International Journal of Computer Trends and Technology
ISSN: 2231 – 2803 / https://doi.org/10.14445/22312803/IJCTT-V71I1P108

Volume 71 Issue 1, 50-55, January 2023

Original Article

Application of Deep Convolutional Neural Networks to Telugu Scripts for Optical Character Recognition

M. Gnaneswari [1], T. Chaitanya Kumar [2]

[1] *M. Tech (CNIS), MVGR College of Engineering (A), Andhra Pradesh, India.*
[2] *CSE, MVGR College of Engineering (A), Andhra Pradesh, India.*

Received: 10 December 2022 Revised: 13 January 2023 Accepted: 23 January 2023 Published: 31 January 2023

Abstract *- This research looks at the process of optical character recognition (OCR) for Telugu scripts. Telugu is an Indian Dravidian language. In English, optical character recognition is widely used, and there is a plethora of smartphone apps available. Telugu has a significantly higher level of complexity due to the number of output classes that can be formed and the inter-class diversity. In addition, there are no good Telugu OCR systems. We employed the Deep Convolutional Neural Network (DCNN) model for Telugu character recognition because of its success in other domains, such as segmentation, object identification, and character recognition. Multiple machine learning algorithms such as AdaBoost, Support Vector Machine (SVM), XGBoost, and Decision Tree (DT) are considered for performance evaluation of the proposed DCNN. The proposed DCNN for recognizing the Telugu scripts has yielded promising results, demonstrating its usefulness when compared to other traditional techniques when experimented with using the IEEE Telugu Character Dataset.*

Keywords *- Text Recognition, Telugu Script, Optical Character Recognition (OCR), Deep Learning, Convolutional Neural Network (CNN), Machine Learning.*

1. Introduction

Machine vision applications benefited greatly from the field of pattern recognition. The field of pattern recognition includes handwriting recognition [12]. Handwriting recognition is a technology that allows a computer to recognize characters and other symbols written in natural handwriting. All of this should be done to maintain the documents in digital form so they can be referenced in the future. In order to engage with a computer system or exchange data with it, users must first enter the necessary data into the system.

Optical character recognition (OCR), which is at the heart of the field of document analysis systems, is the process of transforming written symbols on paper into a machine-processable format [20] [42]. As a result, it is crucial to transition from a paper-based to a paperless electronic information system. For Indian documents, OCR technology is still in its infancy, and most Indian OCR systems can only read documents written in a single script. Every state government must produce an official document containing a national script (Devanagari), an official script (Roman), and a state script, according to the Indian constitution's tri-scripts principle (or regional script). For example, Devanagari, Roman, and Gurmukhi scripts are all used in Punjabi government documents. For OCR researchers, processing such complicated multi-script manuscripts is a difficult challenge.

Recognizing text in natural situations is a significant ask in the field of machine learning and computer vision, which aims to build a computer application that can read scene text automatically from photographs [13]. This technology is presently mostly utilized to improve visual interpretation, which is useful in various real-world applications like automatic traffic signs, car license detection, and autonomous robot navigation.

Deep Learning is a Machine Learning technology that allows us to train machines to predict an output based on a collection of inputs rather than a set of instructions utilizing huge amounts of data. Deep Learning is based on neural networks, which are layers of nodes that operate similarly to neurons in the human brain. Individual layer nodes are linked to those of neighboring layers. Because of the number of levels, the network is thought to be more complicated. In an artificial neural network, signals pass between nodes and are given weights. A node with a higher weight has a greater impact on the nodes beneath it. The weighted inputs are combined into a result in the final layer. Deep learning systems require powerful hardware since they manage large amounts of data and execute complicated mathematical computations. Virtual assistants (Siri and Alexa), language translation systems [14] [15] [16], chatbots used in banking and health applications, and facial recognition methods to distinguish a person in photos are examples of real-world applications of deep Learning.

The following individuals are contributions to this paper:

- An end-to-end system for recognizing words in Indian scripts using image-based data-type recognition.
- Design and implementation of convolutional neural network model for efficient recognition of Telugu characters with an Improved performance.

The following is the order of the remainder of the paper: Section 2 delves into the literature review, Section 3 intothe methodology, Section 4 into the study's experimental setup, Section 5 into the performance analysis and experimentation outcomes, and Section 6 into the conclusion and future work.

2. Literature Review

In the last few decades, the field of optical character recognition (OCR) has seen an intense amount of research for Japanese, Chinese, and Roman scripts. In this aspect, Singh, P.K. et al. [1] experimented with a tree-based approach to distinguish between Matra and non-Matra scripts. The Distance Hough Transform (DHT) is initially used to identify the Matra script from the non-Matra script. The modified log-Gabor filter technique is employed to classify individual scripts into Matra and non-Matra categories. The proposed tree-based approach obtained an accuracy of 94.23%. Kaur, H. & Kumar, M. [2] addressed the necessity of developing word recognition techniques using Indic and non-Indic scripts. Different databases are evaluated for their recognition of the words. Multiple techniques and tools used for the experimental evaluation of word recognition are presented. It is observed that word recognition in printed documents of Indic scripts has a potential area of research.

Ritesh Sarkhel et al. [3] proposed a multi-column, multi-scale convolutional neural network (MMCNN). A prediction model based on a deep quad-tree staggered system is examined to recognise the characters at a faster rate. In this paper, 9 different datasets that are publicly available are tested in this paper. The experimental analysis indicates that the proposed model produced better results when compared to other competing algorithms. Samadiani, N., & Hassanpour, H. [4] elucidated a method for recognizing English characters in multiple fonts using a similarity measure neural network. Instead of a distance measure, a similarity measure metric is used in the proposed neural networks. The experimental evaluation indicates that the proposed model outperforms other compared algorithms.

Shanthi, N., & Duraiswamy, K. [5] proposed a model based on a support vector machine (SVM) to recognise Tamil characters. They built their dataset by collecting samples from different writers. The results showed that the proposed model achieved an accuracy of 82.04% in recognizing Tamil characters. Kunte, R.S., & Samuel, R.D. [6] demonstrated a neural network model for classifying Kannada characters. The deep learning model is trained using the features extracted by Zernike moments and Hu's invariant moment techniques. The results indicate that the experimented model obtained results when compared to traditional algorithms.

Bag, S., & Harit, G. [7] reviewed optical character recognition (OCR) work on the scripts of popular Indian languages such as Devanagari and Bangla. They have discussed the research work published by different authors and their methodologies. It also included the future direction of research for Indian scripts in OCR. Das Bebartta, H., & Mohanty, S. [8] addressed the problem of identifying multiple scripts in a single document using the X-Y-Cut algorithm for Odia and Roman scripts. Initially, they are segmented at line level and then are further identified for characters individually. The proposed methodology resulted in a better recognition model when compared to state-of-the-art techniques.

According to the literature review, several research works have addressed the problem of recognizing characters in different Indic scripts using state-of-art and other machine learning techniques. This paper aims to design and develop a deep convolutional neural network (DCNN) incorporated with transfer learning for the efficient recognition of Telugu characters.

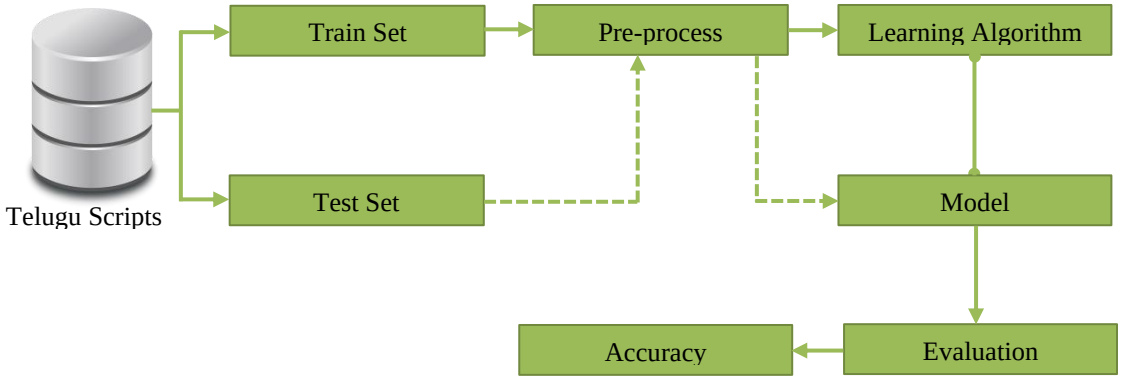

Fig. 1 Architecture of the proposed Optical Character Recognition system

3. Methodology

Character recognition is a significant issue in OCR [39]. The quality of any OCR system is heavily influenced by character recognition algorithms. Despite the availability of several character segmentation techniques that have been developed to date, present methodologies are unable to effectively distinguish characters belonging to some complex writing styles [17] [18]. This research investigates a new model called DCNN for effectively recognising Telugu characters in this regard [10].

The architecture of the proposed optical character recognition system is highlighted in Fig. 1. To maintain uniformity, all the images are preprocessed. The preprocessing steps, such as transforming the image shape into 224 x 224 and then dividing the dataset into training and testing in the ratio of 7:3, are considered the preprocessing steps. After that, the preprocessed data is fed into the DCNN learning algorithm. CNN had to extract feature maps from photos to overcome word recognition problems. The SoftMax component, which predicts label distributions, is used to model these feature vectorsequences.

In Fig. 2., the suggested Deep Convolutional Neural Network classifier is emphasized. The improved classification model was obtained by experimenting with and modifying the hyperparameters of a neural network model with several layers. The resulting model is made up of an input layer with nodes, six fully connected hidden layers with each, and an output layer with an activation function to recognize Telugu characters. The followingformula Eq (1) is used to determine the number of nodes in the hidden layer:

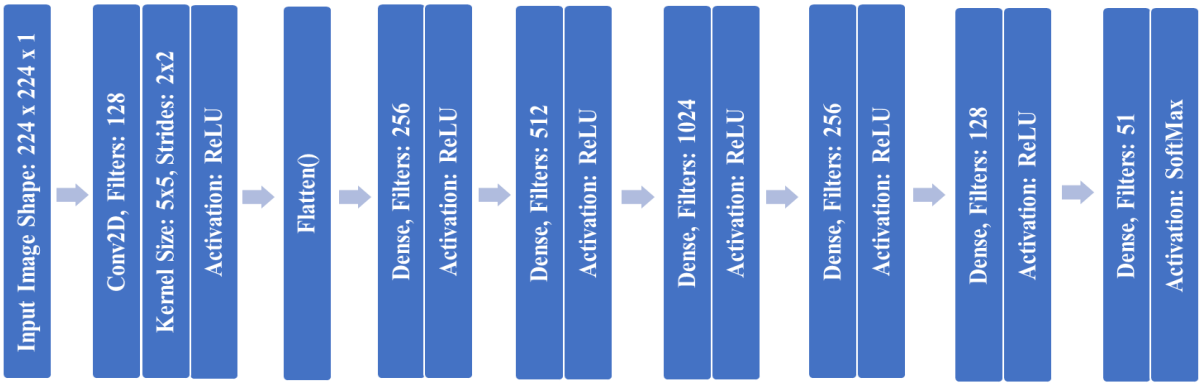

Fig. 2 Architecture of the proposed DCNN

$$M_{nodes} = \frac{N_{features}*2}{\alpha} + 2 \quad (1)$$

$N_{features}$ denotes the number of input features, and α represents an arbitrary value between [2,10]. Except for the last layer, all of the layers use the $ReLU$ activation function, $SoftMax$ activation function is used in the last layer, which is written as in Eq (2) & Eq (3).

$$ReLU: f\big(h_\theta(x)\big) = Max\big(0, h_\theta(x)\big) \quad (2)$$

$$SoftMax: f(y_i) = \frac{e^{y_i}}{\sum_j e^{y_j}} \quad (3)$$

Where $h_\theta(x)$ is obtained using the following Eq. (4):

$$\sum_{i=1}^{n} w_i x_i + bias = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + bias \quad (4)$$

Where $x_i = (x_1, x_2, \ldots, x_n)$ represents the inputs to the unitsin the layer, $w_i = (w_1, w_2, \ldots, w_n)$ indicates the weights of neurons, a constant term $bias$ is added to every layer, to assure the parameters are not passing through the origin.

4. Experimental Setup

The (64-bit) Windows 10 Operating System is used to execute the proposed system, along with an Intel Core i5 processor running at 2.20 GHz, a 1 TB Hard Drive, and 8 GB of RAM. Python is used for the programming interface on the Anaconda platform, which is set up to use deep-learning packages.

For reliable recognition of characters, the Telugu Handwritten Character Dataset was downloaded from the IEEE Data port [11]. All Telugu characters with vowels, consonants, and combining characters such as Othulu (Consonant-Consonant) and Guninthamulu (Consonant-Vowels) are included in the dataset. In comparison to English literature, Indian literature contains significantly fewer characters; for example, English literature contains only 26 while Telugu literature has 1,924 (Achulu(vowels)–16, Hallulu (consonants)–36, Othulu–36, and Guninthamulu–34 * 16 = 544). Sample images from the datasetare shown in Fig. 3. The description of the dataset is shown in Table 1.

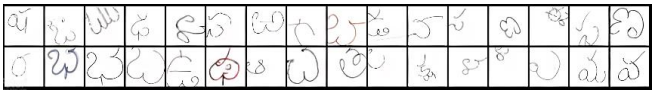

Fig. 3 Sample Images from the dataset

Table. 1 Dataset description

Description	Achulu	Guninthamuluu	Hallulu	Othulu
Categories	16	34	36	36
Size	37.2 MB	957.0 MB	76.5MB	72.0MB

Table 2. Accuracy comparison of DCNN withAdaBoost, SVM, XGBoost & DT

Exp No.	Model	Precision	Recall	F-Score	Accuracy
1	AdaBoost	0.1845	0.2136	0.0912	73.65%
2	SVM	0.6812	0.8147	0.5241	76.89%
3	XGBoost	0.7925	0.8467	0.4256	82.19%
4	DT	0.7561	0.8756	0.6927	91.27%
5	**DCNN**	**0.8869**	**0.9187**	**0.7627**	**94.93%**

5. Performance Analysis & Experimentation Results

Several evaluation metrics, including accuracy, precision, recall, and f-score, were considered for assessing the performance of the proposed system.

Table 2 summarizes the findings of these experiments. Table 3 shows a relevant work comparison based on accuracy. Fig. 4 & Fig. 5 depicts the model's performance in terms of accuracy and loss in relation to epochs. It has been noticed that as the number of epochs increases, so does the accuracy.

Table. 3 Comparison of similar works

Paper, Year	Classifier	Accuracy
[21], 2016	MLP	80.80%
[22], 2016	RNN	86.60%
[23], 2015	Multi-class SVM	88.75%
[24], 2016	Hidden Markov Model	93.82%
This Paper	**DCNN**	**94.93%**

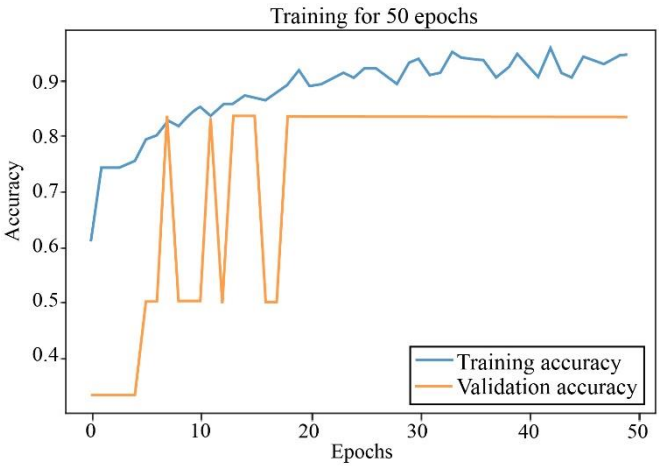

Fig. 4 DCNN performance in terms of accuracy

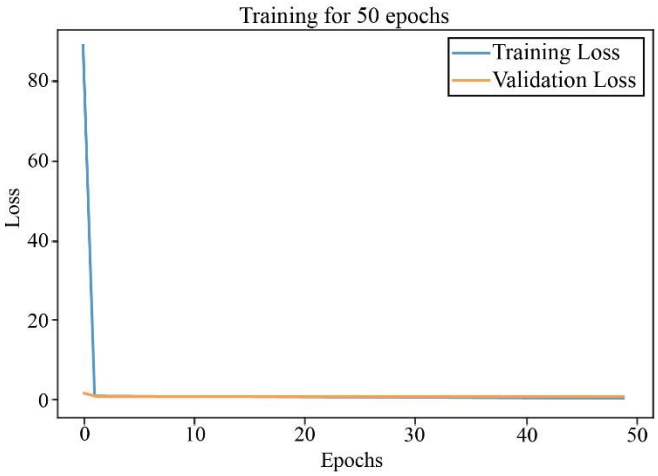

Fig. 5 DCNN performance in terms of loss

6. Conclusion & Future Work

By experimenting with several neural network designs, an effective model for Telugu character recognition is provided in this research. To extract the feature maps in the photos, the full dataset is first preprocessed, and then the preprocessed data is sent to the Deep Convolutional Neural Network model. The extracted feature maps are then employed in the model's final layer to recognize the characters. The experimentation is based on the IEEE Telugu character dataset. The suggested DCNN outperformed all other comparison algorithms with an accuracy of 94.93%. Future work will include designing and implementing various hybrid architectures using deep learning approaches and surveying large datasets for enhanced performance in recognizing words in Indian Scripts.

References

[1] Munish Kumar et al., 'Character and Numeral Recognition for Non-Indic and Indic Scripts: A Survey," *Artificial Intelligence Review*, vol. 52, pp. 2235-2261, 2019. *Crossref*, https://doi.org/10.1007/s10462-017-9607-x

[2] Konkimalla Chandra Prakash et al., "Optical Character Recognition (OCR) for Telugu: Database, Algorithm and Application," *25th IEEE International Conference on Image Processing (ICIP)*, pp. 3963-3967, 2018. *Crossref*, https://doi.org/10.1109/ICIP.2018.8451438

[3] Pradeep Bheemavarapu et al., "An Efficient Method for Coronavirus Detection Through X-rays using deep Neural Network", *Journal of Current Medical Imaging*, vol.18, no. 6, pp. 587- 592, 2022. *Crossref*, https://doi.org/10.2174/1573405617999210112193220

[4] Soumyadeep Kundu et al., "Understanding NFC-Net: a Deep Learning Approach to Word-Level Handwritten Indic Script Recognition," *Neural Computing and Applications*, vol. 32, pp. 7879-7895, 2020. *Crossref*, https://doi.org/10.1007/s00521-019-04235-4

[5] Magesh Kasthuri, and Venkatasubramanian Sivaprasatham, "Cognitive Reading and Character Recognition in Image Processing Techniques," SN Computer Science, vol. 1, no. 131, 2020. *Crossref*, https://doi.org/10.1007/s42979-020-00142-x

[6] S Satyanarayana and P. Srinivasa Rao, "Privacy-Preserving Data Publishing Based On Sensitivity in Context of Big Data Using Hive", *Journal of Bigdata(Springer*), vol. 5, no. 20, 2018. *Crossref*, https://doi.org/10.1186/s40537-018-0130-y

[7] Ismet Zeki Yalniz, and R. Manmatha "Dependence Models for Searching Text in Document Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 1, pp. 49-63, 2019, *Crossref*, https://doi.org/10.1109/TPAMI.2017.2780108

[8] Raymond Ptucha et al., "Intelligent Character Recognition Using Fully Convolutional Neural Networks," *Pattern Recogniti*on, vol. 88, 604-613, 2019. *Crossref*, https://doi.org/10.1016/j.patcog.2018.12.017

[9] Piyush Kiran Redgaonkar et al., "Image Processing Based Pincode Recognizing and Section Wise Courier Sorting System," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 3, no. 3, pp. 16-18, 2016. *Crossref*, https://doi.org/10.14445/23488379/IJEEE-V3I3P103

[10] K. Pramod Sankar, R. Manmatha, C.V. Jawahar, "Large Scale Document Image Retrieval By Automatic Word Annotation," *International Journal on Document Analysis and Recognition*, vol. 17, no. 1, 2014. *Crossref*, https://doi.org/10.1007/s10032-013-0207-2

[11] Pawan Kumar Singh et al., "A Comprehensive Handwritten Indic Script Recognition System: A Tree-Based Approach," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-18, 2018. *Crossref*, https://doi.org/10.1007/s12652-018-1052-4

[12] T.V. Madhusudhana Rao et al., "Matrix Factorization Based Recommendation System using Hybrid Optimization Technique," *EAI Endorsed Transactions on Energy Web*, vol. 5, no. 35, 2021. *Crossref*, https://doi.org/10.4108/eai.19-2-2021.168725

[13] Harmandeep Kaur, and Munish Kumar, "A Comprehensive Survey on Word Recognition For Non-Indic And Indic Scripts," *Pattern Anal Applic*, vol. 21, pp. 897–929, 2018. *Crossref*, https://doi.org/10.1007/s10044-018-0731-2

[14] Ritesh Sarkhel et al., "A Multi-scale Deep Quad Tree Based Feature Extraction Method for the Recognition of Isolated Handwritten Characters of popular Indic Scripts," *Pattern Recognition*, vol. 71, pp. 78-93, 2017. *Crossref*, https://doi.org/10.1016/j.patcog.2017.05.022

[15] T.V. Madhusudhana Rao, P Srinivasa Rao, and P.S. Latha Kalyampudi, "Iridology based Vital Organs Malfunctioning Identification using Machine learning Techniques," *International Journal of Advanced Science and Technology*, vol. 29, no. 5, pp. 5544 – 5554, 2020.

[16] Najmeh Samadiani, and Hamid Hassanpour, "A Neural Network-Based Approach for Recognizing Multi-Font Printed English Characters," *Journal of Electrical Systems and Information Technology*, vol. 2, no. 2, pp. 207-218, 2015. *Crossref*, https://doi.org/10.1016/j.jesit.2015.06.003

[17] Shanthi, N., and Duraiswamy, K., "A novel SVM-Based Handwritten Tamil Character Recognition System," *Pattern Analysis and Applications*, vol. 13, pp. 173-180, 2010. *Crossref*, https://doi.org/10.1007/s10044-009-0147-0

[18] S.Vidya sagar Appaji, P. V. Lakshmi, and P. Srinivasa Rao, "Maximizing Joint Probability in Visual Question Answering Models", *International Journal of Advanced Science and Technology*, vol. 29, no. 3, pp. 3914 – 3923, 2020.

[19] Juhi Ranglani, and Vijay Lachwani, "Automatic Number Plate Recognition (ANPR)," *SSRG International Journal of Computer Science and Engineering*, vol. 3, no. 8, pp. 64-68, 2016. *Crossref*, https://doi.org/10.14445/23488387/IJCSE-V3I8P114

[20] Soumen Bag, and Gaurav Harit, A survey on Optical Character Recognition for Bangla and Devanagari Scripts," *Sadhana*, vol. 38, pp. 133-168, 2013. *Crossref*, https://doi.org/10.1007/s12046-013-0121-9

[21] Vidya sagar Appaji setti, and P Srinivasa Rao, "A Novel Scheme For Red Eye Removal With Image Matching", *Journal of Advanced Research in Dynamical & Control Systems*, vol. 10, no. 13, 2018.

[22] Himadri Nandini Das Bebartta, and Sanghamitra Mohanty, "Algorithm for Segmenting Script-Dependant Portion in a Bilingual Optical Character Recognition System," *Pattern Recognition and Image Analysis*, vol. 27, pp. 560-568, 2017. *Crossref*, https://doi.org/10.1134/S1054661817030142

[23] Arun K. Pujari et al., "An Intelligent Character Recognizer for Telugu Scripts Using Multiresolution Analysis and Associative Memory," *Image and Vision Computing*, vol. 22, no. 14, pp. 1221-1227, 2004. *Crossref*, https://doi.org/10.1016/j.imavis.2004.03.027

[24] P Srinivasa Rao, and P.E.S.N. Krishna Prasad, "A Secure and Efficient Temporal Features Based Framework for Cloud Using MapReduce," *17th International Conference on Intelligent Systems Design and Applications*, vol. 736, pp. 114-123, 2017. *Crossref*, https://doi.org/10.1007/978-3-319-76348-4_12

[25] Mokhtar Hussein, and B. Manjula, "A Hybrid Approach for SVD and Neural Networks Based Robust Image Watermarking," *International Journal of Computer Trends and Technology*, vol. 61, no. 2, pp. 117-121, 2018. *Crossref*, https://doi.org/10.14445/22312803/IJCTT-V61P120

[26] N. Prameela, P. Anjusha, and R. Karthik, "Offline Telugu Handwritten Characters Recognition Using Optical Character Recognition," *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)*, pp. 223-226, 2017. *Crossref*, https://doi.org/10.1109/ICECA.2017.8212801

[27] T.V. Madhusudhana Rao, and Y. Srinivas, "A Secure Framework For Cloud Using Map Reduce," *Journal Of Advanced Research In Dynamical And Control Systems*, vol. 9, no. 14, pp:1850-1861, 2017.

[28] Negi, C. Bhagvati and B. Krishna, "An OCR system for Telugu," *Sixth International Conference on Document Analysis and Recognition*, pp. 1110-1114, 2001. *Crossref*, https://doi.org/10.1109/ICDAR.2001.953958

[29] [Online]. Available: https://ieee-dataport.org/open-access/telugu-handwritten-character-dataset#files.

[30] P Srinivasa Rao, Sushma Rani N, and P Parimala, "An Efficient Statistical Computation Technique for Health Care Big Data using R," *IOP Conference Series: Materials Science and Engineering*, vol. 225, 2017. *Crossref*, https://doi.org/10.1088/1757-899X/225/1/012159

[31] Satish Kumar,. "A study for Handwritten Devanagari Word Recognition," International *Conference on Communication and Signal Processing*, pp. 1009-1014, 2016. *Crossref*, https://doi.org/10.1109/ICCSP.2016.7754301

[32] Chandranath Adak et al. "Offline Cursive Bengali Word Recognition Using CNNs with a Recurrent Model," *15th International Conference on Frontiers in Handwriting Recognition*, pp. 429-434, 2016. *Crossref*, https://doi.org/10.1109/ICFHR.2016.0086

[33] Ms. Dipalee A. Kolte, Maruti B. Limkar, and Sanjay M. Hundiwale, "Character recognition from deblurred motion distorted Vehicle image using Neural Network," *SSRG International Journal of Electronics and Communication Engineering*, vol. 1, no. 4, pp. 1-9, 2014. *Crossref*, https://doi.org/10.14445/23488549/IJECE-V1I4P101

[34] Bikash Shaw, Ujjwal Bhattacharya, and Swapan Kumar Parui, "Offline handwritten Devanagari word recognition: Information fusion at feature and classifier levels," *3rd IAPR Asian Conference on Pattern Recognition*, pp. 720-724, 2015. *Crossref*, https://doi.org/10.1109/ACPR.2015.7486597

[35] Nagesh Vadaparhi et al., "A Novel Clustering Approach using Hadoop Distributed Environment," *Applied Science and Technology*, vol. 9, pp. 113-119, 2014. *Crossref*, https://doi.org/10.1007/978-981-287-338-5_9

[36] Rajib Ghosh, and Partha Pratim Roy. "Comparison of Zone-Features for Online Bengali and Devanagari Word Recognition Using HMM," *15th International Conference on Frontiers in Handwriting Recognition*, pp. 435-440, 2016. *Crossref*, https://doi.org/10.1109/ICFHR.2016.0087

[37] Balajee Maram, Guru Kesava Dasu Gopisetty, and P Srinivasa Rao "A Framework for Data Security using Cryptography and Image Steganography," *International Journal of Innovative Technology and Exploring Engineering*, pp. 2278-3075, vol. 8, no. 11, 2019.

[38] Zeng Runhua, Zhang Shuqun "Improving Speech Emotion Recognition Method of Convolutional Neural Network," *International Journal of Recent Engineering Sciece*, vol. 5, no. 3, pp. 1-7, 2018. *Crossref*, https://doi.org/10.14445/23497157/IJRES-V5I3P101

[39] P.Mahesh Kumar, and P. Srinivasa Rao, "Frequent Pattern Retrieval on Data Streams by using Sliding Window," *EAI Endorsed Transactions on Energy web*, vol. 5, no. 35, 2021. *Crossref*, https://doi.org/10.4108/eai.13-1-2021.168091

[40] Mohammed Aarif, & Sivakumar Poruran, "OCR-Nets: Variants of Pre-trained CNN for Urdu Handwritten Character Recognition via Transfer Learning," *Procedia Computer Science*, vol. 171, pp. 2294-2301, 2020. *Crossref*, https://doi.org/10.1016/j.procs.2020.04.248

[41] P. Srinivasa Rao, Krishna Prasad, M.H.M., and K.Thammi Reddy, "A Efficient Data Integration Framework in Hadoop Using MapReduce," Computational Intelligence Techniques for Comparative Genomics, pp. 129-137, 2014.

[42] R Sanjeev Kunte, and R. D. Sudhaker Samuel, "A Simple and Efficient Optical Character Recognition System for Basic Symbols in Printed Kannada Text," *Sadhana-academy Proceedings in Engineering Sciences*, vol. 32, no. 5, pp. 521-533, 2008. *Crossref*, https://doi.org/10.1007/s12046-007-0039-1