IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Analytical Study of Statistical Features Extraction of Characters for Verification of CAPTCHA in Devanagari Script

P.S. Bodkhe
Associate Professor
Dept. Of Computer Science
G.S. Science., Arts & Commerce.
College, Khamgaon

Dr. P.E. Ajmire
Head & Associate Professor
Dept. Of Computer Science
G.S. Science., Arts & Commerce
College, Khamgaon

Abstract :

Internet registration usually requires one sort of the Completely Automated Public Turing test to inform Computer and Human Apart (CAPTCHA) to avoid illegal registration through computer programs called bots. The CAPTCHA script is formed of distorted or noised characters. This paper presents an analytical study of extraction of varied statistical features for recognition of every character in CAPTCHA image. OCR (Optical Character Recognition) is generic process in area of pattern recognition and is aimed to recognize each character in CAPTCHA image. OCR system involves basic four stages: preprocessing, segmentation, feature extraction, classification which is followed by actual recognition. The technique of feature extraction plays an important role in character recognition. The CAPTCHA character recognition depends on the extracted features. The various statistical feature extraction techniques have been applied over character image. Here, the various types of statistical feature extraction methods are discussed for the application of character recognition for CAPTCHA verification.

Keywords: CAPTCHA, OCR, Bots, Feature extraction, Pattern recognition.

1.Introduction

In the recent time Internet has become a part of life for web users and CPATCHA has become an integral part of the communication and financial transactions on Internet. CAPTCHA (Completely Automated Public Turing test to tell Computer and Human Apart) are the challenges given by online service providers to its users to distinguish human and automated computer programs called bots. It is a defensive mechanism to prevent automated registration, spamming and web bots. CAPTCHA tests can protect web services and resources such as e-mail services, online voting, surveys, chat systems, grievance redressal system, and password authentication. CAPTCHA which acts as a gatekeeper which monitors the entry and check whether it is a process by human being or computer controlled automated system. The researchers are studying the techniques, how to keep these spammers away. The character images are used as CAPTCHA script. These images consist of random combination of characters. A randomly generated text image is taken as CAPTCHA image so that a human user can read it easily while a computer cannot. A number of internet sites like Gmail, FaceBook, Hotmail, Twitter, Yahoo and so on offer free registration to the users to guard against auto-registration, these websites has incorporated CAPTCHA so that all the registrations are done by humans and not by automated computer programs or bots.

OCR is employed for pattern recognition and is targeted to recognize each character in CAPTCHA image. In offline OCR system, raster image of character is taken as an input and then it is taken as an input and then it is processed. Recognition of character starts after generation of character. While in online system, (x , y) coordinates and pressure of electronic tablet is continuously measured on digital pad [1]. The general process for character recognition involves four main steps given in figure 1 [2].

Image Acquisition → Preprocessing → segmentation → **Feature Extraction** →

Classification → Post Processing → Character Recognition

Fig. 1 : General Character Recognition Process by OCR

Each character in CAPTCHA image has unique features that determine its shape, size and orientation. Thus, the technique of feature extraction plays a vital role in character recognition. The CAPTCHA characters recognition depends on the extracted features. There are various statistical features extraction methods which are studied and analyzed in this paper.

2.Devanagari Script

India is a multilingual country. There are more than 1.2 billion populations with 22 official languages and 10 different scripts. Devanagari, an alphabetic script, is used by a number of Indian Languages. Initially it was developed to write Sanskrit but was later used to write many other languages such as Marathi, Hindi, Konkani and Nepali and Rajasthani [3, 4]. Hindi is the third most spoken language in the world. Other Indian languages like Gujarati, Bengali and Punjabi use scripts similar to Devanagari. Devanagari has 34 consonants and 13 vowels. In addition to consonants and vowels, there are compound characters which are formed by combing two or more of basic characters in most of Indian scripts including Devanagari.

3. Feature Extraction Techniques

Feature extraction can be defined as a process of extracting the most representative information from the raw data. Feature extraction is the process of extracting different features from the matrices of digitized characters image. The characters are recognized on the basis of these features. Features of a character can be classified into two categories: Global or statistical features and Structural or topological features [5].

3.1 Structural Features

Structural features are based on topological and geometric properties of the character. Examples of structural features are: number of horizontal lines, number of vertical lines, number of cross points, number of end points, horizontal curves at the top and bottom etc. These features are then used to identify the shape or structure of the character.

3.2 Statistical Features

Statistical methods are used to identify statistical features of a character. The statistical features are derived from the statistical distribution of pixels. These are obtained from the collection of points representing the matrix of character. These features can be easily detected as compared to structural features. Statistical features are less affected by noise or distortions as compared to structural features. These features can be easily obtained as compared to topological or structural features [5].

Representation of a document image (e.g. CAPTCHA image) by statistical distribution of points takes care of style variations to some extent . Although this type of representation does not allow the reconstruction of the original image, it is used to reduce the dimension of the feature set which results in less computing time and low complexity. The major statistical features are used for character representation [6].

3.3 Statistical Feature Extraction Methods

The purpose of the statistical methods is to determine to which category the given pattern belongs There are number of methods used for statistical feature extraction; such as zoning, projection, histograms, crossing and distances, moments, profiling etc [7].

3.3.1 Zoning

Zoning is one of the most popular feature extraction methods. The character image is divided into predefined number of zones and a feature is computed from each of these zones. The frame containing the character is divided into several overlapping or non-overlapping zones and the densities of object pixels in each zone are calculated. The density is calculated as a number of object pixels in each zone divided by total number of pixels [8].

In this method of zoning , four types of statistical features are extracted , namely , intersection and open endpoints, centroid, horizontal peak extent and vertical peak extent, for Devanagari ancient character recognition. A pixel that has more than one pixels in its neighborhood is known as intersection point. A pixel that has only one pixel in its neighborhood is known as open endpoint. The centroid is the point that can be considered as the centre of a two-dimensional image. A character is usually divided into zones of predefined size. These predefined sizes are typically of the order 1*1, 2*2, 4*4 , 8*8 is given in figure 2 [9].

Fig. 2 : A character image : a) one zone b) 4 zones c) 16 zones and d) 64 zones.

3.3.2
Histogram Projections

These features are derived from histograms of horizontal and vertical projections of black pixels in some particular areas of the character. They are extracted from the normalized image of the character so as to obtain normalized histograms of black pixels both on the X-axis as well as on the Y-axis. In order to locate the larger strokes of the character, relative maxima of the histograms could have been extracted. However, extracting maxima from histograms is quite difficult when the number of maxima is not known a priori. One way to tackle this problem is to derive the cumulative histogram of the original one, assuming that each important gap on the Y-axis of the cumulative histogram corresponds to a relative maximum on the Y-axis of the original histogram. Yet, when there is a gap on the Y-axis of the cumulative histogram, the corresponding abscissas are closer than in the case of a slow variation on the Y-axis. Therefore, dividing the Y-axis in enough equal parts and storing as features the corresponding abscissas allows to get information about the location of some structural properties of the character like its larger strokes [10].

The feature extraction principle is illustrated on a basic example in Figure 3 and 4. In the present case, dividing the Y-axis in 11 parts – 10 abscissas are thus retained as features from each cumulative histogram – has been found experimentally to provide the most precise description of the characters [10].

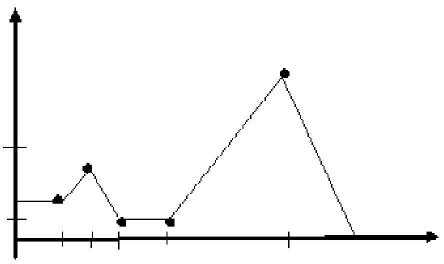


Fig. 3 : An original histogram of projection

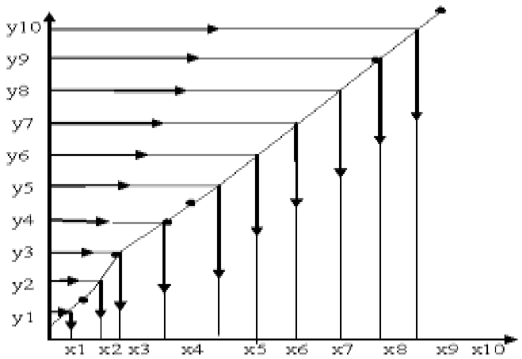


Fig. 4 : The Cumulative Histogram

3.3.3 Crossing and Distances

It refers to the number of crossings of a contour by a line segment in a specified direction. Distance of line segment from a given boundary can be used as one of the features. A horizontal threshold can be established above, below and through the centre of the script. The feature value is the count, the number of times the script crosses the threshold [11]

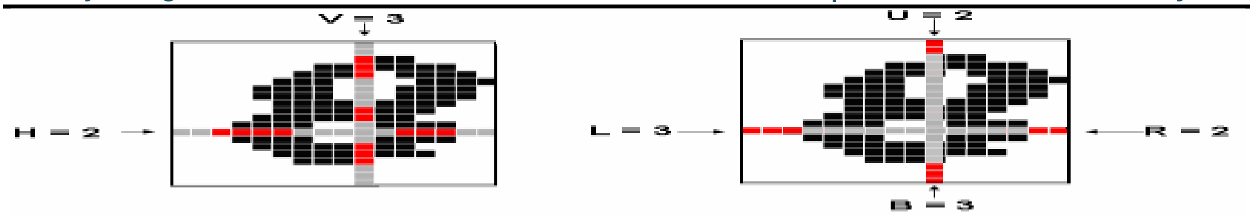


Fig. 5 : Example showing Crossing and Distance

3.3.4 Moments

The moment invariant is an appropriate measure for tracing the noise free image pattern. The features that are invariant to scaling, translation and rotation transformation of a character are useful in order to recognize many variations of the same character. Invariance means that visual properties stay consistent under specific transformations. Rotation invariant should be used to distinguish between such characters as ‘u’ and ‘n’. Size invariant should be used to distinguish between such characters as ‘0’ and ‘o’ [12]. A set of seven 2-D 72 moment invariant features are insensitive to translation, scale change, mirroring, and rotation. These can be derived from seven equations. It computes the moment invariants of an image and obtained seven-element row vector. Using these features a high degree of invariance is achieved. Hu's Seven Moment Invariants are invariant under translation, changes in scale, and also rotation. So it describes an image irrespective of its location, size, and rotation. The 2-D moment of order (s + t) of a digital image f(x ,y) of size m × n is define as follows.

$$m_{st} = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} x^s \, y^t \, f(x,y) \quad \text{where } s = 0,1,2\ldots \text{and } t = 0,1,2\ldots \text{are integers.}$$

The corresponding central moment of order (s + t) is define as follows.

$$H_{st} = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (x - \bar{x})^s (y - \bar{y})^t f(x,y) \quad \text{where } s = 0,1,2\ldots \text{and } t = 0,1,2,\ldots.$$

Where $\bar{x} = m_{10} / m_{00}$ $\bar{y} = m_{01} / m_{00}$

The moments that are computed with their centroid being about the origin are called central moments, denoted by μ_{st}.

The normalized central moment of order (s + t) is define as follow.

$$\eta_{st} = \mu^{\gamma}_{00} \quad \text{where} \quad \gamma = (s + t) / 2 + 1$$

The set of seven moment invariants are derived from these equations, as follows:

$\Phi 1 = \eta_{20} + \eta_{02}$

$\Phi 2 = (\eta_{20} - \eta_{02})^2 + 4\,\eta^2_{11}$

$\Phi 3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$

$\Phi 4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$

$\Phi 5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})$
$[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$

$\Phi 6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})_2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$

$\Phi 7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})]$
$(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$

The seven features of moment invariant are extracted from an image. The value of moments are usually small, so to reducing their dynamic range, log10 transformation is used and sign of original value is also preserve to detect if an image has been mirrored. Figure 6 shows the seven features were extracted from the handwritten characters. Three same characters with different scale, orientation, and transformation, give almost same moment features. Figure 7 shows the features extracted from the printed characters. Examples of same character with different sizes computed almost same moment features. From the experiment it was cleared that Hu’s moments features were invariant with scale in printed characters and invariant in rotation for handwritten characters [12].

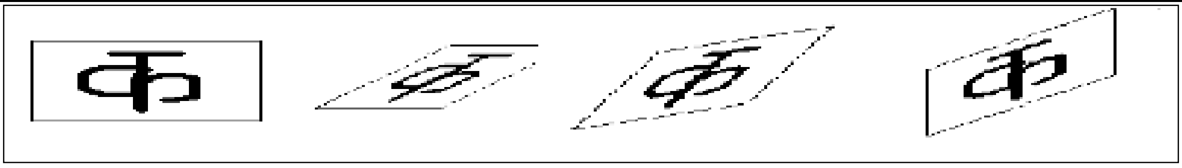


Fig.6 : Moment Invariant features for handwritten characters

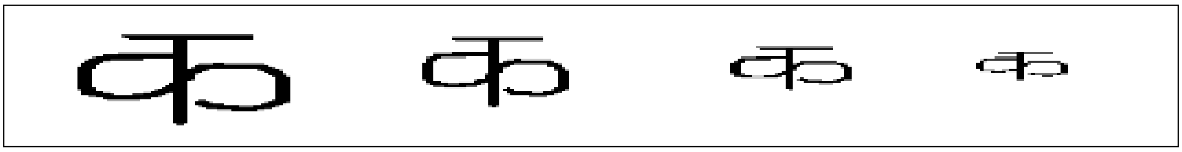


Fig. 7 : Moment invariant features for printed characters

3.3.5 Zernike Moments

Moment functions are defined on images as the weighted sums of the image intensity function. Moment functions of order (p + q) are generally defined as

$$\Phi_{pq} = \int_x \int_y \Psi_{pq}(x, y)\, f(x, y)\, dx\, dy, \qquad (1)$$

Where $\Psi_{pq}(x, y)$ is called the moment weighting kernel.

When applying moment functions to digital images it is often desirable to write them out using the following discrete notation:

$$\Phi_{pq} = \sum_x \sum_y \Psi_{pq}(x,y) f(x,y) \qquad (2)$$

Some properties of the weighting kernel are passed onto the moments themselves, such as invariance features, and orthogonal. Depending on the function chosen for the weighting kernel, the calculated moments can capture different aspects of the input image [13].

Zernike moments are a class of orthogonal moments and have been shown effective in terms of image representation. The orthogonal property of Zernike polynomials enables the contribution of each moment to be unique and independent of information in an image. A Zernike moment does the mapping of an image onto a set of complex Zernike polynomials. Zernike Moments are defined over the unit disk instead of the real plane and exhibit the orthogonal property. Zernike polynomials are mainly used in optometric, where they arise as the expansion of a wave front function in optical systems with circular pupils [14]. Zernike introduced a set of complex polynomials which form a complete orthogonal set over the interior of the unit circle,

i.e., $x^2 + y^2 = 1$. Let the set of these polynomials be denoted

by. $\{V_{nm}(x, y)\}$ The form of these polynomials is:

$$V_{nm}(x, y) = V_{nm}(\rho\ \theta) = V_{nm}(\rho)\exp(jm_\theta) \qquad (3)$$

$R_{nm}(\rho)$ Radial Polynomial defined as

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|/2} (-1)^s \cdot \frac{(n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n+|m|}{2}-s\right)!} \rho^{\,n-2a} \qquad (4)$$

Note that

$R_{n,-m}(\rho) = R_{nm}(\rho)$.

These polynomials are orthogonal and satisfy

$$\iint_{x^2+y^2\le 1} [v_{nm}(x,y)]\ V_{pq}(x,y)\ dx\, dy = \frac{\pi}{n+1}\delta_{np}\,\delta_{mq}.$$

With

$$\delta_{ab} = \begin{cases} 1 & a = b \\ 0 & otherwise \end{cases} \qquad (5)$$

Zernike moments are the projection of the image function onto these orthogonal basis functions. The Zernike moment of order n with repetition m for a continuous image function f(x,y) that vanishes outside the unit circle is

$$A_{nm} = \frac{n+1}{\pi}\iint_{x^2+y^2\le 1} f(x,y)\, V^*_{nm}(\rho,\theta)\, dxdy \qquad (6)$$

For a digital image, the integrals are replaced by summations to get.

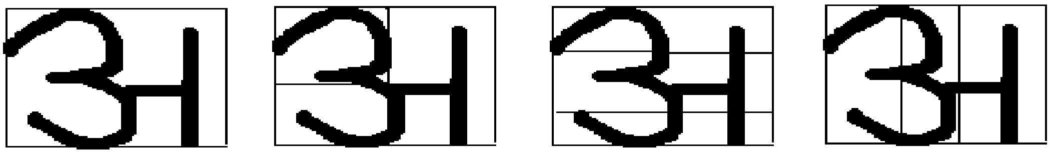
$$A_{nm}=\frac{n+1}{\pi}\sum_x \sum_y f(x,y)\ V^*_{nm}(\rho,\theta),\ x^2+y^2<=1 \qquad (7)$$

To compute the Zernike moments of a given image, the center of the image is taken as the origin and pixel coordinates are mapped to the range of unit circle, i.e $x^2+y^2 <= 1$. Those pixels falling outside the unit circle are not used in the computation. Therefore $|A_{nm}|$ can be used as a rotation invariant feature of the image function. Since $A_{n,-m} = A_{nm}$, and therefore

$| A_{n,-m} = A_{nm} |$, we can use only $| A_{nm} |$ for features[15].

Moment based features are extracted from the each zone of the scaled character bitmapped image. The image is partitioned into zone and features are extracted from each zone. In this paper Zernike moments based feature extraction is proposed for off-line Devnagari Handwritten Basic and Compound Character. The image can be portioned in various ways and it is given in figure 8 [16].

Feature set i : Figure 8 (i) is considered as a whole character image.
Feature set ii : Figure 8 (ii) shows the image divided into four equal zones.
Feature set iii : Figure 8 (iii) shows the image divided into three vertical equal zones.
Feature set iv : Figure 8 (iv) shows the image divided into three horizontal equal zones.



(i) (ii) (iii) (iv)

Fig. 8. Partition of Devanagari Character into feature set

3.3.6 Profiling: Distance Profile Feature

The profile based features are partly inspired by (Shridhar and Badreldin, 1984), in this technique, profile computes the distance (pixels) from bounding box of character image to outer image of character. The distance is traced horizontal and vertical. Author used four profiles of the character which are left, right, bottom and top. Left profiles are computed by traversing horizontally in forward direction and right profiles are computed by traversing backward direction. Similarly, top and bottom profiles are computed. Features derived from profiles of the character [17]. The four profiles are shown in figure 9.

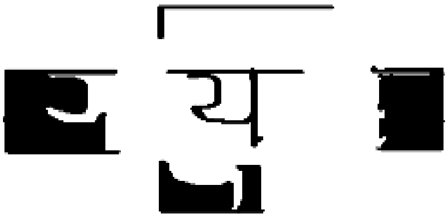


Fig.9. Shows the distance profile features

5. Earlier Work

Feature extraction methods can satisfy this requirement based on specific recognition problem and the input data for recognition. Different feature extraction methods are used according to the suitable domain. A survey of various feature extraction techniques is given in [12, 18]. In [19], Raghuraj et al, presented a scheme to develop complete OCR for five different fonts and sizes of Devanagari characters. Authors used three statistical features: mean distance, histogram of projection based on the spatial position of pixels and the histogram of projection based on pixel value. They used artificial neural network (ANN) approach for classification. Holambe et al [20] presented a review of feature extraction and selection methods for recognition of numerals and characters of the Devanagari script. They used Zernike moment for feature extraction and - KNN classifier based on Euclidean distance. They computed the accuracy recognition rate for vowels, consonants without modifiers as 98%. Yadav et al [21] , proposed an OCR system for printed Hindi recognition, using ANN. They used projection profiles for segmentation and histograms of projection based on mean distance, histogram of projection based on pixel value, vertical zero crossing for feature extraction and back-propagation neural network with two hidden layers for classification. They achieved a recognition rate of about 90.0%. Sethi and Chatterjee [22] proposed a decision tree based approach for recognition of constrained hand printed Devnagari characters using primitive features. In [23], R. Jayadevan et al. encompassed the comparative study of different feature extraction methods proposed by various researchers. They also have discussed some statistical feature extraction methods for recognition of printed Devanagari characters which are summarized in the following table 2 :

Table 2 : Accuracy of Devanagari Printed characters Recognition

Sr .No.	Method proposed by	Type of Feature	Classifier	Data set (size)	Accuracy in %
1	Jayanthi et al. [25]	Statistical	Binary Tree	4863 characters	95.08
2	Choudhuri et al. [26]	Statistical	Tree classifier & template matching	10000 characters	95.42
3	U.Pal et al. [24]	Statistical	Tree classifier and template matching	Unspecified	83.67
4	Divakar et al. [27]	Mean distance, Pixel value and vertical zero crossing	Neural Network	1000 characters	98.5
5	Holambe et al [20]	Zernike moment	KNN classifier based on Euclidean distance	Unspecified	98
6	Divakar et al [27]	Histogram of projection based on mean distance and pixel value , vertical zero crossing	Back-propagation neural network with two hidden layers	650 : word level	90

6. Conclusion:

The analytical study of the methods mentioned in Table 2 concludes the statistical methods used for feature extraction such as: zoning, histogram projections, profiles, crossing and distances, moment invariants, standard deviation, density and mean and feature classification techniques such as SVM, Template matching, Binary tree and Neural Network (KNN) are used for recognition of character. The devanagari CAPTCHA script characters are generally presented with different dimension and size, slanted or rotated to left or right direction, distorted by addition of noises. The statistical feature extraction methods such as Moment invariant and Zernike moments are mostly used to extract the features of characters which are slanted or rotated and dimensioned with varied size. These features can be easily detected as compared to structural features. Statistical features are less affected by noise or distortions as compared to structural features. They are also used to reduce the dimension of the future set. These features can be easily obtained as compared to topological or structural features.

From the above table the accuracy of Devanagari printed characters is 98.5% for Neural Network. The accuracy depends on the features extracted from the image database. In this study, it is found that the use of statistical methods for feature extraction and Neural Network classifier gives better result.

The study shows that the accuracy of recognition of character depends on selection of features and proper classifier. The different statistical feature extraction techniques discussed in this paper are very effective and useful for new researchers. Still the work is going on to improve the accuracy of the above techniques.

7. References

[1] Rekha Teraiya , Gunvantsingh Gohil and Mahesh Goyani “ Chain Code And Holistic Features Based OCR System For Printed Devanagari Script Using ANN And SVM ”, International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.1, January 2012.

[2] Puja Ujwal Talole and P.E. Ajmire ,“ Comparative Study of Feature Extraction and Classification Techniques for Handwritten Devanagari Script”, International Journal of Advanced Research in Computer and Communication Engineering, ISO 3297:2007 Certified, Vol. 6, Issue 2, Feb. 2017.

[3] Sandip Kaur, “Recognition of Handwritten Devanagari Script using Feature Based on Zernike Moments and Zoning and Neural Network Classifier”, A M. Tech. Thesis Report, Panjabi University, Patiala, 2004, pp.

[4] Gaurav Jain, Jason Ko, "Handwritten Digits Recognition", Multimedia Systems, Project Report, University of Toronto, November 21, 2008, pp. 1-3.

[5] Ms. Smita Ashokrao Bhopi and Mr. Manu Pratap Singh , "Feature Extraction Techniques for Marathi Character classification using Neural Networks Models", International Journal on Future Revolution in Computer Science & Communication Engineering, ISSN: 24544248,Volume: 4 Issue : 6, Pages : 70-76.

[6] Vikas J Dongre and Vijay H Mankar, "A Review of Research on Devanagari Character Recognition", International Journal of Computer Applications (0975-8887),Volume 12-No.2.November 2010.

[7] Purna Vithlani and C. K. Kumbharana," Structural and Statistical Feature Extraction Methods for Character and Digit Recognition", International Journal of Computer Applications (0975 - 8887), Vol.120-No.24, June 2015.

[8] Pritpal Singh and Sumit Budhiraja, "Feature Extractication and Classification Techniques in O.C.R. System for Handwritten Gurumukhi Script Survey", International Journal of Engineering Research and Applications, vol. 1, Issue4,pp. 1734-1739, August 2012.

[9] Sonika Narang et al., " Devanagari ancient documents recognition using statistical Feature extraction techniques ", Sådhanå (2019) 44:141, copy Right -Indian Academy of Sciences.

[10] L.Heutte, T. Paquet, J.V. Moreau, Y. Lecourtier , C. Olivier , "A structural/statistical Feature based vector for handwritten character recognition ",Pattern Recognition Letters19 1998 629–641, Received 30 March 1996; revised 1 December 1997.

[11] Gaurav Y. Tawde, Mrs. Jayashree M. Kundargi, " An Overview of Feature Extraction Techniques in OCR for Indian Scripts Focused on Offline Handwriting" ,International Journal of Engineering Research and Applications (IJERA) , ISSN: 2248-9622 www.ijera.com Vol. 3, Issue 1, January - February 2013, pp.919-926.

[12] Anil K Jain, Qinvind Due Trier and Torfinn Taxt, "Feature Extraction Methods For Character Recognition – A survey", Pattern Recognition, vol. no. 4,pp. 641- 662 , 1996, Elsevier Science Ltd.

[13] Y. S. Abu-Mustafa and D. Psaltis, "Image normalization by complex moments," IEEE Trans. Pattern Anal. Machine Intell. vol. PAMI-7, no. 1, pp. 46-55, Jan. 1985.

[14] I.K. Sethi, B. Chatterjee, "Machine recognition of constrained hand printed Devanagri Numerals", J. Inst.Electron. Telecom. Eng. 22 (1976) 532-535.

[15] Prashant S. Kolhe & S. G. Shinde, "Devanagari OCR Using Knn and Moment", International Journal of Computer Science and Engineering (IJCSE) ISSN 2278-9960 Vol. 2, Issue 2, May 2013, 93-100 © IASET.

[16] Karbhari V. Kale, Prapti D. Deshmukh, Shriniwas V. Chavan, Majharoddin M. Kazi, Yogesh S. Rode, " Zernike Moment Feature Extraction for Handwritten Devanagari (Marathi) Compound Character Recognition ", (IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 3, No.1, 2014.

[17] Arjun Singh , Kansham Angphun Maring, "Handwritten Devanagari Character Recognition using SVM and ANN ", International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 8, August 2015. DOI 10.17148/IJARCCE.2015.4825.

[18] Govindan, V.K. Shivaprasad, A.P. "Character Recognition-A Review " , *Pattern Recognition. Vol. 23 N0. 7*, pp 671-683, 1990.

[19] Raghuraj S, Yadav C S, Verma P and Yadav V 2010 "Optical Character Recognition (OCR) for printed Devanagari script using artificial neural network". Int. J.Computer Science & Communication 1(1): 91–95.

[20] Holambe A N, Thool R C and Jagade S M, 2011," A brief review and survey of feature extraction methods for Devnagari OCR". In: Proceedings of the 9th International Conference on ICT and Knowledge Engineering. 99–104.

[21] Yadav D, Sa´nchez-Cuadrado S and Morato J 2013, " OCR for Hindi language using a

Neural Network approach “. J. Inf.Process. Syst. 9(1): 117–140.
[22] I.K. Sethi, B. Chatterjee,“Machine recognition of constrained hand printed Devanagari Characters”, Pattern Recognition 9 (1977) 69-76.
[23] R. Jayadevan et al. , " Offline Recognition of Devanagari Script : A Survey ", IEEE Transactions on system, Man And Cybernetics-Part C: Applications And Reviews. vol. 41, no. 6, November 2011.
[24] U. Pal and B.B.Choudhury, “Printed Devanagari script OCR system”, vol. 10, pp. 12-24, 1997.
[25] K. Jayanthi, A. Suzuki, H. Kanai, Y. Kawasoe, M. Kimura, and K. Kido, “Devanagari Character recognition using structure analysis,” in *Proc .IEEE-TENCON*, 1989, pp. 363 – 366.
[26] B. B.Chaudhuri and U. Pal, “An OCR system to read two Indian language scripts: Bangla and Devanagari,” in *Proc. 4th Conf. Document Anal.Recognit.* 1997, pp. 1011–1015.
[27] Yadav D, Sa´nchez-Cuadrado S and Morato J 2013, “ Optical Character Recognition For Hindi Language Using a Neural network Approach “ . J. Inf.Process. Syst. 9(1): 117–140.