www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
Vol. 02, Issue 07, July 2022, pp : 05-09

e-ISSN : 2583-1062
Impact Factor : 2.205

KANNADA HANDWRITTEN CHARACTER RECOGNITION USING HYBRID RNN

Siddanna S R[*1], Anant Kumar[*2], Avdhesh Dubey[*3], Bal Govind[*4], Diksha Dixit[*5]

[*1]Assistant Professor, Information Science and Engineering, SJB Institute Of TechnologyBangalore, India

[*2,3,4,5]Student, Information Science and Engineering,SJB Institute Of Technology Bangalore, India

ABSTRACT

The objective of this project is to classify characters written in Kannada, a South Indian language. The images of characters are processed using various image processing techniques such as normalization, denoising, resizing etc. using numpy and OpenCV. The dataset used is a combination of the Chars74k dataset. Classifiers have been implemented based on Convolutional Neural Networks, Recurrent Neural Networks and a hybrid approach using Convolutional Neural Networks and Recurrent Neural Networks using OpenCV and Keras and a comparison of their accuracies has been provided. The hybrid classifier has been able to achieve an accuracy of 86.82

1. INTRODUCTION

Kannada handwriting recognition is a function that rec- ognizes the character obtained from the image source and interprets the character. The biggest challenge is recognizing the character because each person has different handwriting. Sometimes people write characters differently than before. Handwritten text can be overlaid and the accupoints make a difference. There are many ways you can write a single character . The focus of this work is to recognize the user- uploaded handwritten Kannada text image and output it as digital text. It consists of many stages, the first stage is the pre-processing of the data, removing unwanted spaces and noise from the image, padding and resizing, and converting the image to grayscale values. .. The datasets is stored as training and test data and is fed as input to the neural network . Text recognition plays a major role in many areas. However, it is difficult to do such work mechanically. For recognition, you need to train the system to to check the text. Character recognition involves multiple steps such as capture, extraction of function , classification, and recognition. Handwriting recognition is a feature of the machine that receives and interprets handwriting input from external sources such as images. The goal of this project is to design a system that can efficiently recognize the actual character of the format using a neural network. The neural computer seems to be trained and unprogrammed The given data is compared to the trained system to provide the user with acceptable output text . The handwriting recognition system takes over the format, and the performs the correct segmentation to the letters and finds the most important plausible words. From ancient times, it was customary to record all important information. Manuscripts range from ancient palm leaf manuscripts to modern paper document . These documents are saved and retrieved whenever needed. Character recognition in this document remains a daunting task. character recognition is one of the most impor- tant applications in the field of image processing, and research is still underway to find the best algorithm or network to achieve maximum accuracy. Many techniques were used. This proposed task performs one of the's basic steps in handwriting recognition for paper documents. The first step is to create a Kannada handwriting database , assigning specific parts of image for testing and the rest for training. Kannada is the official language used by more than 1.2 billion people worldwide. Character recognition is considered one of the most important technologies in the world today. It is used in various fields of , such as artificial intelligence, computer vision, and patterns. Handwriting recognition is divided into two parts.

2. LITERATURE SURVEY

Creating an OCR for Indian languages is a challenge problem. K.S. Sesh Kumar et.al clarified with difficulty in separating texts in Telugu, Tamil, Bengali and Malayalam. Complexity of texts lies in the local distribution of connected parts. Z. Razak et.al proposed methods based on the topic line detection, detection of the base line and tracking contour of the line separation of handwritten characters in Hindi.

M. Arivazhagan et.al , analyzed grid-based effects method also exceeds the standard PCA method. C.V. Aravinda et.al proposed a template based model in Correlation Analysis that showed promising results.

3. PROBLEM DESCRIPTION

Character recognition is a field that has done a lot of progress in recent years and shows a beautiful field improvement-development. Unlike love languages like Span- ish, Portuguese, French, Italian, English etc. have done a lot of manuscript research alphabetic recognition, Indic languages are always limited an unnamed field. The goal is to be able to expand accuracy of handwriting recognition recognition Kannada characters have previous

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
Vol. 02, Issue 07, July 2022, pp : 05-09

e-ISSN : 2583-1062
Impact Factor : 2.205

attempts like once . These characters are individually separated and distinct compiled by in which the authors enter the full details as how to previously collect and process a database. This problem is a problem of offline reading and refers to situations where the system does not exist working and capturing new information in real time. Instead, it has a fixed set of input data.

4. PROPOSED SOLUTION

This section explains the methods that were followed to solve the problem of recognizing characters from Indic scripts, particularly Kannada. It explains how the dataset was acquired and preprocessed as well as the machine learning models that were implemented. Use of Recurrent Neural Network with convolutional neural network and Connectionist Temporal Classification layers. CNN layers will be trained to take out relevant features from the image. RNN propagates relevant information and output is mapped to a matrix. CTC is given RNN output matrix and it decodes it into ultimate text

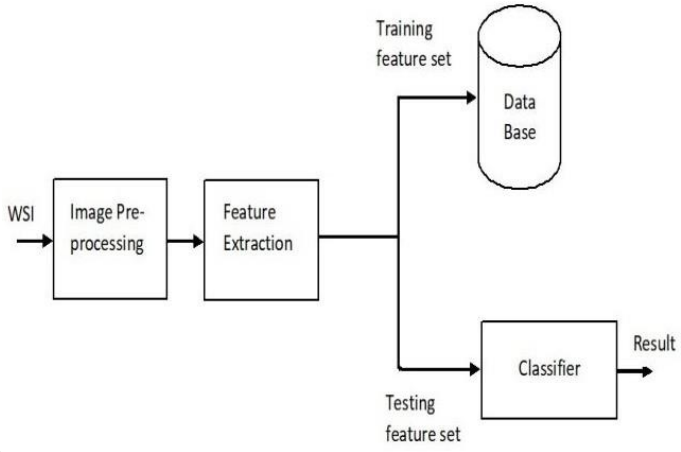

Fig. 1. Proposed system design

5. CONVOLUTIONAL NEURAL NETWORK

Convolutional Networks (LeCun, 1989), also known as Convolutional Neural Networks, or CNN, is special a type of neural network to process known data grid topology. Keras is an in-depth library provides an easy way to create CNNs. It uses Tensorflow as a backend. To use CNN, a method like this used by Prashanth Vijayaraghavan et.al is used. Built using two layers of Convolution and Max Pooling layers layers. After the first stack, Local Response Orientation a layer is added to apply lateral block. Exit layer it is also added by 0.5 times to avoid overcrowding. A A flat layer is then inserted to convert a feature set to 1- Dimensional Vector re-installed on our fully connected internet Crowded layer. Softmax function is used as an output layer find the output as opportunities. Adam is chosen as an optimizer with a dependency level of 1e-2 that was determined to be used method proposed by Leslie N. Smith.The first conversion layer uses 32 window-sized filters. 5x5 and the second convection layer use 64 filters for the same window size. Both layers of convolution use sigmoid as function to activate. As can be seen in Fig. 1, accuracy shoots upwards the first few seasons and then take about 100 more times stability at the highest level of the world.

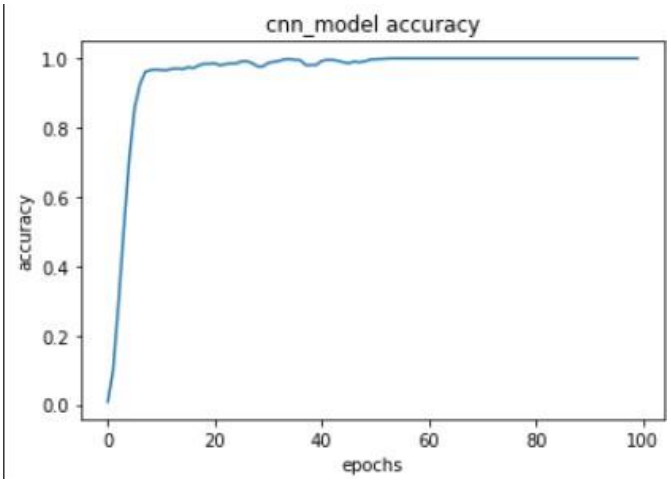

Fig. 2. Variation of accuracy of CNN with respect to the epoch

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
Vol. 02, Issue 07, July 2022, pp : 05-09

e-ISSN : 2583-1062
Impact Factor : 2.205

6. DATA SET

One of the databases used is the Chars74K dataset. It Learning Compared to the Implementation of Supervised and Supervised Models of Recognizing Kannada Manuscripts 775 contains a collection of photos for 657+ classes. Each image contains a single letter or combination characters. Each class has 25 handwritten letters. As an in-depth learning approach has also been used in this paper, the data needed for expansion. This is done through a handwritten note given to 30 volun- teers. Continuation expand the database, using augmentation techniques. This custom database is called PSCube in all paper.

7. METHODOLOGY

Acquired dataset from various sources. Converting dataset into images and sampling them to be used by neural network. The image samples acquired are to be converted into gray scale and proper noise reduction is to be done. RNN-Recurrent Neural Network layers are to be used for feature extraction. CTC is planned to be used for classification. Output-Output will be the correct prediction of the character from the image provides by the user.In the front end, the user has to upload the handwritten text image to the Web application from his browser. The application sends the uploaded image to the trained model where the analysis of the image takes place with various algorithms. The model will have been trained effectively by feeding the data set and creating the Neural Networks. The final feedback of the image is obtained after completing the analysis. The output displays the probability the digital form of the characters in the image.

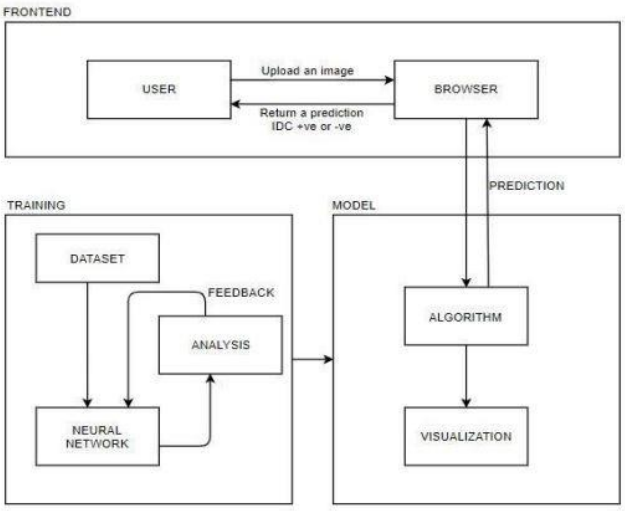

Fig. 3. Architecture Diagram for the Proposed System

8. PROCESS FLOW DIAGRAM

The data will first be extracted from the .png file and the dataset will be processed. Each pixel of the image data is normalized and extracted. This helps in fitting the model in a better way. The next step is training the model, featureselection and dimension reduction. Learning algorithms willbe used to classify data.

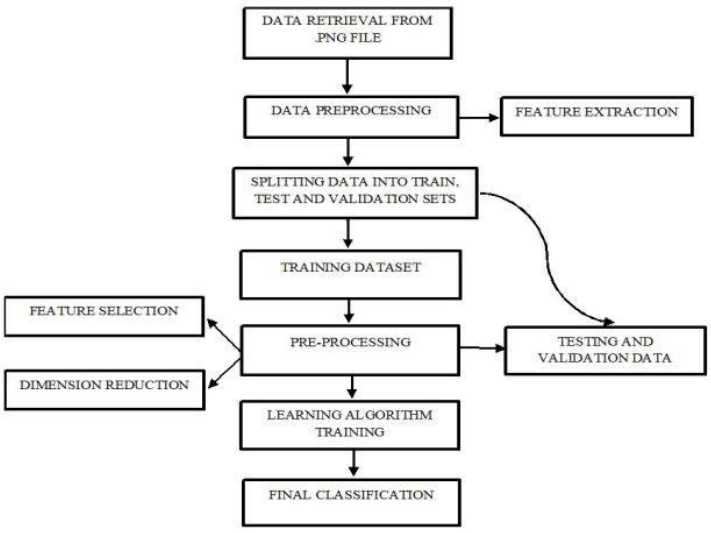

Fig. 4. Process flow diagram

9. CLASS DIAGRAM

First the patches from Whole Slide Images (WSI) is col- lected from the Char74K dataset repository.

1) Then pre-processing is done of the collected data
2) Then the processed data will be sent to the trainingmodel.
3) Now the model is ready with the trained data.
4) Then in testing phase, using the most accurate algorithm,the result of the classification will be provided.

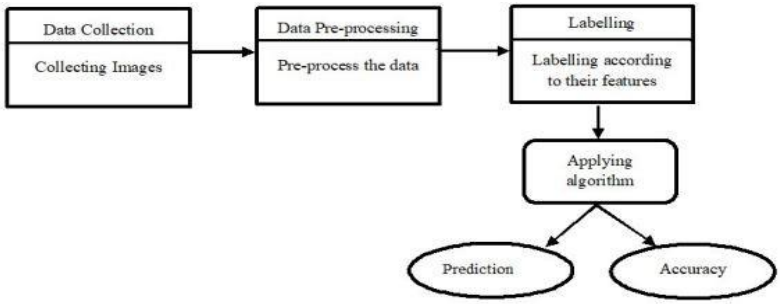

Fig. 5. Class diagram

10. RESULT

Comparing our accuracy's to the ones show in papers and we can see that our CNN architecture outperforms 4 Fig. 1. Time taken by different Classifiers. all other models and shows an improvement in accuracy by 5 even without manual feature extraction as shown in Figure 2. But our machine learning models suffer the same low accuracy as that in Discussed in Section V are the possible reasons for the differences in accuracies. These results were calculated on a cloud GPU. The CNN predicts the handwritten Kannada characters with an accuracy of 100 percent whereas the other models show low accuracy. LSTM shows an accuracy of 97.76 percent whereas Hybrid model shows an accuracy of 86.82 percent. We see a higher accuracy due to the deep learning nature of the CNN algorithm. The time taken by the CNN to train is significantly higher than that of either the machine learning models at 139980 seconds. The LSTM takes 90000 seconds followed by the Hybrid model at 947 seconds. One of the reasons the CNN took almost 1.5 to 2 times longer to train than the machine learning algorithm is due to having to train over multiple epochs, 128 in our experiment.

Results Table:-

Model	Epochs	Accuracy
CNN	100	100%
RNN	100	97.76
Hybrid(CNN+RNN)	150	84.69

Fig. 6. Results Table

11. CONCLUSION

Handwritten character recognition is a very vast area of research and has many applications. For example the medical field is transitioning to electronic records very fastly. But most of the history of records are still paper based. This can be used as a base to convert and store those records electronically through little or no manual effort. There are many other applications of this system in various fields.

12. REFERENCES

[1] Parikshith H, Naga Rajath S M, Shwetha D, Sindhu C M, Ravi P, "Hand- written Character Recognition of Kannada Language,Using ConvolutionalNeural Networks and Transfer Learning."IOP Publication 2021.

[2] Sri. YugandharManchala,"Handwritten Text Recognition using Deep Learning with TensorFlow."IJERT 2020.

[3] Shalaka Prasad Deore,AlbertPravin,"Devanagari Handwritten Character Recognition using finetunedDeepConvolutional Neural Network on trivialdataset."Springer 2020.

[4] R. Jayakanthan, A. Hiran Kumar, N. Sankarram, B. S. Charu- latha, Ashwin Ramesh,"HandwrittenTamil Character Recognition Using ResNet."IJRESM 2020.

[5] J. Memon, M. Sami, R. A. Khan and M. Uddin, "Handwritten Optical Character Recognition(OCR): A Comprehensive Systematic Literature Review (SLR)." IEEE 2020.

[6] Mahmoud Shams, Amira.A.Elsonbaty ,Wael.Z.ElSawy , "Arabic Hand- written Character Recognition based on Convolution Neural Networks and Support Vector Machine." IJACSA 2020.

[7] I Khandokar, Md M Hasan, F Ernawan, Md S Islam, M N Kabir, "Hand- written Character Recognition Using Convolutional Neural Network." ICMSE 2020.

[8] B. Alrehali, N. Alsaedi, H. Alahmadi and N. Abid, "Historical Ara- bic Manuscripts Text RecognitionUsing Convolutional Neural Net- work."IEEE 2020.

[9] Seyed Ali Asghar AbbaszadehArani,Ehsanollah Kabir, Reza Ebrahim- pour,"Handwritten FarsiWord Recognition Using NN-Based Fusion of HMM Classifiers with Different Types of Features."World Scientific Publication 2019.

[10] NagenderAneja and Sandhya Aneja, "Transfer Learning using CNN for Handwritten DevanagariCharacter Recognition." IEEE2019.

[11]] R. Vijaya Kumar Reddy, U. Ravi Babu, "Handwritten Hindi Character Recognition using DeepLearning Techniques."IJCSE 2019.

[12] R. Fernandes and A. P. Rodrigues, "Kannada Handwritten Script Recog- nition using MachineLearning Techniques." IEEE 2019.

[13] Shahbaz Hassan, Ayesha Irfan, Ali Mirza, Imran Siddiqi,"Cursive Hand- written Text Recognitionusing Bi-Directional LSTMs: A case study on Urdu Handwriting," IEEE 2019.

[14] B. Dessai and A. Patil,"A Deep Learning Approach For Optical Char- acter Recognition Of Handwritten Devanagari Script."IEEE 2019.

[15] E. Shaikh, I. Mohiuddin, A. Manzoor, G. Latif, and N. Moham- mad,"Automated grading forhandwritten answer sheets using convolu- tional neural networks."IEEE 2019.

[16] H. Zhang and Z. Cheng,"An Advanced Pyramid Network Technology For Optical CharacterRecognition."IOP Publication 2019

[17] Ali Ahmed Ali, Amani Mallaiah, Suresha. "A New Design Based-Fusion of Features to RecognizeArabic Handwritten Characters." IJEAT 2019.

[18] Maroua Mehri, Pierre Héroux, Rémy Mullot, Jean-Philippe Moreux, Bertrand Couasnon,"Competition on Historical Book Analysis."HAL Open Science 2019.

[19] K. G. Joe, M. Savit, and K. Chandrasekaran,"Offline Character Recog- nition On SegmentedHandwritten Kannada Characters."IEEE 2019.

[20] V. Jayasundara, S. Jayasekara, H. Jayasekara, J. Rajasegaran,S. Senevi- ratne, and R. Rodrigo,"Handwritten Character Recognition With Very Small Datasets."IEEE 2019.

[21]] P. Sahare and S. B. Dhok,"Robust Character Segmentation And Recognition Schemes For Multilingual Indian Document Images."IETE 2019.

[22] Vishal Passricha, Rajesh Kumar Aggarwal, "A Comparative Analysis Of Pooling Strategies ForConvolutional Neural Network Based Hindi ASR."Springer 2019.

[23] Nishatul Majid, Elisa H. Barney Smith, "Introducing the Boise State Bangla Handwriting Datasetand an Efficient Offline Recognizer of Iso- lated Bangla Characters." IEEE 2018.

[24] Nikita Singh, "An Efficient Approach for Handwritten Devanagari Character Recognition basedon Artificial Neural Network."IEEE 2018.

[25] R. Ptucha, F. P. Such, S. Pillai, F. Brockler, V. Singh, and P. Hutkowski,"Intelligent CharacterRecognition Using Fully Convolutional Neural Networks."Elsevier 2018.

[26] T. C. Wei, U. U. Sheikh and A. A. H. A. Rahman, "Improved Optical Character Recognition WithDeep Neural Network."IEEE 2018.

[27] Z. Li, L. Jin and S. Lai, "Rotation-Free Online Handwritten Chinese Character Recognition usingTwo-Stage Convolutional Neural Network." IEEE 2018

[28] Khaled S. Younis, "Arabic Handwritten Character Recognition Based On Deep ConvolutionalNeural Networks."Springer 2017.

[29] P. Dhande and R. Kharat, "Recognition Of Cursive English HandwrittenCharacters," IEEE 2017.

[30] RiteshSarkhel, Nibaran Das, Aritra Das, Mahantapas Kundu, Mi- taNasipuri,"A Multiscale DeepQuad Tree Based Feature Extraction Method For The Recognition Of Isolated HandwrittenCharacters Of Popular Indic Scripts."Elsevier 2017.