An Analysis of Optical Character Recognition (OCR) Methods

Nabeel Ashraf [#1], Syed Yasser Arafat[*2], Muhammad Javed Iqbal[*3]
Department of Computer Science and Information Technology, Mirpur University of Science and Technology (MUST)

Department of Computer Science, University of Engineering and Technology Taxila (UET)
[1]writenabeel@gmail.com,
[2]syed.yasser.arafat@gmail.com
[3]javed1797@hotmail.com

ABSTRACT: *This survey paper presents a comprehensive study of Urdu Optical Character Recognition (OCR) methodologies. The main focus of the study is detail investigation of the techniques used to recognize the Nastaliq, Naskh and other similar scripts fonts. These script fonts are used to write Urdu, Arabic, Pashto and Sindhi etc. languages. Several methods of text recognition and classification of Urdu like cursive scripts are discussed. The survey contains the comparison and description of each method in a brief way which identifies handwritten, printed and online text recognition as well. For each optical character recognition (OCR) the phases of pre-processing, segmentation, feature extraction, classification and finally recognition are discussed. After the comprehensive analysis of all methodologies critics and future work in Urdu cursive scripts, i.e. Naskh and Nastaliq scripts are also proposed.*

1. Introduction

With the advancement in technologies, OCR development has become one of the important need of the time for cursive script languages, especially Urdu text/script. OCR system helps to digitally recognize and classify the current paper records, books, magazines, daily papers, and notes etc. It has additionally become an active research area in the course of the recent years. The need for productive optical character recognizers has become an essential part to change the digital pictures into content. OCR is the most explored design arrangement issues. Several OCR's of the languages have gained mature level in text recognition and classification but there are still some subcontinent languages which needs further improvement like Urdu, Pashto, Sindhi etc [1].

The national language of Pakistan is Urdu. According to census of 2017 whose population about 210–220 million [1]. The

importance of Urdu language can be realized from the fact that there are approximately 70 million speakers of Urdu are also in India as well[2]. Also the urdu speaking community of Pakistan and India which live in the Middle East, Europe, USA and other countries of the world. Several languages like Arabic, Pashto, and Farsi have also followed the Naskh and Nastaliq like scripts.Urdu language is popular due to its vast poetic and classical literature. Several Optical Character Recognition (OCR) is designed to recognize the Urdu language not only to serve the native speakers but also to implement it practically for mobiles, tablets and several operating systems for users [8] . Moreover, OCR systems also help to digitize the documents, understand written books, and scan printed text, newspapers, poetry as well [2].
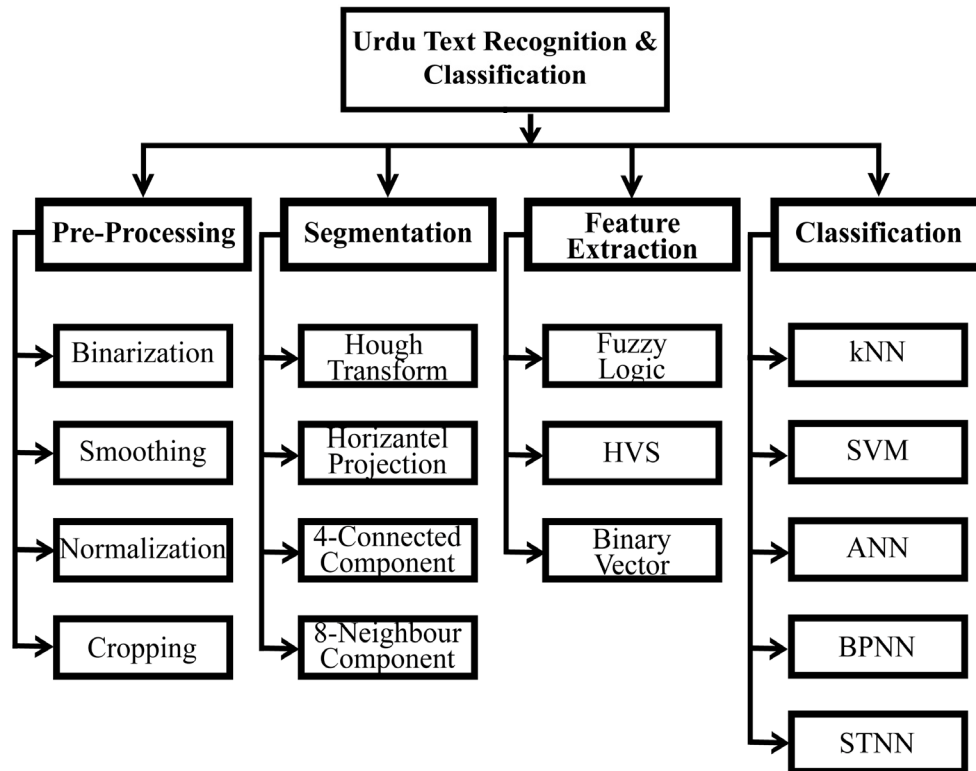

Figure 1. Taxonomy of Urdu Text Recognition & Classification

Following are the major steps of Urdu text recognition and classification OCR shown in figure 1 .

1) Pre-processing:

The stage of preprocessing is important for eliminating unnecessary details from the data. Preprocessing is the most important task to make the image suitable format by making it noise free and clean. Several different approaches of preprocessing are used in Urdu language OCR .For example [5] in first step convert image into gray scale and then its height normalized up to a specific level by keeping aspect ratio constant. After that the extracted pixels are passed to MDLSTM network.

2) Segmentation:

Segmentation, the process of separating one word from another. It separate the several letters of a word and differentiate it into parts. In most cases, we can easily separate the letters there is always space between them.

3) Feature extraction:

In pattern recognition feature extraction is very important phase of OCR extracting unique information for recognition step. Feature extraction involves the extraction of particular and unique patterns from the input image. It helps to enhance the identification power along with reducing the data for classification. The success of a classifier depends on the feature extraction. [2] Multi-dimensional long short term memory network (MDLSTM) for feature extraction is used.[29] used PCA for feature extraction of urdu text.

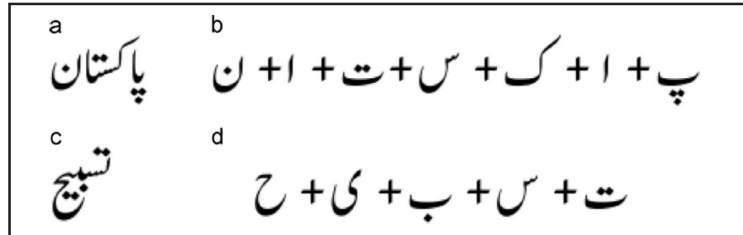Figure 2. Showing different parts of Urdu words called Ligatures

Figure 3. Showing different parts of words [1]

4) Recognition and Classification Approaches:

It's a challenging task to choose a suitable recognition model due to complex nature of Urdu script. Several work is done in past but still there is a big research gap [1] [8]. With the passage of time and advancement in technology classifiers are becoming more powerful and accurate so there is sure hope for better OCR systems in Urdu in future. Several Neural Networks are used by researchers in previous years. [2] implemented a hybrid approach by combining convolutional neural network (CNN) and [26] used 2 dimensional long short term memory networks (2DLSTM) as learning classifiers. [25] proposed convolutional recurrent neural network.[29] implemented Euclidean distance for classification.
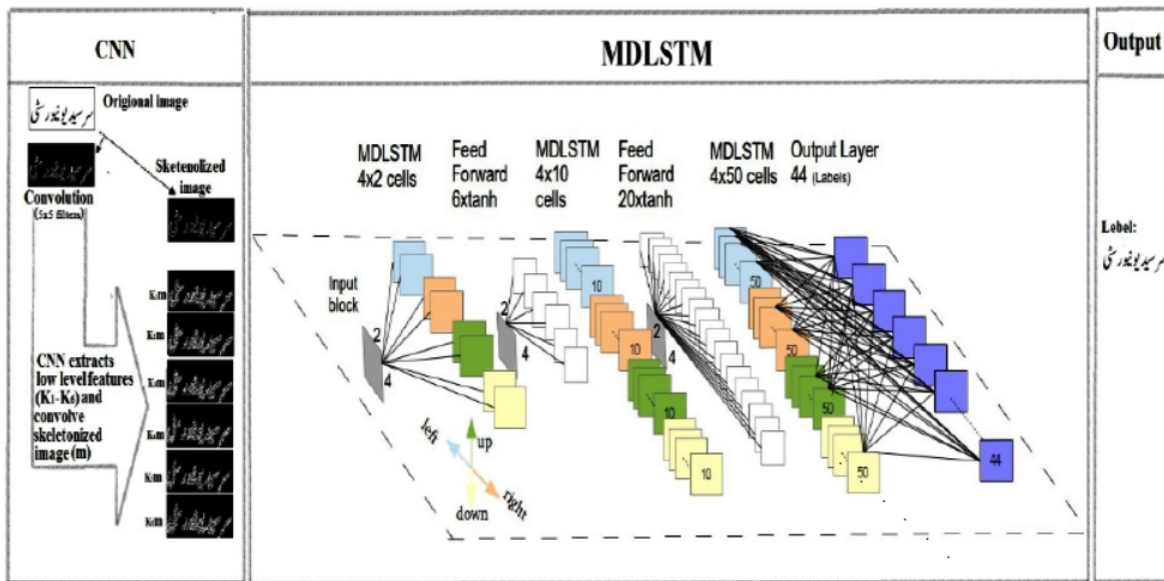

Figure 4. Showing a model of classifiers used of Urdu OCR[2]

This paper provides a detailed survey on Urdu OCR systems in literature review of the papers. Fresh work of some current Urdu researcher is explored briefly. It also contains conclusion and future research gaps identified during the review. This literature survey will depict a comprehensive reflection of modern work on Urdu language.

2. Literature Review

Saeeda Naz et al. [2] proposed an Urdu Nastaliq font recognition using a hybrid methodology of Convolutional Neural Network and Multi-dimensional Long Short term Neural Network. It is used for feature learning and classification of cursive Urdu Nastaliq script. By implementing CNN in the first layer for feature extraction after that those selected features are passed towards Multi-dimensional Long short term Neural Network to extract contextual features and training. Experiment performed on publically available Urdu Printed Text line Image (UPTI) dataset. A recognition Rate of up to 98.12% results on the UPTI dataset [2].

Anwar Ali et al. [3] proposed an optical character recognition system proposed to identify the handwritten Sindhi numeral and expressions excluding any external input device like keyboard or memory.

Following steps are used in proposed OCR process: image acquisition, pre-processing, segmentation, Feature Extraction, and classification. For acquisition of the image, they used digital cameras and scanners to get handwritten and computer -generated fonts for training. Pre-processing is used to achieve by applying several techniques like region filling, binarization based on a threshold, smoothing and to remove noise and improve the quality of images for better results. After pre-processing the segmentation is achieved by matching all the nearest neighbor values from the image. After the segmentation, next step is to extract features which are done by using a binary matrix whose value is 42x24 which is often used in OCR feature extractions. The last phase of this OCR system is classification which is done by implementing support vector machine (SVM) and kNN on binary images.

Safia Shabbir et al. [4] proposed an OCR to recognize printed Urdu Nastaliq font. To reduce the complexity of the proposed OCR system implementation of a segmentation free approach is used in which ligature is the basic unit of recognition. Each ligature represents a set of feature which are clustered accordingly using clustering. Hidden Markov Model has used these clusters ligature for training. After the training, HMM used trained data set to identify the complete words. The average recognition rate of the proposed system is 92%.Morever several diacritics are ignored in this work and no pre-processing technique is used for old and damaged printed text.

The acknowledgment of Arabic content and its subsidiaries, for example, Urdu, Persian, Pashto and so forth is a troublesome assignment because of the intricacy of content. Especially, Urdu content recognition is more troublesome because of its Nastaliq composing style. Nastaliq composing style acquires complex calligraphic nature, which presents real issues to the detect and recognise urdu text [5].
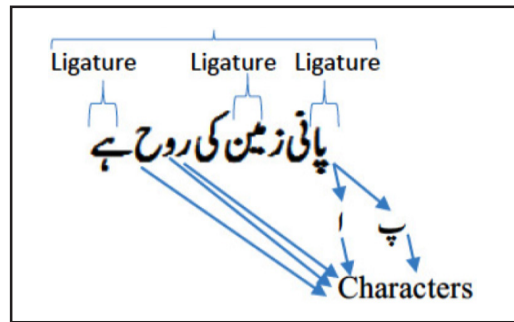

Figure 5. Image showing the difference between character and ligature [26]

Proposed the recognition of Urdu Nastaliq font is done by implementing Multi-dimensional Long Short-Term Memory (MDLSTM) Recurrent Neural Network. For experiment Urdu Printed Text (UPTI) data set is used which contains 10000 images furthermore, UPTI is divided into testing, validation, and training sets. Tests demonstrate that MDLSTM achieved classification accuracy of 98% for the unconstrained Urdu Nastaliq printed content, which essentially beats the best in class systems [5].

On the basis of literature review, several comparisons are made between the studied methodologies with respect to accuracy and other parameters. A detailed summary of the contributions of the various authors in text detection and recognition of cursive

script languages like Urdu is discussed with the help of tables as well.

Study	Year	Dataset	Recognition unit	Accuracy
Saeeda Naz et al. [5]	2016	UPTI	Character	98%
Safia Shabbir et al [4]	2016	Custom	Ligatures	93.7%
Baoguang Shi et al [25]	2017	UPTI	Character	98.12%
Saeeda Naz [26]	2016	UPTI	Character	93.39%
Israr Ud Din [6]	2017	UPTI	Ligatures	92%
K. Khan [27]	2015	Custom	Character	92.06%
Engr. Wafa et al. [7]	2015	Custom	Character	92%
Tofik Ali [8]	2016	Custom	Character	95%
Ibrar A-hmad [9]	2017	Custom	Ligatures	93%-96%

Table 1. Comparison of Recognition accuracy with notable studies

An optical character recognition technique is suggested for printed Urdu Nastaliq textual style in which the basic unit of recognition is ligatures [6]. The proposed method used statistical strategy by implementing Hidden Markov Models (HMM) for feature extraction. Urdu Printed Text images (UPTI) dataset is used to train the classifier. Ligatures from the database for investigation. The classifier works by separating the ligatures form the content lines then the essential parts are grouped according to matching behaviors.

Ligature is independently separated using HMM. Similar groups of words are separated and used accordingly. After the notification, the essential and optional ligatures are independently perceived and later related together by the arrangements of trained data it perceives the entire ligature.

A feature matching technique is implemented on the textual style Jameel Noori Nastaliq font. The proposed method works by converting the whole script into changed editable printed Unicode's [7]. There is a need of adding more words questions with the addition of new databases along with several new Urdu fonts. A different approach of converting Urdu text images into machine-editable code is presented by Tofik Ali [8]. Given UOCR system which converts images into digitally editable text files of Urdu printed Nastaliq scripts by performing step by step operations like binarization, segmentation, feature extraction and classification on that document. It finally produces digitally Urdu editable text file. A SIFT and SURF methods for computing and feature extractions approach. The proposed OCR system works only for digital Urdu images it cannot implement on handwritten text so in future work, it will be a great contribution to enhance this OCR for handwritten.

A denoising autoencoders with the implementation of softmax layer is proposed by Ibrar Ahmad et al. [9]. The implemented technique is used to extract features from crude and distorted pixels and classify it accordingly. Denoising auto encoders are trained from random ligatures from UPTI database which is publicly available. The proposed methodology tested and trained UPTI data set.

Zia Ul Rehman has displayed a novel structure for sentiment analysis in Urdu remarks [10]. The lexicon based engineering works by doling out polarities to the tokens created by Urdu sentence. The lexicon has 7335 sections; 2607 negative and 4728 positives. The general polarity of the sentence is a summation of every single separate term weight. Examination on the informational index of one hundred and twenty-four Urdu remarks from different Urdu site is performed to check the viability of proposed framework.

Generally, the performance of ligatures and line based Urdu language OCR systems are more efficient in working when compared with characters-based OCR systems. [11]. this paper implemented a segmentation approach of Urdu Nastaliq font using content

Study	Dataset	Classifiers	Accuracy
Safia Shabbir et al [4]	Custom	HMM	93.7%
Baoguang Shi et al [25]	UPTI	CNN	98.12%
Saeeda Naz [26]	UPTI	2DLSTM	93.39%
Israr Ud Din [6]	UPTI	HMM	92%
K. Khan [27]	Custom	PCS & DTA j-48	92.06%
Engr. Wafa Qaiser Khan [7]	Custom	PFM	92%
Tofik Ali [8]	Custom	Binarization	95%
Ibrar Ahmad [9]	Custom	SDA	93%-96%

Table 2. Comparison of Recognition classifiers with notable studies

images to line along with ligatures as well. Traditional segmentation methods are effectively utilized by implementing line split to overcome the issues. The proposed algorithm distinguish the related segments from particular line contents and arrange them into auxiliary and specific classes accordingly with respect to features.
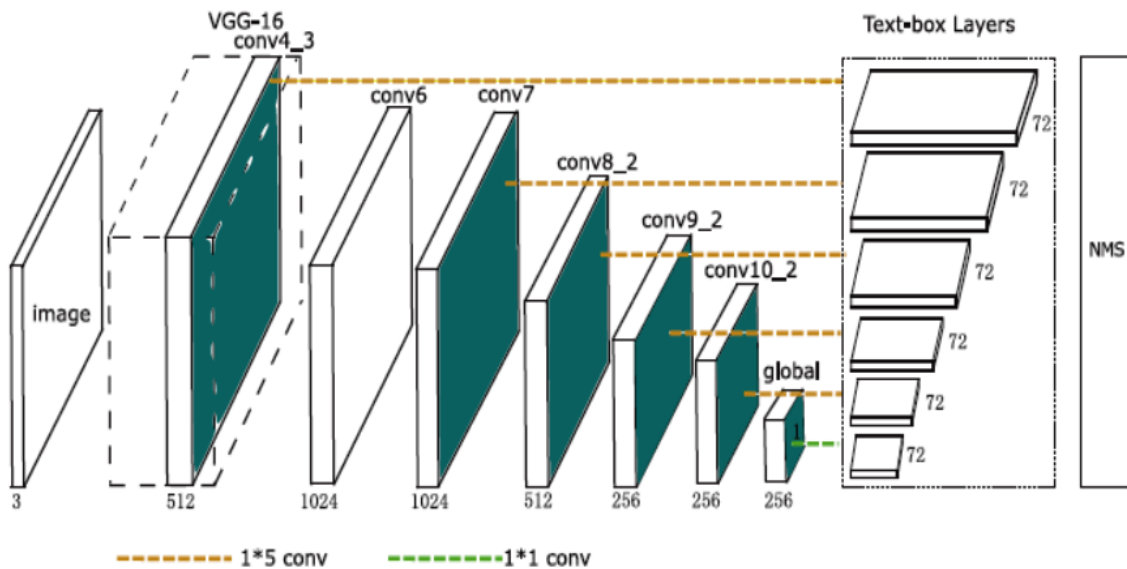

Figure 6. A 28-layer fully convolutional network [28]

Qurat ul Ain Akram et al. [12] a structure for the improvement of font size autonomous OCR is given. The exhibited method is font size independent, however, to enhance the recognition comes about of as often as possible utilized font size text and to make an ongoing framework, isolate recognizer for 14 and 16 text dimensions are created. The Nearest Neighbours interjection procedure is used to resize the pictures at different text dimension and sizes. Israr Uddin et al. [13] introduces a successful sans segmentation also, scale-invariant strategy for recognition of Urdu ligatures in Nastaliq font. The proposed method depends on isolating the primary collection of ligatures from the auxiliary parts and preparing a different hidden Markov model for each. Features catching projection, concavity and bend data of ligatures are extricated utilizing appropriate to-left sliding windows and are encouraged to the models for preparing. The framework prepared and assessed on a sum of more than 2,000 much of the time.

There is rapid development in the field of OCR but still some Urdu fonts like Nastaliq need attention in perspective of OCR. [14]

exhibits a logical preparation on the acknowledgment methods including division on the Nastaliq. Regardless of the differing OCR construct investigates in light of the Nastaliq content, there are as yet unsettled issues in the acknowledgment techniques. The main issue is the absence of a standard database containing an immense scope of letters, digits, and words. Due to limited variations of dataset for Nastaliq content further work needed in dataset. Enhancing the database with a wealth of information for inspecting would add to the compelling upgrade of the ongoing examination in this field. Moreover the calligraphic nature of text also exhibits a challenging task for OCR accuracy and results.
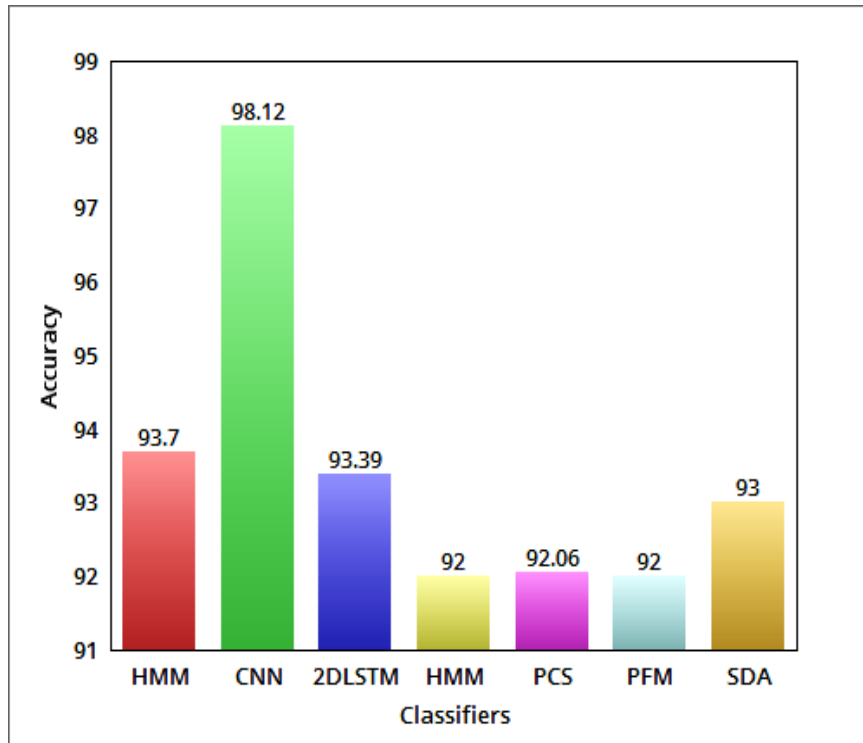

Figure 7. Shows the comparison of Urdu script classifiers

Anshul Mittal et al. [15] have proposed a generic approach to detection and recognition of text from video outlines. Sub-pixel matching is used to deal with pictures of high resolution. For the features extraction from the picture, Support Vector Machine (SVM) is used as a classifier for multi-arranged content/non-content and Recurrent Neural Network (RNN) based classifier is utilized for recognition of content characters. Safdar Zaman et al. [16] presented a method that works by changing the image into signal and after that follow the similar pattern of the image in respective signal format. The approach works under five stages which are shows specifically. Feed Forward Neural Network (FFNN) is utilized for classification.

Zumra Malik et al. [17] propose a video content recognition framework that considers the text in each script as a one of a kind surface. Several text are identified in videos using an artificial neural network and also to figure out how to recognize unique scripts. The framework assessed on video content squares in five unique contents (Arabic, English, Urdu, Hindi, and Chinese) announced promising recognition rates. Notwithstanding the execution of individual textural highlights, distinctive blends of surface measures were explored which acknowledged intriguing outcomes. [18] Presented a detailed survey to discuss the pro and cons of Urdu Nastaliq and Naskh font recognition and classification challenges. Survey tended to the relative complexities involve in baseline detection method used in OCR system for Nastaliq and Naskh cursive contents. Moreover, several issues and challenges recognition strategies are explored. The goal of the paper is to demonstrate that some urdu fonts like Nastaliq have less consideration of scientists in the field of OCR. There is still a need for great contribution in detection and recognition techniques used for Urdu OCR systems.

Presented [19] for the extraction of twisted content line the line detection technique on the basis of ridge base implementation is used. The conclusion of the paper is that the ridge-based text-line extraction strategy is a nonspecific content line discovering
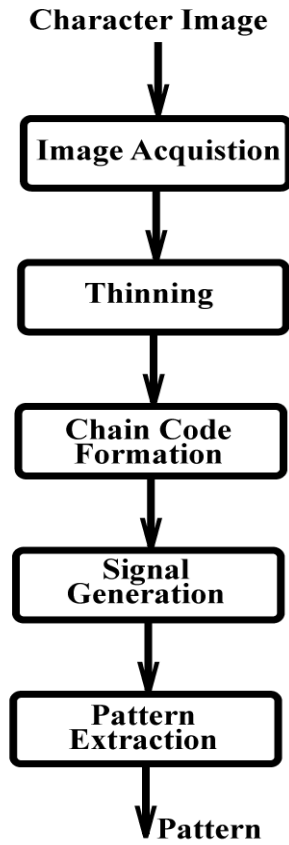
Character Image

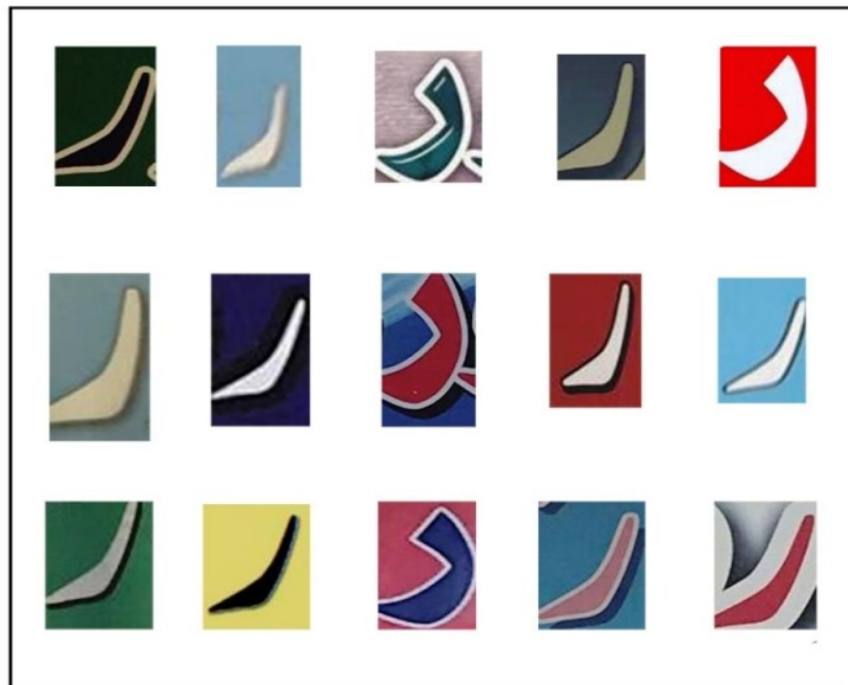

Figure 8. Five phases of OCR [16]

Figure 9. Variation of same Urdu letter in different fonts

technique. It can be vigorously connected to basic and in addition complex report pictures containing diverse testing issues, for example, skewed or potentially twisted content lines and additionally covering text lines, free-form written by hand message lines, sporadic format, clamor also, mutilations. For a broad execution assessment, we have contrasted the edge based strategy and a few domain-specific cutting-edge techniques for countless accessible datasets that have a place with both straightforward and complex sorts of archive pictures.

U. Pal et al. [20] depicts a framework of OCR for printed Urdu content. The recognition exactness of their model is promising; however, more work is required. Character segmentation technique ought to be enhanced to deal with a bigger assortment of characters that happen regularly in pictures got from second-rate quality archives. We additionally need to recognize compound characters to make it a total OCR framework. In general. In future work they are planning to work on the images which contains various text styles and sizes.

Sobia T. et al. [21] suggested Hidden Markov Model (HMM) separated worldwide to recognize the ligature base recognition. Proposed methodology has a capacity to perform acknowledgment effortlessly and productivity. Language independent preparing and acknowledgment procedure was a noteworthy reason for utilizing HMM in creating acknowledgment innovation. Jameel et al. [22] give a survey of distributed research work in phase of Urdu manually written character acknowledgment framework. A few feature extraction, reprocessing, division and recognition techniques are utilized by the scientists and revealed distinctive precision levels yet the utilization of BSpline Bend estimate in the mix with Feedforward Neural Network as classifier has not been made for written by hand Urdu characters recognition. In the ebb and flow look into work, the utilization of this strategy has been proposed to improve the exactness and effectiveness of Urdu Handwritten Character Recognition considering bends like the structure of Urdu characters and consistent and relative invariant nature of B-splines and very versatile capacity of Neural Networks.

Jameel et al. [23] discovered that different attributes, for example, feature extraction, pre-processing, division and acknowledgment procedures have been utilized and revealed diverse precision levels however the utilization of B-Spline Curve has not been found for manually written Urdu content to shape the feature vector in spite of their strength. We have proposed a procedure in such manner with a specific end goal to improve the exactness and proficiency of Urdu OCR.

Saad bin Ahmed [24] proposed new dataset for manually written Urdu dialect named Urdu-Nastaliq transcribed dataset. Being a cursive nature, Urdu Nastaliq text style has no standard dataset accessible publicly. The fundamental thought process of planning UNHD disconnected database is to order Urdu content and make it accessible to examine free of cost. The execution of classifier can be assessed on 2-dimensional BLSTM. Other future assignments may incorporate author identification, apply distinctive element extraction approaches, and apply diverse classifiers to perceive the content and word recognition with the assistance of lexicon and dialect demonstrating.

3. Conclusion

Because of various difficulties, an efficient Urdu content OCR is still a long way and a great deal is yet to be done in this association. Variation in the style of Nastaliq font which makes the undertaking significantly all the more difficult. With the pictures of the printed content, the circumstance is much additionally promising as is apparent by endeavors. In any case, analysts have not been that eager to the extent handwritten content is concerned, regardless of whether disconnected or on the web. The most part of the research deals with ligature based recognition of Urdu contents there is still more efforts needed for character recognition approaches.

Additionally, Urdu line content has briefly explored the works focusing this area is rare. Urdu page acknowledgment is additionally a disregarded zone. Recently work approaches deal with incorporate dictionaries along with composing style limitations, furthermore, and lesser precision with Nastaliq. Generally speaking, the circumstance with Urdu dialect isn't as awful as with Pashto and Sindhi where text recognition is at its simple stage, at first sight in any event. From the point of view without bounds work/bearings, there is a need to create calculations that can fuse boundless or vast vocabulary; a definitive being the ability of multilingual cursive content character acknowledgment. Quickly depicted the properties of Urdu-like content dialects. Purposed methods share a basic character set and simple composing style. For the efficient detection and recognition of baseline estimation and ligature segmentation, there is a need of more critical algorithms that may, in the long run, demonstrate accommodating in handling the confined lexicon issue.

4. Future Recommendations

After the analysis of Urdu OCR literature, we found that there are many researchers interested work in Urdu OCR systems. For example, to best of our knowledge, there is no work done in Urdu-script recognition from markets signboards, street boards and shop names in Pakistan. There is need of Urdu text scanner which can read urdu books and generate its editable files, it will help to save many urdu books. There is also a need of the system which helps the foreigner to understand Urdu by generating its translation in English accordingly. One of the big contribution in urdu language is to make a model which can understand whole urdu famous markets of Pakistan and its translated results can be corporates with Google street view. One of the most active field of urdu research is handwritten urdu detection and recognition it also needs a great attention to overcome its research gaps. This can be a great contribution to the Urdu language if we design a database containing complete annotations it will help to all upcoming urdu researchers in their work. Urdu online screen texts recognition and classification is an important field for future researchers. For example shop names board's images for the Urdu language , traffic sign boards urdu dataset, and other screen printing datasets. This dataset will publicly available for many Urdu language researchers. Fundamentally, the challenges with Urdu, and other scripture contents are complex and committed collective efforts and consistency to acquire good results so it's still a challenging task for the research community.

References

[1] Naz, Saeeda. (2014). The optical character recognition of Urdu-like cursive scripts. *Pattern Recognition* 47 (3) 1229-1248.

[2] Naz, Saeeda. (2014). Urdu Nastaliq recognition using convolutional–recursive deep learning. *Neurocomputing* 243, 80-87.

[3] Sanjrani, Anwar Ali. (2016). Handwritten Optical Character Recognition system for Sindhi numerals. *Computing, In*: Electronic and Electrical Engineering (ICE Cube), 2016 International Conference on. IEEE, 2016.

[4] Shabbir, Safia. (2016). Optical character recognition system for Urdu words in nastaliq font.

[5] Naz, Saeeda. (2010). Urdu Nasta'liq text recognition using implicit segmentation based on multi-dimensional long short term memory neural networks. SpringerPlus 5 (1) 2010.

[6] Din, Israr Ud. (2017). Segmentation-free optical character recognition for printed Urdu text. *EURASIP Journal on Image and Video Processing* 1.

[7] Khan, Wafa Qaiser., Reema Qaiser Khan. (2015). Urdu optical character recognition technique using point feature matching; a generic approach. *In*: *Information and Communication Technologies* (ICICT), 2015 International Conference on. IEEE, 2015.

[8] Ali, Toflk., Tauseef Ahmad, Mohd Imran. (2016). UOCR: a ligature based approach for an Urdu OCR system. Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on. IEEE, 2016.

[9] Ahmad, Ibrar. (2017). Offline Urdu Nastaleeq optical character recognition based on stacked denoising autoencoder. China Communications 14.1, 146-157.

[10] Rehman, Zia Ul., Imran Sarwar Bajwa. (2016). Lexicon-based sentiment analysis for Urdu language. *In*: Innovative Computing Technology (INTECH), 2016 Sixth International Conference on. IEEE, 2016

[11] Ahmad, Ibrar. (2017). Line and ligature segmentation of Urdu Nastaleeq text. IEEE Access 5, 10924-10940.

[12] Akram, Qurat Ul Ain., Sarmad Hussain. (2017). Ligature-based font size independent OCR for Noori Nastalique writing style. Arabic Script Analysis and Recognition (ASAR), *In*: 2017 1st International Workshop on. IEEE, 2017.

[13] Khattak, Israr Uddin. (2015). Recognition of Urdu ligatures-a holistic approach. *Document Analysis and Recognition* (ICDAR), 2015 13th International Conference on. IEEE, 2015.

[14] Sharma, Harmohan., Dharam Veer Sharma. State of the art in Nastaleeq Script Recognition.

[15] Mittal, Anshul. Rotation and script independent text detection from video frames using sub pixel mapping. *Journal of Visual Communication and Image Representation* 46 (2017), 187-198.

[16] Zaman, Safdar., Kanwal Anwar, Riaz Khan. (2016). Image character through signal and pattern formation. *Learning and Technology Conference* (L&T), 2016 13th. IEEE, 2016.

[17] Malik, Zumral. (2015). Video Script Identification Using a Combination of Textural Features. *In*: Signal-Image Technology & Internet-Based Systems (SITIS), 2015 11th International Conference on. IEEE, 2015.

[18] Naz, Saeeda. (2016). Challenges in baseline detection of cursive script languages.*In*: Science and Information Conference (SAI), 2013. IEEE, 2013.

[19] Bukhari, Syed Saqib., Faisal Shafait, Thomas M. Breuel. (2013). Towards generic text-line extraction. *Document Analysis and Recognition (ICDAR)*, *In*: 2013 12th International Conference on. IEEE, 2013.

[20] Pal. U., Anirban Sarkar. (2003). Recognition of printed Urdu script. Document Analysis and Recognition. *In*: Proceedings. Seventh International Conference on. IEEE, 2003.

[21] Javed, Sobia, T. (2010). Segmentation free nastalique urdu ocr. World Academy of Science, Engineering and Technology 46, 456-461.

[22] Jameel, Mohd., Sanjay Kumar, Abdul Karim. (2017). A Review on Recognition of Handwritten Urdu Characters Using Neural Networks. *International Journal* 8 (9).

[23] Jameel, Mohd., Sanjay Kumar. (2017). Offline Recognition of Handwritten Urdu Characters using B Spline Curves: A Survey. *International Journal of Computer Applications* 157.1 (2017).

[24] Ahmed, Saad Bin. (2017). Handwritten Urdu character recognition using one-dimensional BLSTM classifier. *Neural Computing and Applications*. 1-9.

[25] Shi, Baoguang., Xiang Bai, Cong Yao. (2017). An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39 (11) 2298-2304.

[26] Naz, Saeeda (2016). Zoning features and 2DLSTM for Urdu text-line recognition. *Procedia Computer Science* 96 (2016), 16-22.

[27] Khan, K. (2015). Urdu text classification using decision trees. High-Capacity Optical Networks and Enabling/Emerging Technologies (HONET), *In*: 2015 12th International Conference on. IEEE, 2015.

[28] Liao, Minghui. TextBoxes: A Fast Text Detector with a Single Deep Neural Network. *AAAI*. 2017.

[29] Kanwal, Kehkashan. (2014). Assistive Glove for Pakistani Sign Language Translation. *In*: Multi-Topic Conference (INMIC), 2014 IEEE 17th International. IEEE, 2014.