Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, No. 3, August : 2023

HANDWRITTEN GUJARATI NUMERALS RECOGNITION USING DEEP LEARNING MODEL

Dr. Shailesh Chaudhari, Dept. of ICT, Veer Narmad South Gujarat university, Surat (Gujarat), India

Abstract
Handwritten character identification, both offline and online, has grown to be a vibrant and difficult area of research as a result of the advent of machine learning, pattern recognition, computer vision, and deep learning algorithm. Contemporary optical character recognition (OCR) systems place a high value on handwritten character recognition. It is very difficult to identify and recognize correct meaning of handwritten document as different peoples have different ways of writing style. A remarkable research work has been done for Indian and Non-Indian scripts. Gujarati script recognition is quiet difficult due to its large character sets with curvature shape and complexities. This article deals with the handwritten Gujarati digits. There is no benchmark data set is available so a train and test data set is created with the help of more than 200 hundred people. Here, a Deep Convolutional Neural Network (DCNN) is used for training and classification of handwritten Gujarati digits and 98.4% accuracy is achieved.

Keywords: Deep Learning, Convolution, Pooling, Data set.

I. Introduction

OCR is a group of computer vision problems that involves converting handwritten text from digital images into machine-readable text. Many studies utilized several conventional computer vision methods to address the OCR issue, including image filters, contour detection, and picture categorization. These techniques worked well on small, template-based datasets with little variation in the orientation, image quality, etc., but new approaches need to be investigated to make the system more resilient to these variations.

Presently, last few years Deep Learning approaches have been extensively used and improved. Convolutional Neural Networks (CNNs) are particularly crucial for numerous applications [1]-[7]. In these, applications CNN takes images as input and learns several features for classification task. The features obtained by CNN is invariant to translations and distortions. Deep Convolutional Neural Networks provides nearly perfect accuracy for these application along with handwritten text.

Today, handwritten writing is largely divided into online and offline categories. They are sufficiently intelligent but fall short of the threshold of human pervasiveness. Applications like data input [8], office automation [9], postal address reading [10], printed postal code [11], cheque verification systems [12], etc. all heavily rely on handwritten character recognition systems.

Gujarati is one of 22 scheduled languages of the Union in India and is official communication language of Gujarat state. It is spoken by 50 million peoples through out the world. The Gujarati language having large character set and also shape of characters is curvature in nature, which makes the character recognition quite complex as compared to other languages. The Gujarati Numerals character set is shown in Figure 1. The recognition of Gujarati numerals written by hand is the main topic of this paper. Figure 2 displays some illustrations of handwritten Gujarati numbers.

૧ ૨ ૩ ૪ ૫ ૬ ૭ ૮ ૯ ૦

Figure 1 Gujarati Numerals Set

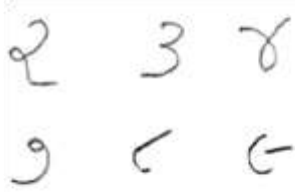

Figure 2 Handwritten Gujarati Numerals

II. Literature Review

OCR has been a thriving study field for a few decades and continues to draw new researchers. Although high level research work has been done but same is not true for many Indian languages.
The topic of handwritten Gujarati numeral recognition is covered in this article. This paper uses research done in other Indian languages as a point of reference. The majority of the time, OCR can tell a printed document from a handwritten one.
The first ever attempt for OCR made in 1977 by rajasekaran et. al. [13]. They used steps of character as features to identify 50 primitive Telugu characters. they used decision tree classifier for the classification purpose. Thereafter Sinha et. al. presented work for devanagari script in 1979 [14]. They used structural features of devanagari script In 1984 Ray et. al.[15] presented nearest neighbor classifier based Bangla character recognition system. In 1998, Chaudhari et al.[16] reported a work that used decision tree classifiers and template matching to detect written Bangla letters.
The first ever work for printed Gujarati characters was presented by Antani et. al. in 1999 [17]. To categorise the printed Gujarati characters, they employ hamming distance and Euclidean distance. A prototype to recognise printed Oriya characters based on stroke properties was put up by Chaudhari et. al. in 2002 [18]. They make use of the idea of water or reservoir flow. In 2003 Laxmi et. al. [19] presented a system to recognize music singers of printed Telugu language. They used moment features and k- nearest neighbor algorithm. In 2005 Dholakia et. al. [20] presented their work for printed Gujarati text. They use horizontal and vertical profile to segment line and word. A technique to recognize the printed South Indian scripts Kannada, Telugu, Malayalam, and Tamil was presented by Manjunatharadhya et. al. in 2008 [21]. In order to highlight the current methods for accurately extracting text from printed images, Saxena et al. [22] presented a comparison study and review report in 2018.
From the literature a very little work is found for Indian languages as compared to printed text. A technique to recognise Tamil hand-printed characters was reported by Chinnuswami et al. in 1980 [23]. For the classification, they employed statistical features and strokes. In 1993, Datta et al.'s[24] research to identify printed and handwritten Bangla characters was reported. They used features like curvature maximum temperature minimum and inflection points. Banashree et. al. [25] in 2007 recognized handwritten Hindi digits using a diffusion algorithm. They used a neural network to classify isolated digits.
In 2008 Rajasekaradhya et.al. [26] proposed zone and centroid based technique for 4 major South Indian languages. In 2009 Shanthi et.al. [27] suggested SVM for and return Tamil characters. They use image subdivision for feature extraction. Nemanath et al. [28] In 2021, proposed an adaptive strategy for offline paragraph recognition that involved pre-processing and sequential CNN and RNN training on the data set.

III. Proposed Methodology

Deep learning is sub-field of Machine learning which mainly concerns with algorithms. It implements neural networks for picking up useful features directly from data without manual intervention. Using training data, the network trains itself to extract specific features, which is used to solve complex problems.It is a particular kind of neural network with multiple hidden layers, and its neurons function by performing a unique linear operation known as convolution. The shape of one function is altered by the other through a mathematical process known as convolution, which results in the creation of a third function.

Another biologically inspired way for the visual cortex to approximate the capabilities of the human visual system is using deep learning approaches, in particular deep convolutional neural networks (DCNN). With the input data, DCNN can easily extract features. To identify the object in an unlimited context, the convolution and pooling layers make use of the information that was retrieved.

3.1 Convonution Layer

CNN is a deep learning model that is inspired by visual images [29, 30] and is made to automatically and accommodatingly learn structural relationship of characteristics. The use of a computer to analyse data is becoming more and more common. A typical CNN is made up of three different types of layers (or "building blocks"): convolution, pooling, and fully connected layers. Convolution and pooling, the first two layers, extract the features, while the fully connected layer, the third layer, maps the features into the output. A crucial component of CNN, a specific kind of linear computation, is a convolution layer. Because a feature could be anywhere in a digital image, pixel values are combined with kernel, which makes CNN extremely effective for processing images. Extracted features may gradually get more complex as the feature mappings from one layer to the next increase in complexity. The back-propagation and gradient descent optimization process is used to reduce the discrepancy between outputs and ground truth labels.

3.2 Pooling Layer

In order to decrease the dimensionality of the feature maps and, as a result, the amount of learnable parameters, this layer offers a widely used down sampling technique. In spite of the fact that filter size, stride, and padding are hyperparameters in pooling operations and are comparable to convolutional processes, there isn't a learnable parameter in any of the pooling layers.

Max pooling is then used to extract patches from the input feature maps, which returns the highest value from each patch while rejecting all other values. In actual use, a max pooling with a 2 x 2 filter and a 2 stride maintains the depth dimension of feature maps.

3.3 Fully Connected Layer

Similar to a traditional neural network, this layer is made up of distinct neurons for each pixel. The previous convolution or pooling layer's output feature maps are often flattened or converted to a one-dimensional format (1D). It is also referred to as dense layers, in which there are as many neurons as expected classes. Here, the "Dropout" regularization technique is applied to reduce over-fitting at the fully-connected layer. Dropout is a regularization method that renders learning insensitive by randomly removing some neurons from the system.

3.4 CNN Architecture

There are main two tasks of CNN structure : the extraction and classification of features. Each node of the convolutional layer is subjected to a convolution operation in order to extract its features[31]. Convolutional, pooling, and fully connected layers, in addition to the I/O layer (FC), make up the majority of CNN's architecture. According on the size of the input image and the kind of application, the number of layers used varies.

The proposed CNN architecture for offline handwritten Gujarati numerals is shown in Figure 3. It has one output layer, two classification layers that are fully connected, and two feature extraction layers that are convolutional and max-pooling. The nodes of the network are linked together, and each layer of the network accepts as input the output of the layer before it. The output from one layer is used as the input for the following layer. Figure 3 also illustrates how CNN's architecture processed the handwritten Gujarati numeral images.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, No. 3, August : 2023

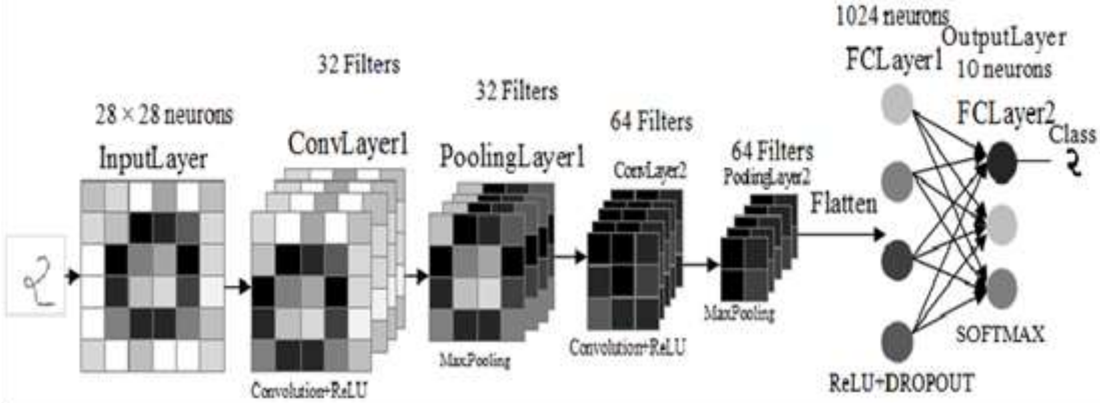

Fig. 3 CNN Architecture for handwritten numeral image processing.

3.5 Parameter Tuning For CNN

Network parameters are crucial in determining how sophisticated the CNN architecture is. In this case, CNN uses two fully connected layers for classification and a sequence of two convolutional and pooling layers for feature extraction. There are 10 output mappings in the first convolutional layer and 20 in the second convolutional layer. The input of the first convolutional layer is a 28 × 28 image. Using the dimensions of the input feature maps (N), the filter dimensions (F), and the convolution stride, the size of the output feature maps is calculated as (((N- F) /S) + 1) [10]. (S). In this instance, each feature map's bias is 1.

In the, first covonutional layer filter size 5 x 5 and stride 1 is applied on input image of 28 x 28. The output of the first covolutional layer is 24x 24 with 10 feature map. As a result, the first convolutional layer's learning parameters are (5 + 5 + 1) 10 = 260, and its connections are 24 * 24 * (5 + 5 + 1) 10 = 1,49,760. ReLU activation function is used to maintain non-linearity. First pooling layer output is 12 x 12 with 10 feature maps.

Similarly,fFor the second convolutional layer, the parameters used to train are ((5 5 + 1) 10) 20 = 5200. Once more, non-linearity is maintained by the ReLU activation function. The output size of second pooling layer is 4 x 4 with 20 feature map.

In the first fully-connected(FC) layer , parameters are 320 × 20 × (5 × 5 + 1) = 1,66,400, and in the second fully-connected(FC)layer parameters are 50 × (320 + 1) = 16,050 and the final layer are 10 × (50 + 1) = 510. Finally, each output node represents a specific Gujarati numeral.

IV. Experimental Result & Discussion

In this study, deep learning was utilized to extract features from images of handwritten Gujarati numerals. CNN is used to try to recognize image objects of handwritten Gujarati numerals.

4.1. Dataset Description

A data set of handwritten Gujarati numeral images that I developed myself is used in this work. The samples of numerical images were gathered from 220 people of various ages. On blank white A4-sized paper, each participant wrote numbers. A data set of 2200 numerical images in the Gujarati script was created as each author provided 10 numeral image samples. Samples were scanned at 300 dpi resolutions after the data collection process in order to obtain image samples. Next, using MATLAB software, bounding boxes were used to classify and confine these data.Each class's input image has been scaled down to a size of 28x28 pixels. Figure 4 displays a few representative images from these datasets. This dateset has no audible noise.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, No. 3, August : 2023

Fig. 4 Data set Sample images

4.2. Performance Evaluation

The proposed offline Gujarati handwritten numeral recognition system's results is based on manually created data set. Out of the 220 samples available for a certain numeral, 150 are randomly selected for training and 50 are used for testing. As a result, there were 1700 and 500 samples in the training and testing sets for numerical recognition, respectively. Training samples are evenly distributed among the underlying 10 classes for numerical recognition. The correctness of the test set is used to determine performance. The training session performance graph is shown in Figure 5.

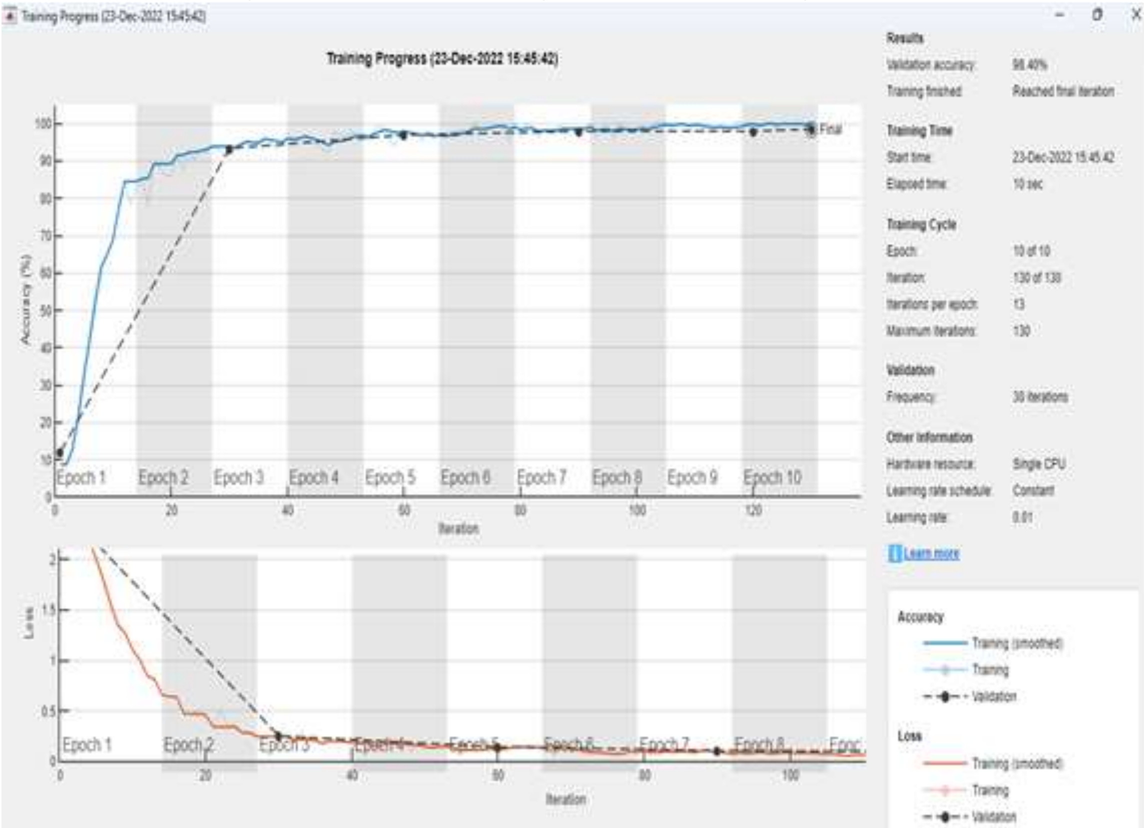

Fig. 5 Training Performance

Table 1 displays the confusion matrix that was discovered during the testing phase. This confusion matrix shows that the performance of numerical recognition for offline handwritten Gujarati script is fairly good. Out of the 10 classes, six of them had accuracy on the data set of 100%. In the suggested system, two classes have recognition accuracy of 98.0%, while the other two have a recognition accuracy of 94.0%. Overall, the Gujarati handwritten numeral recognition system achieves 98.4% accuracy.

Table 1. confusion matrix for offline handwritten Gujarati numerals

Numerals	૦	૧	૨	૩	૪	૫	૬	૭	૮	૯	Total
૦	50	0	0	0	0	0	0	0	0	0	50
૧	0	50	0	0	0	0	0	0	0	0	50
૨	0	0	47	0	1	0	0	2	0	0	50
૩	0	0	0	47	2	0	1	0	0	0	50
૪	0	0	1	0	50	0	0	0	0	0	50
૫	0	0	0	0	0	50	0	0	0	0	50
૬	0	0	0	0	0	0	50	0	0	0	50
૭	0	1	0	0	0	0	0	49	0	0	50
૮	0	0	0	0	0	0	0	0	50	0	50
૯	0	0	0	0	0	0	0	0	1	49	50
	50	50	48	47	53	50	51	51	51	49	98.4

4.3. State Of The Art Comparison

The recognition accuracy of the proposed work is compared with the conventional recognition accuracy for handwritten Gujarati numerals available in the literature and is shown in Table 2.

Table 2 Performance Comparsion

Ref. No.	Author	Method	Accuracy
[32]	Moro K. et. al.	Sparse Representation	80.33%
[33]	Desai A. A.	Neural Network	81.66%
[34]	Maloo M., Kale K. V.	Support Vector Machine	90.55%
[35]	Satange D.,et. al.	Neural Network	93.83%
[36]	Naik V.A., Desai A.A.	Support Vector Machine	95.00%
	Proposed	**CNN**	**98.40%**

V. Conclusion

Researchers from all around the world have contributed to and developed new methods and concepts in the subject of handwritten text recognition, which is a dynamic one. A deep learning based offline Gujarati handwritten numeral recognition system is proposed. This paper also explains the architecture and parameters required to train the network.The proposed system shows an excellent accuracy for Gujarati hand written numeral recognition. From the experimental results, The self-created data set shows that CNN does a good job at classifying handwritten Gujarati numerals. The average recognition accuracy for numerals in this study was 98.4%. The suggested solution focuses on the recognition of pre-segmented handwritten Gujarati numerals, which is not always viable in the actual world. The suggested work's scope can be expanded to include the evaluation of handwritten Gujarati conjunct characters and handwritten Gujarati alphabets. The suggested method can be used to recognize various handwritten independent scripts.

References:

[1] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-Based Learning Applied to Document Recognition,” proc. IEEE, vol. 86, no. 11, pp. 2278–2324, (1998).
[2] O. M. Parkhi, A. Vedaldi, and A. Zisserman, “Deep Face Recognition,” Procedings Br. Mach. Vis. Conf. 2015, no. Section 3, pp. 41.1-41.12, (2015).
[3] C. Couprie, C. Farabet, L. Najman, and Y. LeCun, “Indoor Semantic Segmentation using depth information,” pp. 1–8, arXiv preprint arXiv:1301.3572 (2013).
[4] K. He, X. Zhang, S. Ren, and J. Sun, “Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 9, pp. 1904–1916, 2015.Mahto et al. / Electronic Letters on Computer Vision and Image Analysis 20(2):69-82, (2021)
[5] M. Z. Alom, P. Sidike, T. M. Taha, and V. K. Asari, “Handwritten Bangla Digit Recognition Using Deep Learning,” ArXiv1705.02680v1 [Cs], (2017).
[6] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” Proc. IEEE Comput. Soc. Conf. Compt. Vis. Pattern Recog., pp. 580–587, (2014).
[7] O. Russakovsky et al., “ImageNet Large Scale Visual Recognition Challenge,” Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, (2015).
[8] L. V. Rasmussen, P. L. Peissig, C. A. McCarty, and J. Starren, “Development of an optical character recognition pipeline for handwritten form fields from an electronic health record,” J. Am. Med. Informatics Assoc., vol. 19, no. E1, (2012).
[9] D. Deodhare, N. N. R. Ranga, and S. R. Amit, “Preprocessing and Image Enhancement Algorithms for a Form-based Intelligent Character Recog. Sys.,” Int. J. Comp. Sci. Appl., vol. 2, no. 2, pp. 131–144, (2005).
[10] M. Z. Alom, P. Sidike, M. Hasan, T. M. Taha, and V. K. Asari, “Handwritten Bangla Character Recognition using Inception Convolutional Neural Network,” Computational Intell. Neurosci., vol. 2018, pp. 1–13, (2018).
[11] M. K. Mahto, K. Bhatia, and R. K. Sharma, “Combined Horizontal and Vertical Projection Feature Extraction Technique for Gurmukhi Handwritten Character Recognition,” in 2015 International Conference on Advances in Computer Engineering and Applications (ICACEA), pp. 59–65, (2015).
[12] R. Ghosh, C. Panda, and P. Kumar, “Handwritten Text Recognition in Bank Cheques,” 2018 Conf. Inf. Commun. Technol. CICT 2018, pp. 1–6, (2018).
[13] S.N.S. Rajasekaran, B.L. Deekshatulu, Recognition of printed Telugu characters, Computer Graphics Image Processing 335–360, (1977).
[14] R.K. Sinha, H.N. Mahabala, Machine recognition of Devnagari script, IEEE Transactions on Systems, Man, and Cybernetics, 435–441, (1979) .
[15] K. Ray, B. Chatterjee, Design of a nearest neighbor classififier system for Bengali character recognition, Journal of Institute of Electronics, Telecom Engineers 30, 226–229, (1984).
[16] B.B. Chaudhuri, U. Pal, A complete printed Bangla OCR system, Pattern Recognition 31, 531–549, (1998).
[17] S. Antani, L. Agnihotri, Gujarati character recognition, in: Fifth International Conference on Document Analysis and Recognition (ICDAR’99), pp. 418–421, (1999).

[18] B.B. Chaudhhuri, U. Pal, M. Mitra, Automatic recognition of printed Oriya script, Saadhanaa 27 (1), 23–34, (2002).
[19] C.V. Lakshmi, C. Patvardhan, Optical Character recognition of basic symbols in printed Telugu text, IE(I)Journal-CP 84, 66–71, (2003).
[20] Dholakia, A. Negi, S. Ram Mohan, Zone identifification in the printed Gujarati text, in: Proceedings of the 2005 Eight International Conference on Document Analysis and Recognition (ICDAR05), (2005).
[21] V.N. Manjunatharadhya, G. Hemanthakumar, S. Naushath, Multilingual OCR system for South Indian scripts and English documents: an approach based on Fourier transform and principal component analysis, Engineering Application of Artifificial Intelligence 21 (4), 658–668, (2008).
[22] N. Saxena, H. Parveen, Text Extraction Systems for Printed Images: A Review International Journal Of Advanced Studies Of Scientific Research (IJASSR) Special Issue based on proceedings of 4th International Conference on Cyber Security (ICCS), ISSN 2460 4010, (2018).
[23] P. Chinnuswamy, S.G. Krishnamoorty, Recognition of hand-printed Tamil characters, Pattern Recognition 12 (3), 141–152, (1980).
[24] A. Dutta, S. Chaudhary, Bengali Alpha-numeric character recognition using curvature features, Pattern Recognition 26 (12), 1757–1770, (1993).
[25] N.P. Banashree, D. Andhre, R. Vasanta, P.S. Satyanarayana, OCR for script identifification of Hindi (Devnagari) numerals using error diffusion Halftoning Algorithm with neural classififier, Proceedings of World Academy of Science Engineering and Technology 20, 46–50, (2007).
[26] S.V. Rajashekararadhya, P.V. Ranjan, Effificient zone based feature extraction algorithm for handwritten numeral recognition of popular south Indian scripts, Journal of Theoretical and Applied Information Technology 7 (1), 1171–1180, (2009).
[27] N. Shanthi, K. Duraiswamy, A novel SVM-based handwritten Tamil character recognition, Pattern Analysis and Application, Published Online http://www.springerlink.com/content/k708558648652q12/S, (2009).
[28] G.R. Hemanth, M. Jayasree, S. Keerthi Venii, P. Akshaya, and R. Saranya, CNN-RNN Based Handwritten Text Recognition, ICTACT Journal On Soft Computing, October, Volume: 12, Issue: 01, (2021).
[29] D.H. Hubel , T.N. Wiesel Receptive fields and functional architecture of monkey striate cortex. J Physiol 195:215–243 (1968).
[30] K. Fukushima, Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 36:193–202 (1980).
[31] M. Z. Alom, P. Sidike, M. Hasan, T. M. Taha, and V. K. Asari, “Handwritten Bangla Character Recognition using Inception Convolutional Neural Network,” Computational Intell. Neurosci., vol., pp. 1–13, (2018).
[32] K. Moro, M. Fakir and B. Bouikhalene “Handwritten Gujarati Digit Recognition using Sparse Representation Classifier”, Proccedings of International Arab Conference on Information Technology (ACIT'2016), (2016).
[33] A.A. Desai, “Gujarati handwritten numeral optical character reorganization through neural network”, ELSEVIER journal on Pattern Recognition, pp. 2582–2589, (2010).
[34] M. Maloo, K.V. Kale, “Support Vector Machine based Gujarati Numeral recognition”, International Journal on Computer Science and Engineering (IJCSE), Vol. 3 No. 7 July (2011).
[35] D. Satange, P. E. Ajmire, F.I. Khandwani, “offline handwritten gujrati numeral recognition using mlp classifier”, Volume 3, Issue 8 August (2018).
[36] V.A. Naik, A.A. Desai, “Online Handwritten Gujarati Numeral Recognition Using Support Vector Machine”, International Journal of Computer Sciences and Engineering Vol.6(9), Sept. (2018).