ISSN 2229-5518

A Convolutional Neural Network based Approach for Recognizing Malayalam Handwritten Characters

Jabir Ali V, Jose T Joseph

Abstract—Optical character recognition has leveraged its capabilities to reduce tedious manual work of converting images containing characters to texts for recent decades, recognition of handwritten texts is harder than recognizing printed texts as different handwrittenscript may have different style of writings, slants etc, The Convolutional Neural Network(CNN) has been successfully used to recognize characters in many languages. This paper proposes a CNN architecture for classification of handwritten characters in Malayalam language. . Malayalam is a south Indian language which is used by 33.3 million people in the state of Kerala. This CNN Model has shown a testing accuracy of 97.26% for the classification of 44 handwritten Malayalam characters by using a dataset having seventy three thousand training images and eighteen thousand testing images. In addition, this paper put forward an algorithm to process an image containing handwritten Malayalam characters and output the corresponding Malayalam characters

Index Terms— Classification, Convolutional Neural Network, Dataset Augmentation OCR, Optimizer, Segmentation

1 Introduction

OPTICAL character recognition is the process of converting images containing printed, typewritten or handwritten characters into machine encoded format. OCR has leveraged its capabilities to reduce the tedious manual work work of converting images of printed or handwritten texts to digital form for the recent decades. Different methods have been used in OCR such as Bayesian theory, Hidden Markov Model, Template Matching and Neural Networks. The recognition task of handwritten characters is rather complex and is a great challenge to researchers as the solution should be able to cope with the challenges in identifying the characters from a variety of writing styles, slants of personal interest. Especially in a language like Malayalam which is having a complex structure and very identical nature of character set.

Convlolutional Neural Network (CNN) is a popular deep learning method which has successfully used in many classification problems, It has a great capability to find patterns in two dimensional data and were applied in the area of natural language processing, image classification, face recognition, autonomous driving etc. The proposed method uses a new CNN architecture to do the classification task. Unlike other classical machine learning such as SVM and Random Forest, the feature extraction task is implicitly done in CNN by using the gradient descent algorithm proposed by Yann Le Cunn.This paper also propose method to do the task of segmentation of words and characters from a document image and to make character prediction from the CNN Model created.

• *Jabir Ali is currently pursuing masters degree program in computer applications in College of Engineering Trivandrum,Kerala University ,India , E-mail: jabir366@gmail.com,Jose T Joseph is currently working as Head of the department and Assistant professor, Department of computer applications, College of Engineering Trivandrum ,Kerala University ,India*

2 Malayalam Language

Malayalam is the main language of kerala, a south Indian state. Malayalam is spoken by around 33 million people in kerala and it is one among the twenty two scheduled languages in India. Like other south Indian languages such as Tamil, Kannada and Telugu, Malayalam language has a complex structure. It contains core characters and vowel diacritics. In addition to this, the script is formed mostly by curves and holes. There are characters which looks very similar and having only small distinguishable difference. The modern Malayalam script contains 13 vowel letters and 36 consonants. Apart from this basic characters there are characters which are formed by the combination of other consonants called conjunct consonants. Here a dataset of 44 basic Malayalam characters shown in Fig 1.

3 Literature Review

Shailesh Acharya et al[2] proposed a Large Scale Handwritten Devanagari Character Recognition that used The Deep learning technique convolutional neural network for classifying Devanagari handwritten characters. The employed increment of dataset and added a drop out layer in order to reduce overfitting. They used two models of the network. First model consisted of three convolution layers and one fully connected layer and the second model was a shallow network. The highest testing accuracy for first model was 0.982681 and for the second model was 0.98471.

G Raju et al[3] put forwarded a Malayalam character recognition system which uses gradient based features and Run Length count. The authors also have proposed another

ISSN 2229-5518

character recognition scheme using the fusion of local and global features for the recognition of isolated Malayalam characters. Arora et al. proposed a multiple classifier system using chain code histogram and moment invariants for the recognition of Devanagari character recognition.

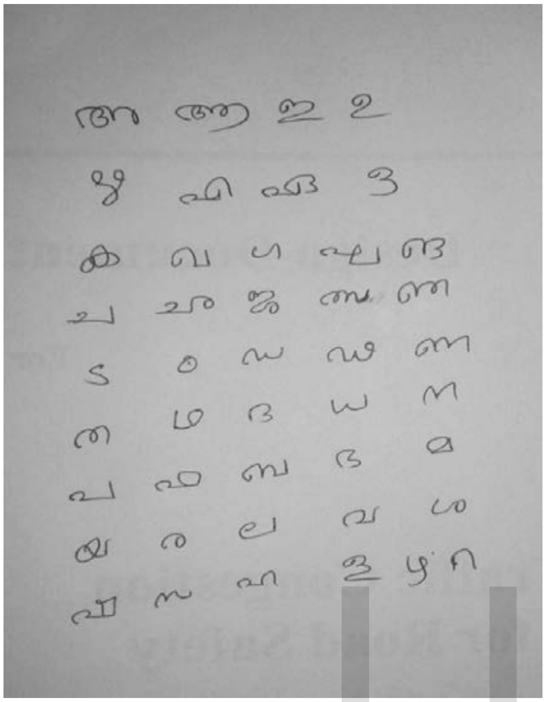

Fig.1 Malayalam characters

Pranav P Nair et al [1] proposed a method for Malayalam Character recognition using CNN. They tested the CNN by training with a Malayalam Dataset constructed by their own; they achieved comparatively high accuracy for the six Malayalam characters.

4 Methodology

4.1 Deep Learning

Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called Artificial neural networks (ANN). In deep learning we don't need to do the manual feature extraction part that we do in classifiers like Random Forest,SVM etc. The ANN extracts features by its own. For character classification This paper propose a unique Convolutional neural network (CNN).

4.2 Convolutional Neural Network (CNN)

Convolution neural network is a special kind of Artificial Neural Network (ANN). It is a feed forward multilayer perceptron which has proven its capabilities for identification of patterns from two dimensional data. It has successfully been applying in problems such as image classification, natural language processing etc.The main difference between a CNN and an ANN is that CNN uses parameter sharing which makes the computation a lot easier. A CNN consists of one input layer, convolution layers, pooling layers, reLU(Rectified Linear Unit) layers,fully connected layers and one output layer. In addition to this, the number of layer in a CNN is subjected to change according to the architecture used and the classification requirements. A simple CNN is shown in Fig 2.

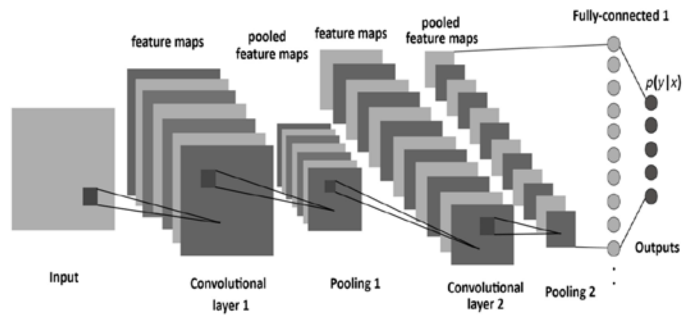

Fig 2. Convolutional Neural Network

5 Proposed System

The proposed system uses a unique CNN Model for the classification task. The overall system architecture is shown in Fig.3 is The system can be effectively used in situations where we need to digitize a Malayalam handwritten script . by using this intelligent system to recognize the handwritten characters and make a digital equivalent of the scripts, we can greatly reduce the manual effort of character recognition and data entry for digitizing a handwritten document .The system consists of four phases:

1. Dataset Creation and Augmentation
2. Defining CNN architecture
3. Training the CNN Model
4. Deploy the Model

5.1 Dataset Creation and Augmentation

Creating a dataset is time consuming and requires a lot of effort. There is no open source dataset available for handwritten Malayalam characters. For this project the dataset was collected from a private organization and it is modified. The dataset consisted images of 44 Malayalam characters. A large dataset is required for training the CNN. In order to attain this, the images that are already obtained is modified. for the images in the dataset, affine transformations such as translation, rotation ,scaling and shearing are applied to get a large number of variations. Fig.4 shows the overall flow of this process.

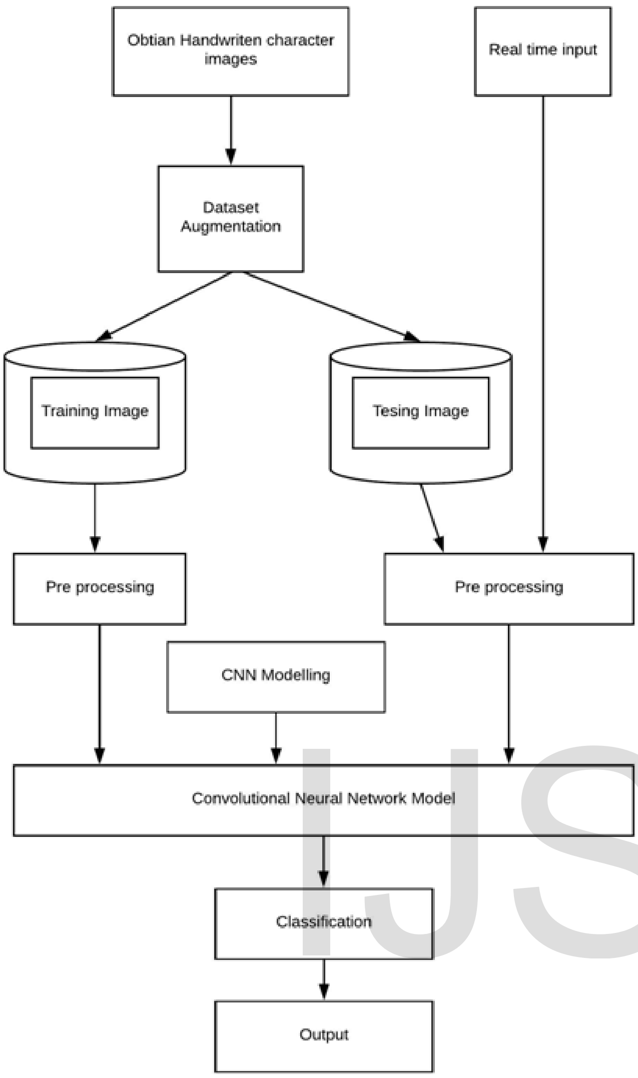

Fig.3 System Architecture

Affine transformation is a linear mapping method that preserves points, straight lines, and planes. Sets of parallel lines remain parallel after an affine transformation. Different translations are used to augment the dataset Gaussian blurring is the result of blurring an image by a Gaussian function. The visual effect of this blurring technique is a smooth blur resembling that of viewing the image through a translucent screen. Salt-and-pepper noise is a form of noise sometimes seen on images. It presents itself as sparsely occurring white and black pixels. Contrast and brightness level of an image is changed. After data augmentation a dataset of 91,902 images were obtained

5.2 Defining The CNN Architecture

A typical Convolutional Neural Network mainly include three types of layers, namely Convolutional layer, pooling layer non linearity layer and fully-connected layer in addition to the input and output layer. There are many standard CNN architectures available which are proven their classification capabilities in many tasks ,eg. LeNet , Alexnet,VGGNet,Inception etc. all these architecture differ due to the difference in number of hyper parameters of the network such as number of convolution layers, pooling layers ,fully connected layers ,the number of filters used in each layers, dropout rate at each layer, L2 or L1 regularization parameters, activation function type (ReLU, Sigmoid, Tanh etc.). After some researches and tries, a CNN architecture is defined for our classification task.

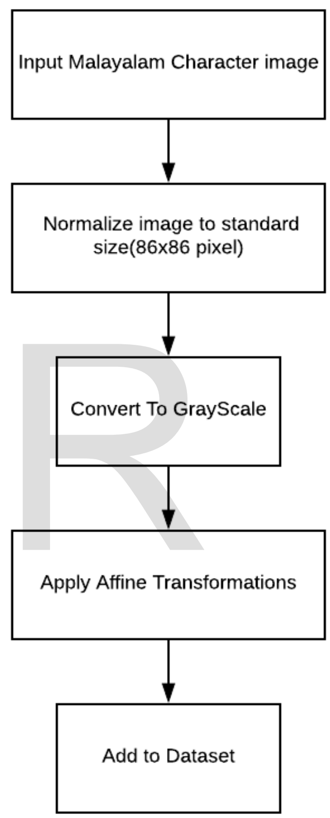

Fig.4 Dataset creation and augmentation

The input size of images dataset is fixed to 86x86x1, so the initially the network is designed to occupy tensors(3D arrays here) of shape [86,86,1].The architecture of CNN model is shown in Fig.5.This model uses 3x3 filters for convolution,2x2 filters for max-pooling and ReLu(Rectified Linear Unit) as activation function in Non linearity layer. Batch normalization is applied to achieve global minimum faster, Adam is used as optimizer for gradient descent algorithm to perform, During training .a single image of a Malayalam character with one-hot encoded label is passes through all the layers and according to the applied gradient descent strategy here, the weights are updated. The output layer contains 44 classes each represents a character in our dataset.The Architecture of CNN Model is as follows

1. .Input layer of size 86x86x1
2. Convolution layer of 32 filters of size 3x3 with zero padding and ReLu activation function
3. Max-pooling layer of 2x2 filter size
4. batch normalization
5. Convolution layer of 64 filters of size 3x3 with zero padding and ReLu activation function
6. Max-pooling layer of 2x2 filter size
7. batch normalization
8. Convolution layer of 128 filters of size 3x3 with zero padding and ReLu activation function
9. Max-pooling layer of 2x2 filter size
10. batch normalization
11. Convolution layer of 256 filters of size 3x3 with zero padding and ReLu activation function
12. Max-pooling layer of 2x2 filter size
13. fully connected layer of 512 neurons
14. drop-out(0.25)
15. fully connected layer of 1024 neurons
16. drop-out(0.5)
17. fully connected layer of 512 neurons
18. drop-out(0.5)
19. output layer of 44 neurons

5.2 Training the CNN Model

The model has to be trained with the dataset created. For this, we need to prepare the dataset to input and labels format so that the model understands it. for this. The dataset is processed and a list is created which consists of numpy array of images along with the one-hot encoded labels. The processed dataset is divided into training set and Testing set in the ratio of 80:20 . and again the training set is further divided into training and validation sets in the same 80:20 ratio, there were 58816 images for training,14705 images for validation and 18381 images for testing. Validation set used here to test during the training time itself so that we can discern whether our model overfits or not during each epoch. An epoch is the complete iteration of training over the entire dataset. To reduce the complexities that may occur due to memory inefficiencies while training, the entire dataset is divided into batches and each batch is given to the network for training. After successful training,the model weights are saved.

5.2 Deploy the Model

After successful training, the Model has to be deployed in such a way that the user can input a real time images of Malayalam handwritten script and make predictions using our saved CNN Model. the flow chart in Fig.6 describes how the model can be implemented in real time. The image of Malayalam handwritten character has to be read. After applying proper segmentation mechanisms, the words are separated and then characters are separated. Each character that is segmented are given to model and prediction is made. For each predicted class the character is mapped to corresponding Malayalam Unicode character.The Segmentation algorithm is as follows

1. Read image
2. Convert image into grayscale
3. Apply morphologyEx operation with structuring element of kernel sizes suitable to get the words as a single contour
4. Store the contour bounding boxes as a list
5. For each bounding box ,apply morphologyEx operation with structuring element of small kernel size
6. Find contours and draw bounding boxes
7. Sort the bounding boxes in the increasing order of x axis
8. Crop each bounding box coordinate from input image and store it in a list

6 Results and Discussions

The CNN model is created and the hyper parameters are tuned to get the increased accuracy. This model gave a training accuracy of 97.14 with loss 0.3718, validation accuracy of 96.04 with loss 0.3242 and a testing Accuracy of 97.26**.** The proposed system is successfully implemented in realtime. The implementation is done using Keras with Tensorflow as backend. For image aquistition and segmentation tasks,OpenCV is used.The prediction of realtime input is shown in Fig 5

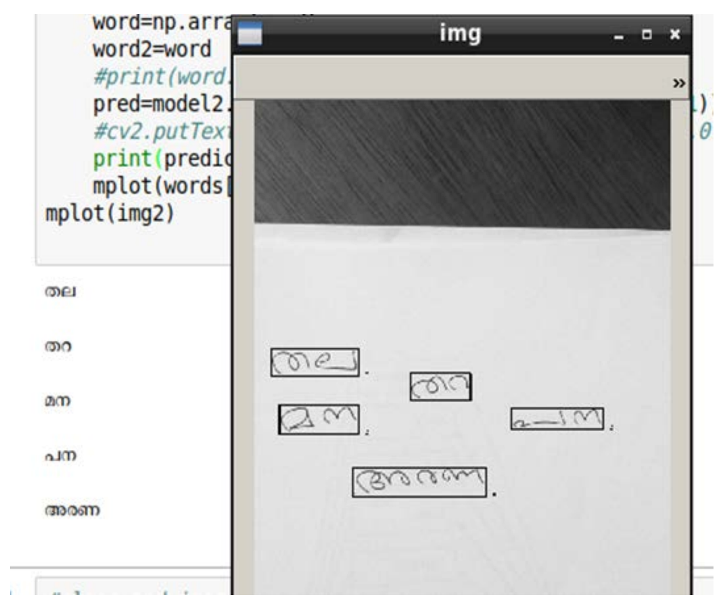

Fig 5.Prediction of real time input

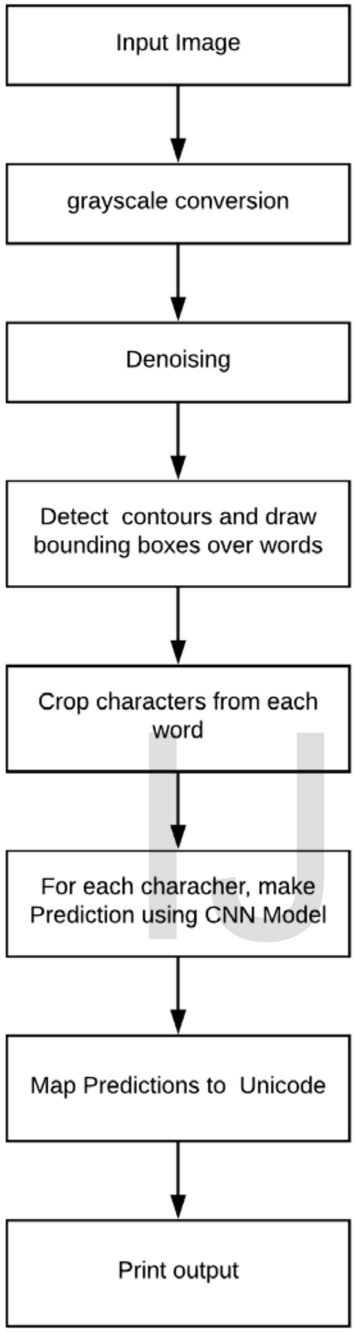

Fig.6 Deployment of the Model

6 Conclusion

In this paper, we have proposed a Convolutional Neural Network Model for classifying Malayalam handwritten characters. This Model is trained using a dataset having over 90,000 images of 44 malayalam handwritten characters and an accuracy of 97.26% is achieved in testing the model.

An Algorithm for processing real time input images containing Malayalam handwritten characters is also proposed in this paper. It includes image acquisition, grayscale conversion,binarization, word segmentation,character segmentation and prediction. Further refinement of the system includes improving The CNN Model with all the glyphs and diacritics in the Malayalam language. The future scope of this work lies in proper recognition of characters from connected writings and recognition of vowel diacritics from a script.

Acknowledgment

First of all, we thank to almighty for successful completion of this research work. Moreover, we express our deep gratitude to our family, friends and teachers for their whole hearted support

References

[1] Pravav P Nair, Ajay James and C Saravanan. "*Malayalam Handwritten Character Recognition using Convolutional Neural Network*", International Conference on Inventive Communication and Computational Technologies, October-2017

[2] Shailesh Acharya, Ashok Kumar Pant, Prashnna Kumar Gyawali "*Deep Learning Based Large Scale Handwritten Devanagari Character Recognition*", 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), 2015.

[3] Raju, Bindu S Moni, Madhu S. Nair, "*A Novel Handwritten Character Recognition System Using Gradient Based Features and Run Length Count*", Sadhana Indian Academy of Sciences, Springer India, 2014, p.1-23.

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton. "*Imagenet classification with deep convolutional neural networks*", In Advances in neural information processing systems, pages 1097–1105, 2011

[5] Abdul Rahiman M and Rajasree M S, "*An Efficient Character Recognition System for Handwritten Malayalam Characters Based on Intensity Variations*", International Journal of Computer Theory and Engineering, Vol. 3, No. 3, June 2011

[6] Yann LeCun, Leon Bottou, Yoshua Bengio and Patric Haffner,"*Gradient-Based Learning Applied to Document Recognition*", PROC.OF THE IEEE, NOVEMBER 1998