1 HANDWRITTEN TEXT RECOGNITION FOR INDIAN LANGUAGES

1.1 INTRODUCTION

The absolute prominence of information mainly interpreted by three bodies: Person, Equipment and Medium. When all three bodies emerge in decent way produce fruitful communication. An active and adequate communication helps to spread knowledge among the people in the world. Usually, person refers their first language to produce their expression, which serves as a vehicle for expressing their desires, ideas and mean [1]. One of the most popular and traditional methods to produces person's expression is "Pen and Paper", which led to most easy, cheap and convenient way to record handwritten data [2]. Also, people use handwritten paper document as it is convenient to complete work speedily. It is used as evidence too.

Handwritten data have been excellent means of communication for thousands of years and till today. In country like India, these documents were generated mainly in regional languages. This is due to the fact that people uses their first language for communication.

In India, there are many places where handwritten paper data records are generated which are in regional languages. To name a few, information gathered by filling a form, questionnaire, bank forms, medical forms, admission forms, etc. In current digital society, these handwritten paper documents needs to be in digital form. It has benefit of searching through keyword, sorting, easily accessible from anywhere, security, cost reduction, easy movement, easy storage, convert into speech, and process through word processor.

To digitize handwritten documents, scanning and capturing interfaces are used where result are in image form [3]. Though these image formats consider as a digital document, the information contained in it cannot be searched through, interpreted or processed.

Another way to digitize is entering data through keyboard; causes great human labour [2,4,5]. Furthermore, to input text in regional language through keyboard is difficult due to the multiple key combinations require to key in for regional language as keyboard is meant for English script [6]. So, preferable approach is the usage of scanning & capturing interfaces. The handwritten text recognition (HTR) is technology that helps for this [7-10]. It is subtype of optical character recognition (OCR), a field of pattern recognition [2,9,11].

In last few decades, many researchers have put up their high exertions in the field of OCR and accomplished recognition in accurate manner for many foreign languages like English, Japanese, Chinese, Korean, and Arabic [3,9,12-18]. In recent years, the research in automated OCR is at the peak for many of the Indic languages like Devnagari, Gurumukhi, Bangla, Oriya, Kannad, Malaylam, Tamil and Marathi [4,5,19-26].

The Gujarati is official language of Gujarat state, spoken by 65.5 million people in the world giving it 26th most spoken native language **[27,28]**. Despite of that, work in handwritten Gujarati OCR is very limited [29-31]. Most of the work in Gujarati language focused on printed character recognition or handwritten numerals recognition [32-35]. Few researchers worked in the area of isolated handwritten Gujarati alphabet and numeral recognition and still in the search of good accuracy [27].

The work concentrating on recognition of Gujarati text from handwritten paper document is rare due to segmentation of text, large character set, variation in character shape and diacritic mark [27]. The segmentation is a crucial stage in document image processing due to varying text characteristics, fluctuating lines and irregularity in character shape [6,36-38]. Hence automated offline handwritten text recognition (HTR) is still an open problem, specially for Gujarati text.

In this research focus is towards handwritten text recognition but to understand the complexity involved in handwritten text recognition we start next section with brief description of OCR, its history and types.

1.2 OCR AND ITS TYPE

Pattern recognition is a branch of machine learning that focuses on the recognition of patterns and regularities in data [39]. It is the process of classifying input data into objects/classes based on key features. An optical character recognition (OCR) is challenging field of research in pattern recognition. It is an actively studied topic in academic, research and industry because of its enormous application potential [3,17,40]. The process of translating optically acquired images of machine printed, typed or handwritten text document into machine-encoded text such as ASCII or UNICODE is called OCR [2,9].

The OCR makes document text searchable, which is huge advantage to be able to search through documents by various means. It saves considerable amount of time compare to searching through paper based documents. The text generated by OCR can be used in various machine processes such as data entry from the paper data records, postal code recognition, address extraction from postal document, automated evaluation system, text-to-speech conversion, writer identification, and converted to text for processing through word processor etc.

1.2.1 OCR: History

The OCR was first dreamed by Tusheck in 1927 and Handel in 1933 to build automatic machine that can read characters and numbers [41,42]. They conceptualized template matching using optical and mechanical component for character printed on paper and patented their idea.

The prototype of reading upper case type written character was developed by J. Rainbow in 1954 at the speed of one character per minute. By 1967, OCR systems become commercially available in market by IBM. During this period Toshiba and Hitachi also came-up with their first OCR system. However, this system was too expensive [43].

Research prior to 1980s was limited due to computers were lacking in powerful hardware and data acquisition devices. The next decade 1981-1990, took significant growth with respect to printed and typewritten character recognition

[8,44-46]. The hardware prices went down and OCR systems are available as software package in mid of 1980s.

The OCR approaches have evolved over time as needs and technologies have changes [47,48]. The growth in OCR system is at the peak during 1990s as image processing and pattern recognition techniques were effectively and efficiently combined with artificial intelligence. The advancement continues due to more powerful computers and electronic equipment like scanner, camera and electronic tablets. During this time researchers focused more on hand-printed numeral, character and joined/cursive handwriting recognition [41,43,49,50]. Unfortunately the success of OCR could not carry on to handwriting recognition as for the recognition of isolated handwritten numerals, due to the variability in people's handwriting [34,45,50,52]. The OCR problems still exist with unusual character shapes, different types of fonts and with poor quality documents [42].

1.2.2 Type of OCR

OCR technology divided into two category based on the data acquisition process as online and offline [44]. The classification of OCR is shown in Figure 1. In online character recognition system, character is processed for recognition as soon as it is written on a pressure sensitive surface, from which real time information, such as the order of the stroke made by the writer, pen up/down switching and pressure data are obtained. Due to these inputs online recognition system performs better compare to offline recognition system [44].

The offline OCR system is dealing with the scanned document image which is previously typed, printed or handwritten to recognize data. It can be performed days, months or years later. The offline OCR can be categorized based on its input document as typed or printed character recognition (PCR) and handwritten character recognition (HCR) [41,53].

In offline HCR, recognition carried out only after all writing is complete. The input document image is two-dimensional data containing handwritten text. The HCR is significantly more challenging than PCR [10,54,55]. The PCR deals with well

define shape and size of the characters, while the HCR faces all the challenges that are faced by PCR which includes page quality, stroke thickness, layout typeface, size, etc. Apart from these, the HCR is more complex due to numerous variability of writing styles of different persons and the same person can write differently in different state of mind [10,54,56].

Figure 1 Classification of OCR
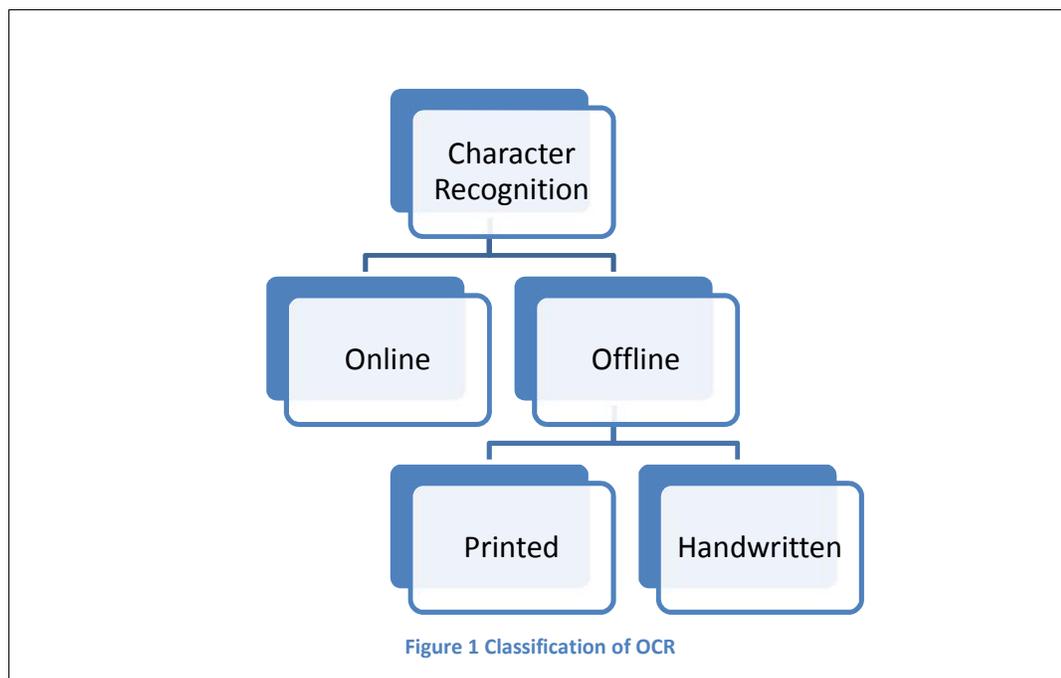
The HTR can be further categorized based on how text is written i.e. text written in predefine boxes, separated by space, run on discretely written, cursive writing or mix mode writing [6,57]. The HCR refers to recognition of handwritten character in any form while the HTR, more specifically refers to read scanned document images containing text lines and convert them into editable text [58].

The HTR system faces all the challenges that are faced by isolated HCR system moreover complexity involved with the segmentation of document image. Furthermore, the document written in Indian languages like Gujarati adds complexity in HTR system. The Indic languages has large character set compare to English, varying writing style and modifiers attached to character shape increases the level of complexity [11].

Even a well-spaced handwritten document faces challenges in segmentation and recognition due to linguistic characteristics like character shape etc. Looking to these difficulties researcher decided to work on handwritten text recognition for Gujarati script. In the next section, towards the finding solution, researcher discussed general architecture used for HCR.

1.3 GENERAL ARCHITECTURE FOR HANDWRITTEN TEXT RECOGNITION

The general architecture of HTR system consists of phases namely preprocessing, segmentation, feature extraction and recognition as shown in Figure 2 [9,13,29,34,51,53,54,56,60]**.** Different researcher focuses on different phases. The selection of the phases for the solution is based on strategy used for solution, script style, its characteristics and application. Before discussing general architecture, different approaches that are used for overall solution require to study. There are mainly two types of approaches used to design text recognition, which is either holistic or analytic [13,40,49,54,55,61,62].

The holistic approach is based on how human recognize the text that is while reading characters. Human doesn't isolate each of the character in the word. It is top-down approach to recognize full word, which will eliminate segmentation of word to further into characters. Automation of this kind of system works well with limited vocabulary. This approach is useful in case where segmenting word into character is complex task [57,62,63]. Mainly used with script which uses cursive writing i.e. English handwritten cursive text, Arabic text document etc.

The analytical approach is to segment word into characters. Then recognize them during recognition phase. Most of the character recognition system uses analytical approach and recognizing individual characters primarily by their shape and appropriate for Indic script [13,49,54,55,61,62,64]. The analytical approach is bottom up approach. It has advantages of unlimited vocabulary over limited vocabulary of holistic approach.

Figure 2 General architecture of HTR
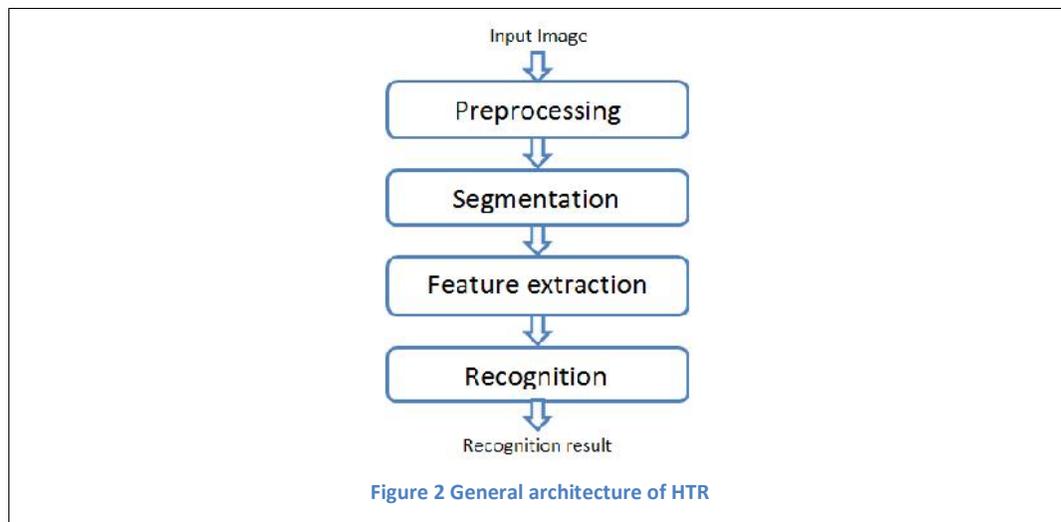
Preprocessing: The preprocessing phase deals with various steps which enhance the document for the next phase segmentation. These steps include conversion of document image from colour to grayscale scale, removal of unwanted pixels values called noise removal, representing document in foreground as text and background i.e. binarization, correcting skew of document using skew detection & correction, cropping document to only text area, improving pixels value using morphological processes and thinning of character to one pixel width etc. based on input method, characteristics of text document image and applications.

The main objectives of preprocessing are reduction in image data, reduce noise present in the scanned image, stroke normalization, skew detection and correction and normalization of character image. The detail discussion on various preprocessing techniques is given in [6].

There are many tools and techniques are available to remove noise from the image data such as smoothing, sharpening, median filter, wiener filter, contrast adjustment, and thresholding [6,65]. For binarization of document two types of thresholding methods are used, local and global [65]. Binarization using a global thresholding method suggested by Otsu is widely used [27-29,32,34,56,66,67]. It works based on minimization of the weighted sum of within-class variances of the foreground and background pixels using gray level histograms [39,65,68].

The process of document skew detection and correction is considered mandatory in preprocessing of offline text recognition. A skew is generally introduced into the image while scanning due to improper alignment. The skew detection methods are based on Hough transformation, cross correlation, vertical and horizontal projection profile [37,69].

Thinning removes irregularities in the character shape by making character width uniform as 1 pixel [65,70]. It reduces the amount of pixel data and helps to extract features independent from character stroke thickness. Morphological thinning algorithm is widely used [65,71,72].

Segmentation: Segmentation is one of the most vital steps in HTR system [65,73,74]. The goal of segmentation in HTR is to partition the document image into connected regions such that each region should exhibit uniformity and homogeneity with respect to one or more characteristics.

The process of segmentation decided based on overall recognition approach. It is mainly classified as holistic or analytical. In holistic approach, the lowest unit of segmentation output is word. In case of analytical approach, the lowest unit of segmentation output is character. The segmented lowest units word or character is processed for feature extraction and recognition.

In analytical approach, the isolated OCR system normalize segmented unit which is character converted into fix size unit for feature extraction. In handwritten OCR the process of normalizing segmented unit changes the shape of character especially for Indic script, which leads to recognition error [6].

In HTR system the segmentation process divides text document into segments of lines called line segmentation. The output of line segmentation algorithm is each pixel in a text image is assigned to a line. Further segmented line divided into words or characters referred as word or character segmentation. The output is each pixel of the segmented line image is assigned to a word/character.

Segmentation of handwritten text into lines and character is one of the important phases in HTR process. The process of line segmentation from unconstraint handwritten document is complex due to freely flowing text which

causes uneven distance between lines. While character segmentation comparatively easier than line segmentation and can be achieved by applying connected component analysis or by vertical projection [22,75].

The output of segmentation is then processed for feature extraction in next phase for recognition.

Feature extraction: A feature is object's characteristics or primitive which helps to describe the object in meaningful manner so it aids the recognition process and help in discrimination of objects. The aim of feature extraction process is to extract or generate relevant information i.e. features from the segmented unit (i.e. character) to assist the task of character recognition.

These extracted set of features is called feature vector or feature set, which is used in recognition of characters using classification methods. The selection of features are in such a way that it can be classify to its character class and maximize the rate of recognition with least amount of elements. This stage is critical because the quality of the features influences the recognition process.

Classification: The extracted set of features is given as an input to the recognition process. The recognition or classification engine used to classify character based on feature set into its character class using its training algorithm.

For classification of character different researchers uses different technique based on problem statement formulated and strategy used. The classification techniques for character recognition are structural and rule based techniques, statistical technique and neural networks [5,11,34,76,77]. For handwritten character variation neural network is robust due to it can adapt and learn from stable features over a large training data [29,76,78,79].

1.4 THE GUJARATI SCRIPT

The Gujarati language is an Indo-Aryan language native to the Indian state of Gujarat. It is part of the greater Indo-European language family which is descended from old Gujarati [80]. It is one of the 22nd official language recognized by the Government of India, used in western part of India, and official language of Gujarat as well as in the union territories of Daman and Diu.

The Gujarati script is descended from Brahmi. It was adapted from the Devanagari script to write the Gujarati language. In accordance with all other Indic scripts, it is written from left to right with a specific alphabet similar to Devanagari, which is used to write Hindi and Sanskrit [66,81]. It is not case-sensitive and written without head line called "Shirolekha".

Gujarati script has 15 vowels, which are called "Swar", 34+2 consonants (2 conjuncts are used as consonants), called "Vyanjan", and 10 digit symbols. Gujarati consists of 14 diacritic mark special symbols also called "Matra", corresponding to each vowel, which is attached to consonants to modify their sound. [81].

All these diacritic marks are placed at the top, left, right or bottom part of the consonants and it can be attached at different positions for different consonants. These diacritic marks can occur in different shapes depending on the consonant to which it is attached. The Gujarati consonant character set, vowels, numerals, diacritic marks and some of the conjuncts characters are shown in appendices I to V respectively [27,32,35,38,80-82].

The Gujarati script follows an abugida writing system like all Nagari (Devnagari) [80]. The abugida, also known as avugida, is a segmental writing system in which consonant-vowel sequences are written as a unit: each unit is based on consonant letter and vowel notation is secondary. This vowel notation is attached to consonant letter at various positions. In Brahmic script family, the term Akshara is used for the units.

The Gujarati Akshara is logically divided into two parts based on main character and position of the diacritic mark that is core character area and area at top, bottom, left and right of the core character area as shown in Figure 3. A core character area is place where consonants or vowels are written (vowels which are written in character form like "અ a ", "ઇ i I", "ઈ I", "ઉ u u", "ઊ u", "ઋ r u").

The area at top, bottom, left and right of the core character area is a placeholder for upper, lower, left and right diacritic marks respectively.

Figure 3 The parts of Gujarati "Akshara"
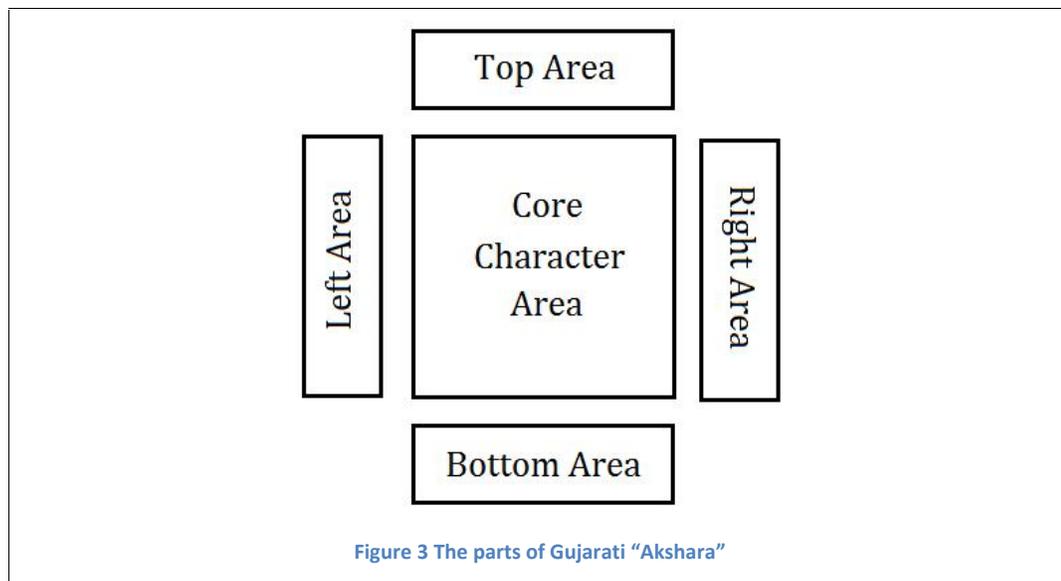
The Gujarati text line formed by a word followed by other words, separated by punctuation mark used for word separation generally a space and ends with statement end punctuation marks like full stop ("." Baseline dot)/exclamation makr("!") or question marks("?"). Each word is formed by Akshara, followed by another Akshara if word contains more characters.

1.5 CHALLENGES IN HANDWRITTEN TEXT RECOGNITION

Working with good quality machine print text document provides recognition accuracy of nearly 100%. The diversity of human writing styles, spacing variations, irregularities of handwriting – all these factors make handwriting text recognition a much more challenging than machine printed or even hand printed text documents. The development of effective HTR system with high recognition accuracy is still an open problem even for language like English. Some of the general issues in development of HTR are:

- Writer variations: Different people have different handwriting due to character size variation, material used for writing i.e. pen, pencil, quality of ink, and age of writer etc. This causes variability in shape of same character image.

- Document segmentation complexity: Due to writing style variations of different writer adds complexity in segmentation of document. Segmentation

is considered to be one of the most important phases of any pattern recognition system. The handwritten text document having more than one line of text. To recognize characters from text document these lines need to segment into individual line and then into words and/or characters.

- Skew in document: There are two types of skew in document. Global skew in the document due to improper alignment of the paper while scanning process adds complexity in segmentation process. The handwritten documents tend to have internal line skew or non-uniform text line skew that is orientation fluctuate within the line while writing.

- Noise in the character due to extending lines: Writer may write some script in cursive manner or joining characters by extending lines of previous character to the next character. This extended line unnecessarily changes the shape of character which adds difficulties in its recognition.

- Noise in input document: The main sources of noise in the input image is due to quality of paper used for writing, text written on both sides of paper "Back Paper Noise" and due to the scanner's brightness and sensors. All these noises contribute to decrease accuracy of OCR system. As a result, of this having a noise correction routine in place becomes inevitable.

The HTR become more complex for Indic scripts due to the language characteristics. Here we are presenting some of them:

- The Indic scripts have large character set which make classification problem more difficult. The rate of recognition for isolated character is yet not reach to satisfactory level for some of the Indian scripts due to large character set and its peculiarity [15,29,83].

- Indic script that do not have head line ("Shirolekah") poses segmentation problem compare to the script which have head line. Like in Hindi script the presence of head line gives clue to detect lines. There is no such clue for script like Gujarati for line segmentation.

- The problem becomes more complex due to the presence of diacritic marks also called as "Matra" which is attached to character at various places based on script style. While attaching it with character it may join with character and modifies shape of the character. Diacritic mark recognition is also complex, attributed to the complex nature of the language.

- Almost all Indian scripts characters are peculiar in nature making them hard to recognize by machines. The shapes of characters are varying even by single writer.

- Some of the characters shape are exactly same or having very minor variation making recognition difficult.

- There are various isolated dots and small symbols, which are vowel modifiers, namely, "Anuswar", "Visarga", "Chandra Bindu", and "Rakar (reph - curved upward dash)" which add up complexity in segmentation and recognition.

- In many Indic scripts like Gujarati has composite characters which can be formed by joining two characters also known as conjuncts. These conjunct characters are increasing difficulties in segmentation as well as in recognition.

1.6 NEED OF WORK

In today's world, user wants the computer to have user-friendly interactions. Due to use of computers and computing devices as well as smart phones in all aspects of human life, it is always being desired that computer system should recognize native languages to help common men in their routine tasks. But routine use of physical documents is still prevalent in most of the communications in daily life. Hence there is a great demand of the software that can automatically extract, analyse, interpret and store information from the physical documents for later usage [10].

As noted earlier, HCR has received extensive attention in academic and industry since last some decades. The research on OCR of handwritten alphabet and numerals has been mostly found to concentrate on Roman script [6,51,54] related to English and some European languages and scripts related to Asian languages like Chinese, Korean, and Japanese [9,12-17]. The English script is one of the most

thoroughly studied in the field of HCR due to internationally accepted language as well as availability of standard datasets for comparing results [8,51,53,54,84].

Recently, the research on offline printed and handwritten OCR for different Indian script has been started and in pace for Kannad, Tamil, Bangla, Oriya, Devanagari, and Gurmukhi, but effort for creating text recognition for Indian scripts are still verdant [4,5,19-26]. Despite of that, very less amount of work is found in case of Gujarati HCR [31] and rare in text recognition. The work on Gujarati script is limited to numerals or isolated character recognition as well as it is lacking in rate of recognition.

The first effort for developing Gujarati OCR found in [32] with small subset of printed Gujarati alphabet. In [38], attempt for Gujarati script recognition is noted with zone identification in printed text. The isolated Gujarati HCR using Euclidean distance classifier achieved overall 26.86% recognition rate. The work concentrating on Gujarati handwritten numerals are found in [29,34,45,77].

Gujarati printed character recognition using modified version of Hidden Markov Model found in [35,82]. In [85] worked on identifying similar appearing Gujarati printed character and achieved 67% overall result using structural features along with wavelet features.

The research work on handwritten Gujarati character recognition is found mostly concentrated on numeral recognition. The rate of recognition for handwritten Gujarati numeral are 90.55 in [77], 80.5% in [30], 88.79% in [29], 82% in [34] and 96.32% in [86].

The significant work on isolated non-numeral Gujarati characters is found in [27] with the success rate of 63.1%. The structural analysis of Gujarati consonant is presented in [59,67] for development of handwritten character recognition with experimenting on five consonant and three numerals. They achieved 88.13% and 91.33% of recognition rate only for five consonant character and three numerals respectively with dataset size 150 for each character.

Limitation of these approaches is that they used structural characteristics in their feature vector. The handwritten character tends to have various types of noise

and gap between the lines. Due to this, noisy and broken characters lead to different subset of characters or misclassified or unclassified characters. Also, to note that their effort was on recognition of isolated Gujarati characters, so dataset is also limited to that purpose.

Handwritten text recognition system for Gujarati script is able to read scanned handwritten Gujarati document images and convert it into editable Gujarati text. One can find research on printed Gujarati text recognition which can read printed document image and convert it into editable text but limited efforts in HTR in Gujarati. Following researchers made efforts by considering creating system which can read handwritten Gujarati words/text.

The extraction of modifier from characters of Gujarati text document is suggested in [87]. The approach was based on formation of Gujarati word as a complete consonant-vowel cluster i.e. as a distinct entity. In Gujarati two modifiers occurs purely in lower zone, three occurs at purely in upper zone and two purely in middle zone, while four modifiers present in middle as well as in upper zone. They used connected component method based on 8-connectivity of pixel relationship. The algorithm suggested by them is having 60% accuracy with some of the connected characters while it gives 82.4% accuracy for without connected characters.

Chaudhari S. & Gulati R. discussed issues of segmentation associated with Gujarati text document in [88]. They divided segmentation problems in three categories which are line segmentation, word segmentation and character segmentation.
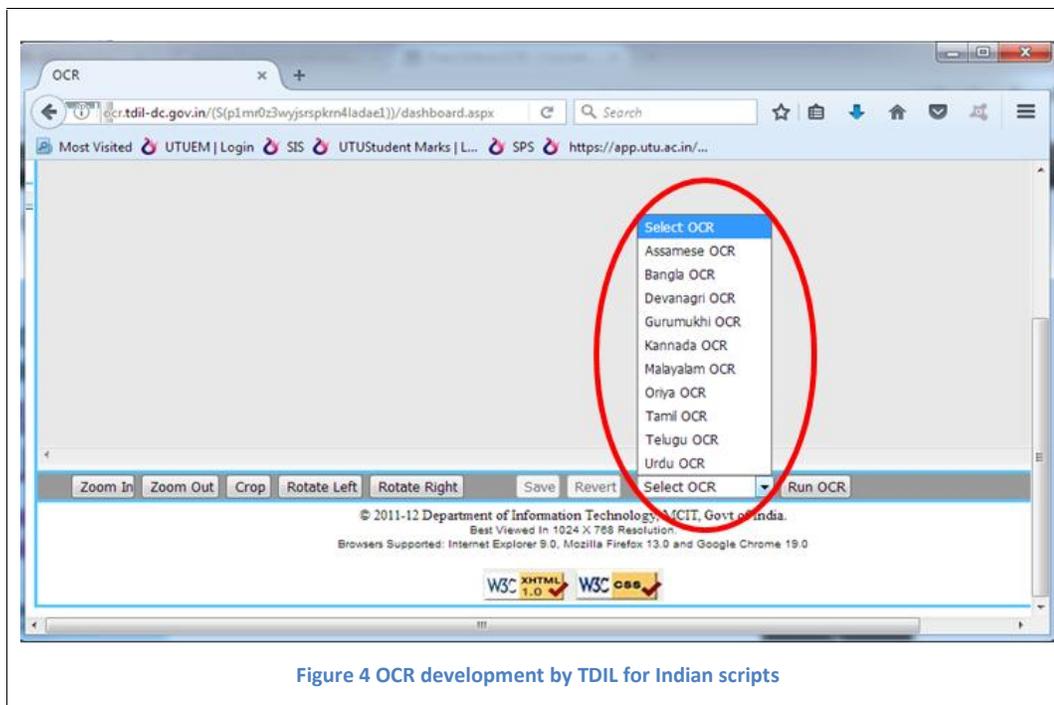
Figure 4 OCR development by TDIL for Indian scripts

The Ministry of Communication & Information Technology (MCIT), Government of India, has started a program called Technology Development for Indian Languages (TDIL) in 1991. One of the stated objectives of the TDIL is to develop information processing   tools   and   techniques for Indian languages.

To fulfil the objective OCR system development for most of the important Indian language scripts have been taken up by different labs and academic institutions as project. The OCR developed for different Indian scripts are shown in Figure 4[1] under TDIL programme. It shows absence of Gujarati script in language selection menu.

1.7 AIM AND OBJECTIVES

The main objective of research is that to create offline Gujarati handwritten text recognition (OGHTR) system which is capable to segment handwritten Gujarati document and generate Gujarati text. The system needs effective Gujarati character

---

[1] Source: http://ocr.tdil-dc.gov.in/(S(2sph4ccx3sgr3j1bjr055xp0))/default.aspx accessed date: 22/09/2017 10:29 am

recognition to recognize handwritten Gujarati text. The other objectives of research are:

- To create a model for handwritten text recognition for Gujarati script.
- Segmentation of Gujarati text documents into core characters, and diacritic marks.
- Recognition of isolated Gujarati characters includes alphabet, numerals and diacritic marks.
- Study of Gujarati character's features.
- Develop methods to extract character features.
- Segmentation of character attached with diacritic marks.
- Design and development of tool for preprocessing of handwritten text documents.

1.8 STATEMENT OF THE PROBLEM

The regional languages of India are very peculiar in nature and require more efforts to keying in to computer for digitization purpose. The handwritten text which is written in regional language can be used as input in place of keyboard for keying data, which reduces the great amount of efforts need for typing. Such applications not just reducing the efforts of keying in but opens door for many applications which can be thought for regional languages.

The amount of work in the field of HTR is very limited for regional languages of India compare to foreign languages. Gujarati text segmentation and recognition offers more challenges like most other Indian scripts as well as added difficulties of language characteristics.

Based on above discussion and objective of the study I have decided the title of the thesis as "**Design and development of text line segmentation and recognition of offline handwritten Gujarati text**".

1.9 THESIS OVERVIEW

In Chapter 1 Introduces need of handwritten text recognition for Indic scripts. It discusses general architecture for HTR system and brief description of its phases

and approach. The challenges involve in HTR for Indic script and specifically for Gujarati is presented as our focus is towards HTR system for Gujarati script. The current status of research work in the Gujarati HCR is also discussed.

Chapter 2 is dedicated to literature review. It presents previous work in segmentation, feature extraction and classification approaches used by different researcher for international and national languages. Later in the chapter review with respect to handwritten OCR solutions for Gujarati script is presented.

In Chapter 3 presented the architectural solution towards the solution of creating offline handwritten text recognition system for Gujarati. It briefly discusses overall approach towards the solution and its phases.

In Chapter 4, the first stage of OGHR system, preprocessing is presented. Preprocessing of image document is essential for improving quality for better processing. It deals with transformation of text document into the form which is suitable for segmentation. The preprocessing steps designed in this research are for Gujarati HTR system.

Chapter 5 describes the segmentation phase in detail used for Gujarati text. The segmentation is one of the most vital phases of any recognition system. Our segmentation approach is based on combination method. The output of segmentation phase is categorising into core character and diacritic marks.

Chapter 6 explains the strategy used for recognition of diacritic mark and core character. For diacritic mark recognition tree based classifier is used as they are few in numbers as well as position is also available from segmentation output. For recognition of core character isolated handwritten Gujarati character recognition system is implemented. The neural network is trained with extracted features from handwritten character dataset. Two novel features have been used to recognize character is presented along with the other features. The chapter ends with discussion on text generation step.

Chapter 7 deals with the results and analysis of experiment of HTR system. It discusses results of segmentation phase followed by results of IGHCR module. The overall result is compared to know effectiveness of different feature vectors. The

results are also compared with recent works. We also determined effectiveness of one added preprocessing step in IGHCR module. Finally, chapter concluded with OGHTR results.

Chapter 8 discusses the conclusion drawn during the experiment and analysis of research work.