ABSTRACT
Many HOCR systems was developed for different foreign language and efforts on Indian script still verdant. Due to the varying writing style of each individual HOCR becomes a critical challenge since last few decades. India was known to world for its unity in diversity. The geographical and cultural vastness of the country various spoken languages and writing scripts were invented and utilized for daily communication. Courageous endeavors was made towards HOCR for various Indian scripts. The development of HOCR system was considered an active area of research due to the complexity and structure of Indian scripts. MODI Script was an ancient script of India as compared with other Indian scripts. Large number of historical documents exists and they need to be preserved and explored therefore special attention must be given to MODI script. This research paper is an attempt towards an investigation of recent advances in Handwritten Optical Character Recognition system for MODI script.

Keywords
MODI Script, HOCR, Segmentation, Feature Extraction, Recognition, Image Classification

1. INTRODUCTION
Many researchers were attracted towards handwritten character recognition which is the most promising area of Pattern recognition, Image processing, Natural Language processing and Document analysis. HOCR was a process that provides intelligence to a computer that converts human writings into text. Due to advancements in technology pattern recognition and image processing lot of improvement was observed in recognizing handwritten characters.

Handwritten character recognition problem was defined based on general model of OCR with six stages 1) Preprocessing, 2) Segmentation, 3) Feature Extraction, 4) Classification and Recognition 6) Post processing. General framework of HOCR was shown in figure 1.


Rest of the paper was organized as, History of MODI script was given in section 2, Properties of MODI script was given in section 3, Pre-processing was discussed in section 4, Segmentation in section 5, Feature extraction methodologies in section 6, Techniques for classification and recognition in section 7, Post processing in section 8. Concluding the review in section 9, finally, Future research directions was discussed in Section 10.

2. HISTORY OF MODI SCRIPT
According to a people's linguistic survey of India conducted by David Lalmalsawma [7] India speaks 780 languages and 86 different scripts out of which 220 languages was disappeared in last 50 years and also bespeak a threat that 150 languages could disappear. MODI Script was one of the ancient script and vanished from official work during 19th century before Devanagari was officially adopted. Total time span where MODI was utilized for written communication at different stages was about 600 years. Meanwhile MODI became very popular and frequently used to write 'Marathi' (primary language of Maharashtra in western India) [8,9].

Origin of this script still uncertain. According to Savant M.Y.[10], Rakesh A. Ramraje [11] MODI alphabets was developed by 'Hemadpant' or 'Hemadri' in 12th Century. Hemadpant was an administrator in the kingdom of ‘Mahadev Yadav’ and ‘Ramdev Yadav’ (1187-1318 at ‘Devgiri’). According to history expert Chandorkar, MODI script was evolved from Mouri (Bramhi) script of Ashoka period. Besekar D.N. & Ramteke R.J. [4], Besekar D.N. & Ramteke R.J. [5] and Naren Ranadive [9] states that MODI alphabet was invented during the 17th century and frequently used for writing purpose in the era of 'Peshwai' (Pune) and 'Chatrapati Shivaji Maharaja'.

According to Salgaonkar et.al. [12] oldest available MODI document was of 1429 and 1389 A.D. preserved in Bharat Itihas Sanshodhan Mandal (BISM) Pune.

According to Anshuman Pandey [13] earliest in 12th century MODI script was known as 'proto-MODI' or 'Adyakalin', 13th century ‘Yadavakalin’, ‘Bahamanikalin’ of the 14th-16th century, ‘Shivakalin’ of the 17th century and ‘Chitnisi’ during 18th century called as ‘Peshvekalin’ between 1818 to 1952.
3. PROPERTIES OF MODI SCRIPT

The MODI was written by ‘Boru’ or ‘lekhan’. ‘Lekhan’ was a pen created with the help of ‘Bambuu’ which need to lift too often for dipping in the ink. The MODI script was formed by 46 distinctive letters in which 36 were consonants and 10 vowels. These characters were known as basic characters.

As compare to Devanagari there were 48 distinctive letters includes 36 consonants and 12 vowels. Long ‘i’ and Long ‘u’ was not used in MODI Script.

Before the commencement of writing in MODI characters a horizontal line was drawn across the page. Characters were written with respect to the horizontal line with full of moulds and curves to avoid lifting of ‘Boru’ too often. According to Savant M.Y. [10] no punctuation marks and conjuncts were used in this script. Termination symbol for sentence or word was not used in MODI Script. Speed of writing was increased due to elimination of these symbols.

Consonants used in MODI script was shown in figure 2 and Vowels used in MODI Script was shown in figure 3.

Vowels was written as independent letters, or also by using diacritical marks at above, below, before or after to the consonant to which they belongs. Pattern of writing the vowels was known as modifiers and newly formed characters were called as conjuncts. Use of these modifiers was shown in figure 4 and some conjunctions were shown in figure 5.

Sometimes two or more consonants were combined to produce new shape called compound character. Similar structural representation of characters was found in MODI, Devanagari, Hindi, Sanskrit and Marathi. Some examples of compound characters was shown in figure 6.

Pattern recognition of MODI Character was very challenging due to the inconsistency in the writing style, similarity in the character shapes, presence of modifiers and various other features of MODI scripts. The actual differences were visible in the behaviors of characters in certain environments, such as calligraphic uniformity of the characters, cursive nature of the script and degraded document. Similarity in calligraphic representation of 38 basic characters was shown in figure 7.

Figure 2: The consonants used in MODI Script.

Figure 3: Vowels used in MODI Script.

Figure 4: Use of Modifiers in Historical MODI Script

Figure 5: Conjunctions in Historical MODI Script

Figure 6: Compound characters in MODI Script

Figure 7: Uniformity in calligraphy of MODI characters
Principal interest of this review was to analyze different methodologies implemented in HOCR with special attention to MODI script. Despite of advancement in technology, HOCR for MODI Script was still in initial phase. Up till now only six studies were available for MODI script. Besekar D.N.[1] stated a morphological approach for recognition of five numerals from MODI script. Feature set was created by using blobs, vertical lines, horizontal lines, concavities present in the numerals. Recognition and classification was done by using decision tree and mathematical morphology. In this case 75% recognition rate was achieved for 0 and 1 and lower for 4, 7 and 9.

Chain code and image centroid based recognition model was presented by Besekar D.N [2] for vowels in MODI script. In this work median filter was used for removing noise, global threshold for binarization, flood fill for avoiding boundary breaks and size normalization. Two layer feed forward Neural Network and SVM was used for classification and achieved 65.3% to 73.5% recognition rate.

Structural similarities was used by Ramteke A.S. & Katkar G.S [3,6] for recognition of MODI characters. Structural similarity approach was used for measuring image quality and image quality metrics. measured structure similarity (SSIM), KNN and back propagation NN was used for classification and achieved 91 to 97% recognition rate.

Zone based approach was used by Besekar D.N. & Ramteke R.J. [4] for recognition of offline handwritten numerals. In this study preprocessing was done by using median filter for removing noise, global threshold for binarization, flood fill for avoiding boundary breaks and size normalization. Feature set was created with the help of four equal square zones of size 15×15 and their polar coordinate, Variance, Theta angle and Rh distance. Using variance table for classification this study achieved 93.5% recognition rate.

Theoretical analysis of MODI Script recognition was done by Besekar D.N. & Ramteke R.J. [5]. Devanagari, MODI and Roman scripts were compared in this work and found that structural features were difficult to extract for MODI script. In this work internal and external segmentations were discussed and advised Internal segmentation for MODI script. In this analysis structural as well as topological features were suggested. This study also explains that HOCR for MODI script was a difficult task as compared to other handwritten script due to the cursive nature, variations in character, handwriting habits and synonymous structure of characters.

Summarily this review concludes that specific steps and results obtained in preprocessing was not shown by [1,3,4,5,6]. Methods and results of segmentation was unclear in [1-6] and neither of these studies were considered post processing for MODI script. Analysis of these six studies were shown in table 1.

<table>
<thead>
<tr>
<th>Author and Reference</th>
<th>Type of Data</th>
<th>Preprocessing</th>
<th>Segmentation</th>
<th>Feature Extraction</th>
<th>Classification and Recognition</th>
<th>Post Process</th>
<th>RR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Besekar D.N. [1]</td>
<td>Five MODI numerals (0,1,4,7,9)</td>
<td>Morphological operations</td>
<td>Unclear</td>
<td>blobs, vertical &amp; horizontal lines, concavities.</td>
<td>Mathematical morphology and Decision tree.</td>
<td>Unclear</td>
<td>75%</td>
</tr>
<tr>
<td>Besekar D.N. [2]</td>
<td>MODI vowels</td>
<td>Median filter Skewing, Binarization, Edge detection, Flood fill and Normalization</td>
<td>Unclear</td>
<td>Chain codes, Centroid, Normalized chain code histogram.</td>
<td>Two layer feed forward NN and SVM</td>
<td>Error correction and Unicode</td>
<td>65.3% to 73.5%</td>
</tr>
<tr>
<td>Ramteke A.S. &amp; Katkar G.S. [3]</td>
<td>MODI characters</td>
<td>Depends on method of data acquisition</td>
<td>Unclear</td>
<td>Structural similarity approach</td>
<td>SSIM and KNN, back propagation NN.</td>
<td>Unclear</td>
<td>91 to 97%</td>
</tr>
<tr>
<td>Besekar D.N. &amp; Ramteke R.J. [4]</td>
<td>MODI numerals</td>
<td>Morphological operations, Median filter Skewing, Binarization, Edge detection, Flood fill and Normalization</td>
<td>Unclear</td>
<td>Four square zones of size 15x15, Polar coordinate of zone, Variance, theta angle and Rh distance</td>
<td>Variance table</td>
<td>Unclear</td>
<td>93.5%</td>
</tr>
<tr>
<td>Besekar D.N. &amp; Ramteke R.J. [5]</td>
<td>MODI characters</td>
<td>Depends on method of data acquisition</td>
<td>Unclear</td>
<td>Structural similarity approach</td>
<td>SSIM and KNN, back propagation NN</td>
<td>Unclear</td>
<td>91 to 97%</td>
</tr>
</tbody>
</table>

4. PRE-PROCESSING
Preprocessing was used to produce clean and inflectionless document to be used in later stages of HOCR. Various infections were noticed in old MODI historical documents. Plenty of such historical documents were about 600 years old [8,10] preserved with the help of cloth covers [1-6]. Keeping all these factors preprocessing stage became essential for the development of MODI HOCR. An review on offline handwritten character recognition was done by Nafiz Arica et.al. [14] and stated various essential preprocessing steps like noise removal, filtering, morphological operations, noise Modeling, normalization, Skew normalization and baseline extraction, slant normalization, size normalization, contour smoothing, compression, thresholding and thinning. Generally following techniques were used in various studies for preprocessing.

4.1 Noise Reduction
Noise removal was a process to remove noises produced at the time of capturing or transmission of the image. In case of historical MODI documents many types of noises were
introduced by scanning process such as disconnected line segments, bumps and gaps in lines, filled loops, local variations, rounding of corners, dilation, erosion, spike noise, random noise or independent noise, salt and pepper noise, black and white dot and dust particles. J. Serra [15] stated that it was necessary to eliminate various imperfections for best possible results in OCR.

4.2 Filtering
Filtering was a process to remove unwanted component from image. These unwanted components were produced by uneven writing surface, poor sampling rate of the data acquisition device. The book written by Ramesh Jain et.al. [16] stated various applications of image filtering techniques. Rohit Verma & Jahid Ali [17] stated use of filtering techniques and found that adaptive filter BM3D was the best choice for removing salt and paper noise also shows that median filter was reliable choice for filtering. Filtering techniques like smoothing, sharpening, threshold, contrast adjustment and textured or colored background elimination were studied by [18-22].

4.3 Morphological Operations:
Morphological operations are playing very important role for enhancing the accuracy of HOCR. Erosion, Dilation, Closing, Opening and skeleton, thinning were basic morphological operations. Although unexpected hoodesness was produced by careless use of Morphological techniques the good quality of image was produced by careful use [14].

4.4 Skew angle detection
Inconsistency of scanning process and writing style were major constrains in HOCR for various languages. In MODI script writing style of characters were slightly tilted, curved and overlapped within the image and calligraphy of MODI script hurts accuracy of HOCR [6]. Therefore skew angle detection and correction became essential for HOCR in MODI script. Naazia Makkar & Sukhjit Singh [23] stated that projection profile, Hough transformation, K-NN Clustering, Thinning and Cross Co-relation for skew angle detection and correction.

4.5 Normalization
Normalization process used to modify all characters in a unique and standard form. Slant normalization and Size Normalization were commonly used methods [14]. Slant normalization focuses on angle of characters and Size Normalization focuses on size of the character. Due to calligraphic effect normalization become essential for MODI HOCR.

It was assumed that the preprocessing stage gives clear and normalized document with sufficient amount of shape information and high compression.

5. SEGMENTATION
Segmentation, divides the document into subcomponents like words, lines and characters. Accuracy produced by segmentation directly affect to the recognition rate. Segmentation became major concern area in many HOCR. Still there was a room for research in segmentation of touching, overlapping, cursive and complex scripts.


5.1 External Segmentation
External segmentation was used to extract paragraphs, sentences, or words with the help of structural analysis and functional analysis. Structural analysis was used to divide the document into blocks and functional analysis was used to find location and size of the components.

5.2 Internal Segmentation
Internal segmentation was used to extract isolates letters from cursive script. Internal segmentation was divided in Explicit Segmentation, Implicit Segmentation and Hybrid Segmentation [26,27,28]. Explicit Segmentation was used to find interconnections and cut the image according to the interconnections. In case of implicit segmentation a class of model images was designed and compared with the image. Image was then segmented according to match found in system search.

6. FEATURE EXTRACTION
Feature extraction process plays very crucial role in development of promising HOCR systems. Large no of feature extraction methods have been reported by researchers working on HOCR and effectiveness of these methods has been assessed in different studies using different forms of data. The book written by Mark Nixon and Alberto Aguado [29] stated that feature was a point of interest according to image classification and recognition also stated various feature extraction techniques. A review of feature extraction was done by Dong ping Tian [30] which summarizes features according color histogram, color moments, texture, shape, moments, scale and shape transformation.

Oivind Due Trier et.al. [31] stated use of template matching, deformable templates, unitary image transforms, zoning, geometric moment invariants, zernike moments and orthogonal moments for feature extraction. Gurpreet Singh et.al. [32] evaluated use of distance profile features, projection histogram features, zoning density features, zernike moment features and hybrid features for Gurumukhi Script.

Pritpal Singh & Sumit Budhiraja [33] stated zoning density with background directional distribution features for recognition of Gurumukhi script.

7. CLASSIFICATION AND RECOGNITION
Various techniques were developed since last few decades for classifying an unknown sample into a predefined class. Generally classification and recognition in HOCR was done by using holistic or analytic strategies [34,35,36]. Top down approach was used in holistic strategy for recognition of full word and segmentation problem was eliminated in this case. On other hand bottom up approach was used in analytic strategy for recognition of characters. Character segmentation was major concerned in this case. Errors at segmentation level was reflected in classification and recognition.

Priyanka Sharma & Manavjeet Kaur [37] discussed statistical model, syntactical or structural model, template matching model and neural network based model for pattern recognition, and also discussed use of Linear discriminant analysis (LDA), Quadratic Discriminant Analysis (QDA), Maximum entropy classifier, Decision trees, Kernel Estimation & K-nearest neighbor (KNN), Naive Bayes classifier, Artificial Neural Networks (ANN) and Support Vector Machine (SVM) for pattern classification.
Charles C et al. [38] evaluated various methods for online as well as offline handwritten character recognition with special interest to Chinese script also discussed various types of recognition methods for cursive script.

Various recognition and classification techniques were reviewed by Naﬁz Arica et al. [14] and Gaikwad V.A. & D.S. Bormane [39] like template matching, statistical techniques, structural techniques, Neural Networks (NNs) for offline handwritten character recognition. Ramzi Haraty & Catherine Ghaddar [40] stated Artiﬁcial Neural Network (ANN) for Arabic text recognition and classiﬁcation and found that promising results were produced by two generalized feed forward neural network.

8. POST-PROCESSING

A post processor was an integral part of any OCR system. Errors produced in recognition was major concern for implementation of post processing stage. An important role of correcting errors and resolving ambiguities in OCR was given to post processing in HOCR. Unfortunately, HOCR was still imperfect as it occasionally produces syntax and semantic errors in output text.

Post processing was done at word level, at sentence level and at semantics level. The most common post-processing technique at the word level was dictionary look-up method where an dictionary was developed to compare generated output of the OCR [41]. A dictionary can produce spell check and provide some alternatives for the outputs produced by recognizer. Ultimately utilization of a dictionary for correcting the minor mistakes was became a solution for OCR systems [41,42,43].

Accuracy of the HOCR was increased when post processing was implemented by using semantic information instead of using this context information, whereas contextual information was very important for development of vocabulary and extracting the meaning of the isolated word or phrase or statement. Therefore the incorporation of context and shape information in all the stages of OCR systems was necessary for meaningful improvements in recognition rates.

This review concludes that many promising methodologies were developed for OCR and HOCR in last few decades for Indian scripts like English (cursive) [26], Arabic [27], Gurumukhi [32], Persian[35], Devnagari [36], Tamil [44], Telugu [45], Bengali [46], and Kannada [47], Urdu [48], Gujarati [49] and Malayalam[50,51].

System architectures of OCR techniques for these studies were stated in Table 2. Although this work was done for different languages and can be considered for further enhancing of HOCR in MODI script.

Table 2. Existing OCR system architectures for various languages.

<table>
<thead>
<tr>
<th>Author and Reference</th>
<th>Language</th>
<th>Preprocessing</th>
<th>Segmentation</th>
<th>Feature Extraction</th>
<th>Classification and Recognition</th>
<th>Post Process</th>
<th>RR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amjad rehman et al. [26]</td>
<td>English (cursive)</td>
<td>Digitization, Noise filter and Thinning</td>
<td>Implicit and Explicit segmentation</td>
<td>Hybrid features with Statistical features and Profile projection features</td>
<td>ANN, Lexicon</td>
<td>79.23% to 80.91 for Segmentation</td>
<td>--</td>
</tr>
<tr>
<td>Lasie Dinggo et al. [27]</td>
<td>Arabic</td>
<td>Median filter, Baseline estimation, Binarization, Thinning</td>
<td>Holistic approach</td>
<td>Structural and statistical features</td>
<td>Decision tree and Probability function</td>
<td>Lexicon</td>
<td>82% to 99.22% for Arabic words</td>
</tr>
<tr>
<td>Gurpreet Singh et al. [32]</td>
<td>Gurumukhi</td>
<td>Morphological operations, Skew detection and correction, Noise removal</td>
<td>Zone based word segmentation</td>
<td>Distance Profile, Projection Histogram, Zoning density, Zernike moments and Hybrid feature</td>
<td>SVM, KNN</td>
<td>--</td>
<td>72.83% to 99%</td>
</tr>
<tr>
<td>Mir Mohammad Alipour [35]</td>
<td>Persian</td>
<td>8-bit Gray Scale image, Binarization, Gaussian filter, Spatial filter for Smoothing and Thinning</td>
<td>Horizontal and Vertical projection profile</td>
<td>Structural features, Statistical features and Global Transformations.</td>
<td>Decision tree and Holistic (Global) or Analytic strategies</td>
<td>--</td>
<td>72 % to 98% Average RR 90.56%</td>
</tr>
<tr>
<td>Gunvantinsinh Gohil et al. [36]</td>
<td>Devnagari</td>
<td>Binarization, Noise removal, Scaling, Skew detection &amp; correction, Hierarchical approach</td>
<td>Projection profile and graph based, Hierarchical approach</td>
<td>Binary features, Freeman Chain Code, PCA and FDA</td>
<td>ANN, SVM</td>
<td>--</td>
<td>54.10 to 76.90% for ANN 60.02 to 84.41% for SVM</td>
</tr>
<tr>
<td>C. Suresh Kumar and Dr. T. Ravichandran [44]</td>
<td>Tamil</td>
<td>Noise removal, Skew detection and correction, smoothing, foreground extraction by threshold, Skeletonization, Thinning</td>
<td>Horizontal and Vertical projection profile</td>
<td>Height, width of character, no of horizontal and vertical lines, no of circles, no of arcs, centroid, pixels and Position in region.</td>
<td>SVM, Self Organizing maps (SOM), Radial Basis Function (RBF), RACHSU and Fuzzy NN</td>
<td>--</td>
<td>100% for train set 97% for test set</td>
</tr>
<tr>
<td>Rama J. and B. Parvathavarthini [45]</td>
<td>Telugu</td>
<td>Skew detection and correction, Binarization</td>
<td>Horizontal and Vertical projection profile</td>
<td>Structural features</td>
<td>Feed Forward Back Propagation Neural Network</td>
<td>--</td>
<td>90%</td>
</tr>
<tr>
<td>Farhana Yeasmin Onue [46]</td>
<td>Bengali</td>
<td>Binarization, Noise removal, Skew detection &amp; correction, Page Layout</td>
<td>Horizontal and Vertical projection profile</td>
<td>Structural features</td>
<td>Decision tree, MLP, Kohonen NN, HIM</td>
<td>Spell checking &amp; editing</td>
<td>98%</td>
</tr>
<tr>
<td>H. Imran Khan et al. [47]</td>
<td>Kannada</td>
<td>Binarization, Thinning Noise removal</td>
<td>--</td>
<td>Chain code</td>
<td>k-Nearest Neighbor</td>
<td>--</td>
<td>100%</td>
</tr>
<tr>
<td>Tabassam Nawaz et al. [48]</td>
<td>Urdu</td>
<td>Binarization, Noise removal, Filtering</td>
<td>Projection profile</td>
<td>chain code, shape and visual properties</td>
<td>Template matching</td>
<td>Unicode</td>
<td>89%</td>
</tr>
<tr>
<td>Roja Shegla Sopita &amp; Vishnukumar Dhasad [49]</td>
<td>Gujarati</td>
<td>Binarization, SOM</td>
<td>--</td>
<td>Clusters and mean value</td>
<td>k-Nearest Neighbor</td>
<td>--</td>
<td>97.78% for consonants</td>
</tr>
<tr>
<td>Rosemol Emmanuel and Jiva George [50]</td>
<td>Malayalam</td>
<td>Binarization, median filtering, Skew correction, Sobel edge detection</td>
<td>Horizontal and Vertical projection profile</td>
<td>Template</td>
<td>Template matching</td>
<td>Linguistic rules</td>
<td>--</td>
</tr>
<tr>
<td>Jomy John et al. [51]</td>
<td>Malayalam</td>
<td>Noise Removal, Binarization, Edge detection, Normalization</td>
<td>Morphological operations</td>
<td>4 and 8 connected Chain code</td>
<td>A two layer feed forward neural network &amp; SVM</td>
<td>error correction &amp; mapping Unicode</td>
<td>61.2% to 72.1%</td>
</tr>
</tbody>
</table>

International Journal of Computer Applications (0975 – 8887) Volume 115 – No. 19, April 2015
9. CONCLUSION
Principal objective of this research work was to analyze different methodologies used in HOCR with special attention to MODI Script. Regardless of paucity of the script and availability of the documents various problems like color representation, degradations, noise, calligraphy and boldness of characters were produced and can be tackled in preprocessing stage.

Segmentation was most studied and most concern part in various HOCR. Critical problem was faced for Segmentation in case of MODI script due to its cursive nature, lot of variations in writing style, overlapping lines, overlapping characters, presence of erratic ‘shirorekha’. Even though, technology was developed, overlapping character and overlapping line segmentation remained mysterious for HOCR and OCR.

Plenty of feature extraction methodologies were invented and invention was continued in this sense. Feature extraction for MODI HOCR was in infancy stage as compare to the amount of work done in various Indian or non Indian scripts. This review concludes that only structural similarity approach and one pattern of zoning was insufficient for reliable and accurate implementation of HOCR in MODI Script as there was various Low level, Shape and Size based, Region based, Texture based, Moment based and Zone based feature extraction methods were available.

Lot of work was done for classification using ANN, SVM, KNN, Bayesian classifier, and Decision Tree with respect to Indian and non Indian script. Although SSIM, SVM, KNN and NN were discussed for classification of MODI characters, implementation strategy and results produced for different environment was still unclear. Distinct representations of same character and contiguous representation of distinct character caused serious problems in classification of handwritten MODI script.

According to Rajesh Khillari [52], Wikipedia [53], Sattkaryotak Sabha [54], The Times of India [55], The Hindu [56,57] and MID DAY INDIA NEWS [58] millions of MODI historical documents were preserved in Tanjavar’s Saraswati Mahal, Oriental manuscript section of Chennai’s Connemara University, Bharat Itihas Sanshodhan Mandal, Pune (BISM), Rajwade Sanshodhan Mandal, Dhule, museums in London, Paris, Spain, and Holland, South Asia, Europe, Denmark and other countries. These manuscripts were full of invaluable and priceless treasures from various fields like Literature, Science, Fine Arts, Ayurved, Pharmacy, Chemistry, Social Sciences, Psychology, Drawings, Paintings, Music, Astrology, History, Charms and Spells.

More scientific efforts were required to explore historical information stored in such a priceless treasure house by the development of HOCR for compound, conjunctions, numerals, and basic character sets of MODI script and various methods were yet to be implemented at different stages of HOCR for MODI script like segmentation, feature extraction, classification, recognition, and post processing.

10. FUTURE WORK
Research and development in MODI script was necessary to unfold the mysteries written on golden pages of history. An initiative was taken by Government of India for the development of Indian language by starting a program TDIL [59]. Main objective of this work was to study and develop various aspects of Indian languages.

CDAC, a well-known institute was also involved in the development of fonts and translators in collaboration with IIT Kanpur [60]. ISCII (Indian Scripts Standard Code for Information Interchange) [61] was developed by Indian government in 1988 for various Indian languages. Presently lot of room was available to do the research in the area of handwritten MODI script or character recognition form handwritten character documents or from handwritten MODI historical documents.

An example of handwritten MODI historical documents was shown in Figure 8.

Figure 8: Handwritten MODI historical document.

11. REFERENCES


