A., M.Pickering, and K.Shafi, (2018). Urdu Natural Scene Character Recognition using Convolutional Neural Networks. *2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR)*, 29–34.
- Banik, P., U.Bhattacharya, and S. K. Parui, (2012). Segmentation of Bangla words in scene images. *Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing - ICVGIP '12*, 1–7.
<https://doi.org/10.1145/2425333.2425403>
- Baran, R., P.Partila, and R.Wilk, (2018). Automated Text Detection and Character Recognition in Natural Scenes Based on Local Image Features and Contour Processing Techniques. In W. Karwowski and T. Ahram

(Eds.), *Intelligent Human Systems Integration* 42–48. Cham: Springer International Publishing.

U.Elakkiya, and M. Safa, (2018). Text Detection in Natural Scene Images, (March).

Guzel, M. S. (2017). A Novel Framework for Text Recognition in Street View Images. *International Journal of Intelligent Systems and Applications in Engineering*, 5(3), 140-144. *International Journal of Intelligent Systems and Applications in Engineering*, 5, 140--144. <https://doi.org/10.1039/b000000x>

Huang, W., Y.Qiao, and X. Tang, (2014). Robust Scene Text Detection with Convolution Neural Network Induced MSER Trees. In D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars (Eds.), *Computer Vision -- ECCV 2014* (497–511). Cham: Springer International Publishing.

Kaur, T. (2015). Text Detection and Recognition from Natural Scene, 4(7), 3211–3216.

Kaushik, D., and V. S. Verma, (2018). Review on Text Recognition in Natural Scene Images. In B. Panda, S. Sharma, and U. Batra (Eds.), *Innovations in Computational Intelligence : Best Selected Papers of the Third International Conference on REDSET* 29–43. Singapore: Springer Singapore.
https://doi.org/10.1007/978-981-10-4555-4_3

Kulkarni, C. R., and A. B. Barbadekar, (2017). Text Detection and Recognition: A Review. *International*

- Research Journal of Engineering and Technology (IRJET)*, 4(6), 179–185.
- Liao, M., B. Shi, X. Bai, X. Wang, and W. Liu, (2014). TextBoxes : A Fast Text Detector with a Single Deep Neural Network, 4161–4167.
- Park, J., T. N. Dinh, and G. Lee, (2008). Binarization of Text Region based on Fuzzy Clustering and Histogram Distribution in Signboards, 2(7), 2319–2324.
- Revathi, G. (n.d.). Indian Sign Board Detection and Classification Using Image Processing Techniques, 192–196.
- Sahare, P., and S. B. Dhok, (2017). Review of Text Extraction Algorithms for Scene- text and Document Images. *IETE Technical Review*, 34(2), 144–164. <https://doi.org/10.1080/02564602.2016.1160805>
- Science, C., and M. Studies, (2015). Text Detection and Extraction from Natural Scene: A Survey, 331–336.
- Sharma, S., and J. Prakash, (2012). A Survey of Image to Text Detection Methodology. *International Journal of Advanced Research in*, 3(1), 46–49. Retrieved from <http://www.akgec.org/journals/jan-june 12/10-San.pdf>
- Tam, A., H. Shen, J. Liu, X. Tang, and H. Kong, (n.d.). Quadrilateral Signboard Detection And Text Extraction.
- Tian, S., X. Yin, S. Member, Y. Su, and H. Hao, (2018). A Unified Framework for Tracking Based Text Detection and Recognition from Web Videos, 40(3), 542–554. <https://doi.org/10.1109/TPAMI.2017.2692763>