

## **A Systemic Study of Indian Ancient Script by determining the Issues for Effective use of Handwritten Character Recognition**

**Dr. Kirti Nilesh Mahajan**

*Professor, Bharati Vidypeeth Deemed to be University, Institute of Management & Entrepreneurship Development, Pune  
kirtimahajan75@gmail.com*

**Mr. Niket Tajne**

*Research Scholar, Bharati Vidypeeth Deemed to be University, Institute of Management & Entrepreneurship Development, Pune  
nikettajne22@rediffmail.com*

### **ABSTRACT:**

Massive amount of ancient documents are created by our ancestors. This all documents are published already on the digital libraries on single click in worldwide. Also this documents published in Paper Media. Transcription of this documents are either prepared by manually or by the help of digital technology. These all ancient Documents are already available in raw images which are very useful to investigate the ancient philosophies of any field. However these raw images to be transliterate in electronic format by adopting the some image processing algorithms for appropriate transcription of the ancient documents. The character recognition difficulty is meager high due to some causes. This Research Paper investigates the problems faced in digital transcription of ancient handwritten text documents. The Research paper is organized in Three Section the First Section highlighted the Introduction about ancient script like MODI script, Handwritten Character Recognition techniques. The Second Section is all about the Problems investigating of MODI Character and last section is highlighted the Concluding remarks of analytical research.

**Keywords :** Image Processing, Character Recognition, Handwritten Character Recognition techniques.

### **1. INTRODUCTION**

Indian content is the best example of diversity. This diversity reflected in the peoples living culture, food, and communication. In India at every miles of distance this variation of the culture is exemplified. For communication purpose the people used varies languages. Not only in the orally but writing wise also. These all spoken languages has some historical reference some has individually referenced and some has mixed of two or three languages. In this section of paper the historical background and its significance are discussed, also the introduction about the handwritten character recognition is deliberated.

The significance of transliteration is very much high. This kind of computation techniques is not only used for the converting the one script / language into another script / language but also open the new techniques, thoughts, to showcase the pathway for the newcomer to do the extraordinary work in the same field. It's also helpful for the each and every single person of the society. Transliteration sometimes work on the basis of the optical character recognition, handwritten character recognition, Neural Network, Pattern Recognition to achieve various objectives of specific work.

As the significance of ancient handwritten document are very much countless. The issues and problem of regarding these wonderful manuscript have to be discuses and find out appropriate elucidation

about the transliteration of particular manuscript in this paper the introduction about the ancient script called MODI script and the issues related with the handwritten character manuscript are deliberated.

## 1.1 HISTORY OF MODI SCRIPT

MODI script is one of the most popular Ancient scripts used in western and southern part of India. The use of MODI script is for the administrative and quotidian purpose. The principal meaning of MODI word is inspire by the Persian word “Shikasta” which means breaking in English language and the Meaning of MODI word in Marathi language, is “Modne” which means also breakdown in the parts[1][2]. There is some substantiation about the origin of the script, says that the proper MODI alphabet originated in the 12th century which was used for Administrative work, Education, and other routine activities which called as “Adyakalin”. Then in 13<sup>th</sup> century it was used for the same purpose which called as “Yadvakalin”. In next three century i.e. from 14<sup>th</sup> century to 16<sup>th</sup> century this script called as “Bahamanikalin” in this period the script majorly used in the administrative work. In the court of Maratha emperor Shivaji Raje Bhosale the script is rapidly use due to the cursive type of nature of MODI script. The MODI script used in the period of 17<sup>th</sup> century is the period of Maratha King Chatrapati Shivaji Maharaj hence the MODI script used in this period is called as “Shivakalin”. The next period of the MODI script is in period of “Peshwa”. The “Peshwa” used this script in the 18<sup>th</sup> century. Because the MODI script is used in “Peshwa” era the script name was “Peshwakalin”. The writing style major change happened in “Peshwakalin” MODI script that is the reason the script was categorized in four different segment naming as “Chitnisi”, “Bilavalkali”, “Mahadevpanti”, “Ranadi” these four categorize is based on the writing style of MODI script. The next period of MODI script is also naming as “Pashwakalin” and the time duration is from 18<sup>th</sup> century till the mid of 20<sup>th</sup> century[3][4][5]. The following Fig. 1 show the period of MODI script development according to their century.

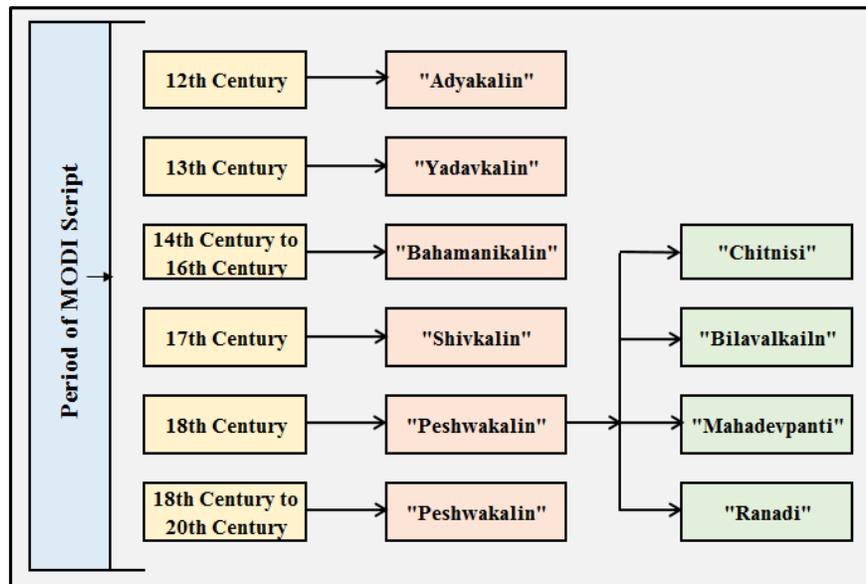


Fig. 1 Period of MODI Script with century

In period of 18<sup>th</sup> century till the mid of 20<sup>th</sup> century the area of MODI script utilization is not restricted to the precise area but it utilize in the Education, Healthcare, Ayurveda, Astrology, literature, Science, Fine Arts, , Pharmacy, Chemistry, social sciences, Psychology, Drawings, music, , History, and many more other fields. The manuscript of MODI script is available in so many museums. Not only in Indian cities, but in other countries of the world[5][6].

The MODI alphabet are variant of the Marathi Language and it has 14 vowels, 34 consonant letters and 10 Numbers[7][8]. Following Fig. 2, Fig 3, Fig 4 are show the MODI Vowels, MODI Consonants and MODI Numerals Respectively.

MODI Script Vowels						
अ	आ	इ	ई	उ	ऊ	ए
a	aa	i	ii	u	uu	E
ऐ	ऑ	औ	अं	अः		
ai	o	au	am	aha		

Fig. 2 MODI Vowels

MODI Script Consonants						
क	ख	ग	घ	ङ	च	छ
ka	kha	ga	gha	nga	ca	cha
ज	झ	ञ	ट	ठ	ड	ढ
ja	jha	nya	tta	ttha	dda	ddha
ण	त	थ	द	ध	न	प
nna	ta	tha	da	dha	na	pa
फ	ब	भ	म	य	र	ल
pha	ba	bha	ma	ya	ra	la
व	श	ष	स	ह	ळ	श्र
va	sha	ssa	sa	ha	lla	shr
ह्र						
dnya						

Fig. 3 MODI Consonants

MODI Script Digits						
०	१	२	३	४	५	६
shunya	ek	don	tin	char	pach	saha
७	८	९				
sat	aath	nau				

Fig. 4 MODI Numbers

The Significance of such a beautiful manuscript which was not only used for the administrative purpose but also it was used in every field like medical, food, arts, culture and also maintain the daily routine documents. Now a days these MODI script ancient documents are golden pages for every archaeologist and people who have an interest in ancient manuscript. In the between of 16th century till 20th century this script was very much popular and it's used everywhere. Due to the change in Social and Political views in the mid of 20th century in western part of India the 'devnagari' script naming as 'balbodh' was used everywhere as a primary language[4]. From that period of time the MODI script was gradually push out of its usage from

everywhere and from that time this beautiful arts of writing gets cease to exist. Fortunately there are many of MODI documents are preserved by the some states like Maharashtra, Goa, and Karnataka[9][10][11].

## 1.2 HANDWRITTEN CHARACTER RECOGNITION

Handwritten Recognition is the computer based technique to identify the received handwritten input from source format into the target format. Most of the time the handwritten character recognition system uses the image as a source so the sensed data is always offline. Handwritten Character recognition is the major type of Optical character recognition in the field of pattern recognition another type of optical character recognition is the Printed Character recognition. The following Fig. 5 show the different type of Optical character recognition[12][13].

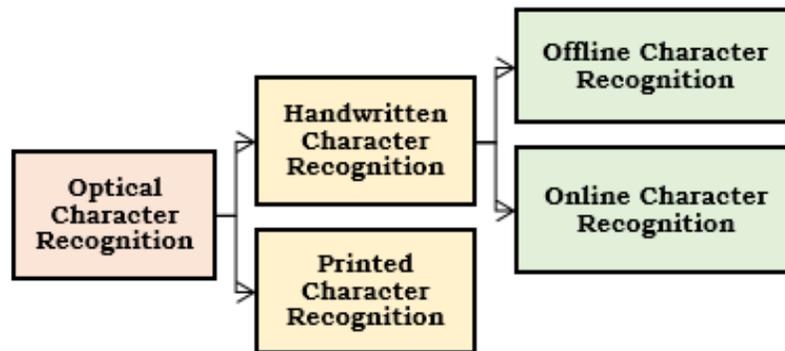


Fig. 5 Types of Optical Character Recognition

The Optical Character Recognition Techniques is the one of significant way to identify whether the Printed character or Handwritten character. Now a days the Optical character recognition is used in all domains like banking, healthcare, administrative, etc. The best advantage of the optical character recognition is to reduce the human efforts for data entry, storage of the document manually, etc. The handwritten character recognition system is categorized in to two parts and these are Off-Line Character Recognition and On-Line Character Recognition system respectively.

### 1.2.1 ON-LINE CHARACTER RECOGNITION SYSTEM

In the Handwritten character recognition system the On-Line Character Recognition System is used the digital gadgets like digital Pen, Touch Pad / Digital Pad, or any other digitalized gadgets from where the hand motion of the user can perceive. The On-line handwritten character recognition system is basically work on the 2-Dimensional data acquisition mechanism. Here any pointing device who is responsible to accept the string of coordinate's i.e.(x, y) is used. The On-Line Handwritten Character Recognition system is more accurate than the Off-Line Handwritten Character Recognition reason is that data acquisition of the user writing is dynamic by nature. The second reason is recognition of On-line happened more accurately is interactivity which means the tracing the error is very fast. The On-Line character recognition system is also very much fast because in On-Line character recognition system the unwanted bit removal process called 'Noise Removal' process is not consider because the data acquisition is happened with direct interaction through the devices not by the another way.

### 1.2.2 OFF-LINE CHARACTER RECOGNITION SYSTEM

In the Handwritten character recognition system the Off-Line Character recognition system is the second type. It takes the Raster images for the character recognition. In Off-Line character recognition the

data acquisition is done by scanned the handwritten document through the Computer Scanner or any other scanning device how scan the image from the piece of paper and then do the further process. The best advantage of Off-Line Character recognition system it is capable to immune the various writing styles to handle by it. Hence it takes the little more cost in terms of Processing and Time for the detecting the hand stroke of handwritten documents. Off-Line Character recognition system takes more time for progressing phase for Noise removal.

## 2. ISSUES IDENTIFICATION IN MODI SCRIPT CHRACTERS

Significant of the MODI script are huge. This ancient script is already available in the various research center, libraries, and in some universities. Due to enormous demand of the MODI script so its major requirement is Transliteration. Transliteration is the segment of Natural Language Processing (NPL). It is use full in the cross language information reclamation without changing the original essence of the source script / language. There are many problems in identification of MODI character like writing style of script, difference of MODI vowels, consonants, etc. these all are elaborated more properly below.

### 2.1 THE ISSUE OF WRITING STYLE OF MODI SCRIPT

The MODI script is by nature is cursive in nature that is the main reason were this language used in various department of the Maratha King Chhatrapati Shivaji Maharaja[14]. The MODI script also called as shorthand writing script because it write the character very much faster without pull up the pen from the paper[15]. The Major impact of the cursive nature of the MODI script is reflected not only on Marathi language but it also reflected on the Gujarati language too[16]. The following Fig. 6 demonstrate the nature of MODI script cursive type which was written by Maratha King Chhatrapati Shivaji Maharaja which is 344 years old found in the Capital of Maratha Kingdom called “Satara” in 2018[17].

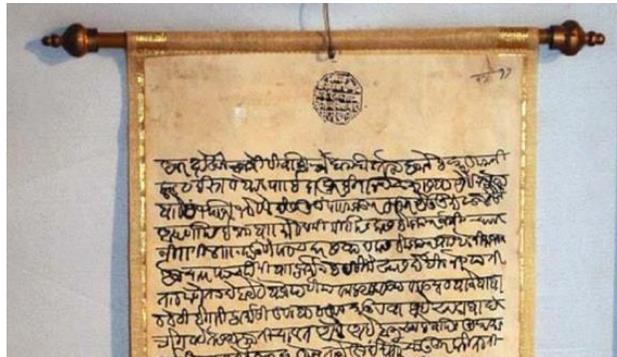
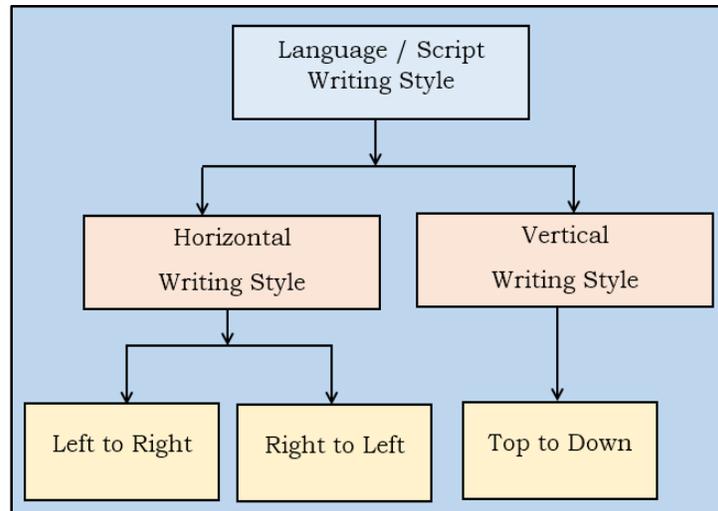


Fig. 6 Old Letter Written by Shivaji Maharaj

### 2.2 THE ISSUE OF SPECIFIC STRUCTURE OF MODI SCRIPT

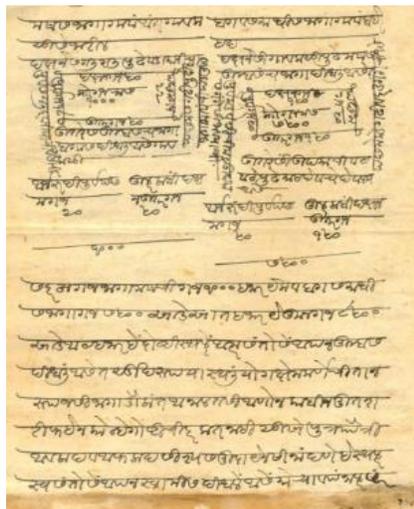
In the Literary Techniques the sentence arrangement and word arrangement is specific for the particular language that called as “Form”[18] were the writing for the particular language or script gives the proper structure of reader and writer also. On the basis of writing structure many languages in overall world writing style is from left to right which is one type of Horizontal writing style in this type the English, French, Italian, African, Some Indo Aryan languages, Japanese using hiragana and katakana etc. languages has the writing style from left to right. The second type of Horizontal writing style is right to left some languages has right to left writing style like Urdu, Arabic script, Persian script, Iranian languages. The second type of the writing style is Vertical writing style. In this writing style the Chines, Japanese, Korean languages writing style is top to down style following fig. 7 show the type of language writing style[19][20].



**Fig. 7 Type of Language / Script writing Style**

As we discuss the type of Language / Script writing style are categorized in to two major parts. The MODI manuscript many ancient document are demonstrate the type of horizontal writing style of left to right. But some documental evidence this statement as ambiguous because there is no any structure or style for writing the MODI manuscript. Following fig. 8 show the account related document in the period of Maratha Ruler[21].

The stone inscription written in MODI script was found in the Thanjavur located on temple. The stone inscription written in MODI script is constructed in the period of Kulothunga Chola and it is maintained by the Pandyas and Maratha rulers subsequently. Following fig. 9 show the stone inscription written in MODI script. Because it is written on Stone so embossed character are not displayed properly and the some character are embossed without any alignment[22].



**Fig. 8 MODI manuscript from the Maratha Ruler period**



**Fig. 9 Stone Inscription written in MODI script from the Kulothunga Chola period**

## 2.3 THE ISSUE OF NOISE IN MODI SCRIPT HANDWRITTEN DOCUMENTS

The Image noise is the combine blend of random variation of different color pixel information and brightness from the digital image. The digital noise or image noise is form between the process of degradation and restoration. If noise is exist in the digital image it produces unrequired effects in image such as blurred lines, dots, objects, unrealistic edges, and disturbs background scenes, color. The Noise in the existing digital images is the big issue for the transliteration and the other purpose of the data acquaintance. Noise in the images itself stated that the age of the documents and the availability and applicability of the digital content. The noise present in the digital content is deteriorating the quality of the digital images. In the original image pixel  $O(x,y)$  the Noise is mixed which is called as  $N(x,y)$  the resulting image is formed called as  $ON(x,y)$ [23][24].

$$ON(x,y) = O(x,y) + N(x,y)$$

Following are the some types of Noise model which are majorly belong to Photoelectronic, Impulse and Structured Noise of digital model[24][25]. These noises occurs in either, Binary images, Gray images or Color images.

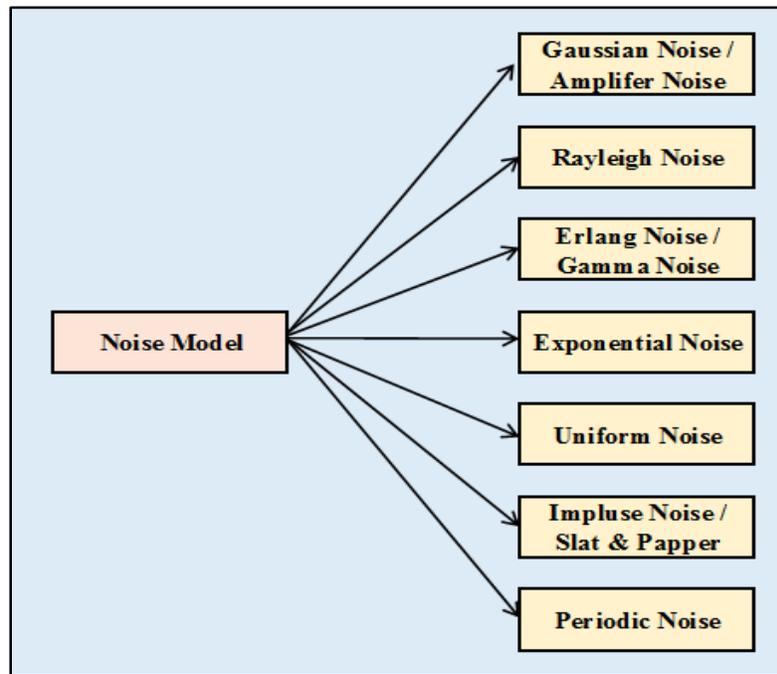
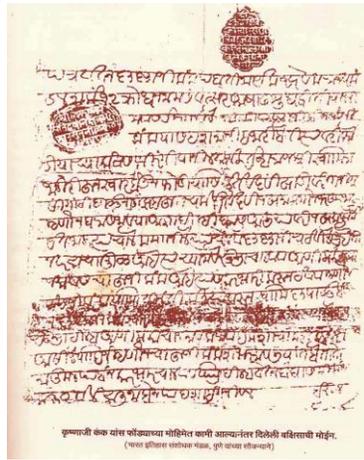
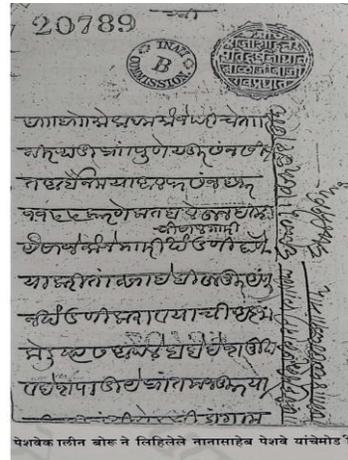


Fig. 10 Types of Noise Model exist in Digital Images

Following fig. 11 and fig. 12 show the existence of the Noise in the MODI script digital images. The Fig. 11 show the Amplifier Noise or Gaussian Noise structure and the Fig. 12 show the Impulse Noise called Salt & pepper noise.



**Fig. 11 Prize Winning letter given by Maratha King**



**Fig. 12 Letter written by Nanasahab Peshwa**

**2.4 The Issues of Writing style of MODI character**

The character arrangement, word formation play major role in the literature. Unfortunately character are misarranged the complete meaning of the sentence or rather context of the literature is changed. In every language / script the MODI script also has 14 vowels and 34 consonant but the some are similar in arrangement and writing also which change the total meaning of the sentence. Following are the category explicate the different variation of MODI script and similarity of MODI with Marathi and Gujarati language. Following Table 1 show the different writing style of the one MODI script handwritten character and Table 2 show the similarity between the MODI script, Marathi, Gujarati language.

Devanagari script	Writing Style - 1	Writing Style - 2	Writing Style - 3	Writing Style - 4	Writing Style - 5
स / Sa					
च / Ch					
ढ / Dh					
ख / Kh					
ण / Na					

**Table 1. Differentiation writing style of one letter of MODI Script**

Devanagari script	MODI Scrip	Marathi	Gujarati	Devanagari script	MODI Scrip	Marathi	Gujarati
ग		ग	ગ	घ		घ	ધ
छ		छ	છ	झ		झ	
ड		ड	ડ	ढ		ढ	ढ
ण		ण		त		त	ત
ध		ध	ધ	भ		भ	
श		श	શ	ष		ष	ષ
ळ		ळ		क्ष		क्ष	ક્ષ

Table 2. Similarity between MODI, Marathi and Gujarati

### 3. CONCLUSION

In this Research paper the significant of the MODI Script has been discussed, where different perspective of the MODI script are formed, arranged, and deliberated properly. In the Second half of the research paper the different mechanism of the Handwritten Character Recognition has been discussed and studies the various types of handwritten character recognition system. In the last section of the research paper the few problems found in the MODI manuscript for writing and transliterating which are based on different literature and evidence are discussed. The few problems of ancient scripts are basically highlighting the lacunas of MODI script why it is not been so popular even if it is so beautiful in writing and reading.

The discussed problem may give the help to the all fraternity of Computer Science and Application to resolve the problem of reading and transliterating the MODI Script. The Problem regarding the MODI script handwritten character is described with proper evidential background. It focus on the writing style of the MODI script and the some evidence show the unambiguous about the writing style and structure of the MODI script. It also explain the types of Noise exist in the Digital images. In the last Problem of the MODI handwritten character is, show the similarity between MODI, Marathi and Gujarati and it also show the one type of character written in different form in MODI script writing which some time unclear for the transliteration and Natural language processing. The differential style of the MODI script is defiantly the good area for the research. This research work is helpful to all people how wants to do the something extra ordinary work in the field of pattern recognition.

## REFERENCES

- [1] Mr. Rakesh A. Ramraje, "History Of Modi Script In Maharashtra", Global Online Electronic International Interdisciplinary Research Journal (GOEIJR), 2013.
- [2] Shubhangi Bhatambrekar and Niket Tajne, "Context free grammar for MODI script", International Conference on Computer Science, Data Mining and Mechanical Engg, 2015.
- [3] Sadanand A. Kulkarni, Prashant L. Borde, Ramesh R. Manza and Pravin L. Yannawar, "Recognition of Handwritten MODI Numerals using Hu and Zernike features", Research Gate, 2014.
- [4] Dr. Chatterjee Shekhar, "Designing a New Digital Font for Modi-Script", Ergonomics International Journal, 2018.
- [5] Kirti Nilesh Mahajan and Niket Pundlikrao Tajne, "Transliteration of Indian Ancient Script to Braille Script using Pattern Recognition Technique: A Review", International Journal of Computer Applications, 2017.
- [6] Pradeep Chakravarthy, "IIC India's Best-Kept Secret: The Sarasvati Mahal", The Occasional Publication series by India International Centre, 2010.
- [7] Anshuman Pandey, "Proposal to Encode the Modi Script in ISO/IEC 10646", Department of History University of Michigan Ann Arbor, Michigan, U.S.A, 2011.
- [8] Thakre Rajendra Bhimraoji, "Typography Day 2014 - Typography and Culture Reviving Modi-script", Research Gate, 2014.
- [9] Article by The Times of India, "Band of researchers, enthusiasts strive to keep Modi script", [online] Available : <https://timesofindia.indiatimes.com/city/pune/Band-of-researchers-enthusiasts-strive-to-keep-Modi-script-alive/articleshow/30761335.cms> [Accessed by Feb 21, 2014].
- [10] Article by The Times of India , "Decoding the Modi script: Lifting a veil on history", [online] Available : <https://timesofindia.indiatimes.com/city/goa/decoding-the-modi-script-lifting-a-veil-on-history/articleshow/61710711.cms> [Accessed by Nov 19,2017].
- [11] Article by Bangalore Mirror, "Karnataka to decode Modi, letter by letter", [online] Available: <https://bangaloremirror.indiatimes.com/bangalore/others/karnataka-to-decode-modi-letter-by-letter/articleshow/58367812.cms> , [Accessed by April 26, 2017].
- [12] Bunke Horst, Liwicki Marcus, "Recognition Of Whiteboard Notes: Online, Offline And Combination", Research Gate, 2008.
- [13] Surya Nath R S, Afseena S, "Handwritten Character Recognition – A Review", International Journal of Scientific and Research Publications, 2015.
- [14] Dr. Kirti Nilesh Mahajan1, Niket P. Tajne, "A Methodology to Analyze Handwritten MODI Character in Digital Image", International Journal of Creative Research Thoughts (IJCRT), 2018.
- [15] MODI script Tools by center for development of advanced computing(CDAC) [online] Available :[https://www.cdac.in/index.aspx?id=lu\\_modi\\_script](https://www.cdac.in/index.aspx?id=lu_modi_script)
- [16] Article by Times of India, "Dead Modi script comes alive in courts and coaching centres" [online] Available :

- <https://timesofindia.indiatimes.com/city/mumbai/Dead-Modi-script-comes-alive-in-courts-and-coaching-centres/articleshow/36574115.cms> [Accessed by Jun 16, 2014]
- [17] Article by The Better India, “Rare 344-Year-Old Letter Written by Shivaji Before His Coronation Found in Satara”, [online] Available : <https://www.thebetterindia.com/147177/news-history-letter-shivaji-coronation-satara/> [Accessed by June23, 2018]
- [18] The Ultimate HSC English Literary Techniques Cheatsheet by Maddison Leach, [online] Available : <https://www.artofsmart.com.au/literary-techniques/>
- [19] “English is read from left to right, but are some languages written from right to left or from top to bottom”, [online] Available : <https://www.dictionary.com/e/righttoleft/>
- [20] "Horizontal and vertical writing in East Asian scripts", [online] Available : [https://en.wikipedia.org/wiki/Horizontal\\_and\\_vertical\\_writing\\_in\\_East\\_Asian\\_scripts](https://en.wikipedia.org/wiki/Horizontal_and_vertical_writing_in_East_Asian_scripts)
- [21] Kulkarni Sadanand A. Borde Prashant L. Manza Ramesh R. and Yannawar Pravin L., “Review on Recent Advances in Automatic Handwritten MODI Script Recognition”, Internation Journal of Computer Application, 2015.
- [22] Article by Times of India, “Stone inscription of Maratha ruler found in Thanjavur temple” , [online] Available : <https://timesofindia.indiatimes.com/city/trichy/stone-inscription-of-maratha-ruler-found-in-thanjavur-temple/articleshow/68982861.cms> , [Accessed by Apr 22, 2019 ]
- [23] Priyanka Kamboj1 and Versha Ran, “A Brief Study Of Various Noise Model And Filtering Techniques”, Journal of Global Research in Computer Science, 2013
- [24] Article by Anisha Swain “Noise in Digital Image Processing” [online] Available : <https://medium.com/image-vision/noise-in-digital-image-processing-55357c9fab71> [Accessed by Aug 27, 2018].
- [25] Rafael C. Gonzalez, Richard E. by Prentice Hall Publication , “Digital\_Image\_Processing\_2ndEd”.