Telugu Alphabet Classification

1. Introduction

In present days there are many interactive learning apps for kids to start learning alphabets with fun on mobile / tablet screens. The majority of apps are limited to the English language that has a set of 26 alphabets. These alphabets can be easily drawn with slant lines, sleeping and standing lines, curves (left, right, upward and downward), and hooks (left, right).

Coming to Indian languages, the Telugu language is a prominent south Indian language. One of the reasons why handwriting recognition is relatively difficult for Telugu scripts when compared with other languages is the complexity of the orthography in Telugu. Telugu is a syllabic language and contains 41 consonants, 16 vowels, and 3 vowel-modifiers whose combinations form more than 10,000 syllabic orthographic units. Telugu characters consist of rounded shapes and very few vertical strokes. The orthography is composed with various vowel sound syllables (maatras) modifying the basic consonants. Vottus are symbolized pure consonant sounds and can be combined with other consonant or vowel-modified consonant symbols. Vottus and maatras can be positioned at locations surrounding the base character (shown in Fig.1). Telugu proves to be one of the challenging scripts for handwriting recognition.

Fig. 1 Four characters showing few combinations of vottus and maatras

Dataset plays an important role in the design of supervised learning systems. Datasets available for Telugu characters are very limited and not available as open source, hence a dataset is collected from different sources consisting of handwritten characters. When it comes to handwritten text, each person may write the letters differently, even the same person may write it differently at times. Coming to kids who just started to learn, writing letters are not fixed to width or height and even to the shape of a letter. A kid's handwritten English alphabets match to their original ones, but the handwritten telugu alphabets barely match. Then the collected dataset is pre-processed. In this project, Convolutional Neural Networks (CNNs) are used to identify the handwritten characters. CNNs have revolutionised computer vision and pattern recognition, mainly for offline handwriting recognition. The main advantage of this neural network is that the model can learn from the previous learned data and future predicted data. It can learn while working or predicting.

2. Related Work

Many of the research works focused on Support Vector Machine (SVM), Neural Networks (NN), k-Nearest Neighbour (k-NN) for classification with statistical, topological and structural features for character recognition.

In one of the papers, they proposed a hybrid approach for recognition of handwritten Telugu CV type characters, as offline images, using a combination of CNN, PCA, SVM and multi classifier system techniques. The best individual classifier trained on MNIST data yielded a test performance of 98.25% while the ensemble classifier for the same data achieved a test performance of 98.5%.

Riya Guha [1] in their paper proposed a new CNN architecture called Dev-net architecture to recognize Hindi characters. It is a 6-layer CNN architecture which requires less time and memory space as compared to the existing CNN models. It has achieved 99.6% accuracy. It also performed better than already existing CNN architectures like LeNet, ResNet, AlexNet etc. The architecture of the proposed recognition scheme for palm leaf scribed text. It is divided into three major phases: 3D data acquisition, Pre-processing, and Character recognition.

K.C. Movva [2] proposed that the usage of distributions of few properties of stroke points like local variance and local moments can be used to represent a stroke and also suggested some features to improve the shape recognition scheme. A stroke could be viewed as a sequence of points from pen-up to pen-down. Each Telugu character is composed of some strokes and thus individual stroke identification and classification has led them to individual character classification. In this paper a K-NN and A-NN based models were compared. They found that the artificial neural network (ANN) model has been proven to give good classification accuracy even when they use few samples per character which makes it efficient and suitable for handwritten character recognition.

3. Materials and Experimental Evaluation

3.1 Dataset

1. Since datasets for telugu language are limited and are not open, among very few datasets that are available online, for this paper the majority of the dataset is chosen from Kaggle [3]. Telugu alphabets from random websites (google images) were used replacing the random images in dataset that were downloaded earlier

2. Format and Structure(Image dataset) : Dataset is in image format (jpeg)

”Ee”

3. Preprocessing: Resizing images is a difficult step in computer vision. Primarily, our machine learning models train faster if images are smaller. An input image which is two times as large requires our network to learn from four times as many pixels which

increases the time to generate outputs. Also, many deep learning model architectures require all images to be in the same size but our raw collected images may vary in size. So, we resized all our images to a uniform size using Pooling Layers > GlobalAveragePooling2D. This pooling layer downsamples the input along its spatial dimensions (height and width) by taking the average value over an input window for each channel of the input.

4. Number of Classes: There are 5 different classes (“A”, “Aa”, “E”, “Ee”, “U”)
5. Class Distribution: Each class has 200 sample image data.
6. Training and Testing: 80% of the data is used for training, the rest 20% for Validation”

3.2 Methodology

In spite of being popular, ANNs were unable to handle large datasets in recognition / classification tasks. To overcome these, a new machine learning paradigm, deep learning, was introduced. It is a stacked neural network that is composed of several layers.

Earlier versions of neural networks, such as the first perceptron were shallow, composed of one input and one output layer; and one hidden layer in between. In deep-learning networks, each layer of nodes trains on a distinct set of features based on the previous layer’s output. A model will be efficient if hidden layers have the ability to learn complicated features from observed data. Deep neural networks show notable performance on unseen data. Some popular deep neural network architectures are recurrent neural networks (RNNs), CNNs, deep belief networks, auto-encoders and generative adversarial networks. In general, CNNs are considered as a machine learning architecture, which has a capability to learn from experiences like multilayer neural networks with back propagation. For the requirement of minimal pre-processing, CNNs use a variation of the multilayer perceptron. CNNs are composed of an automatic feature extractor and a trainable classifier, having important layers

- Convolutional Layer (CL)
- Pooling Layer (PL)
- Fully-Connected Layer(FCL)

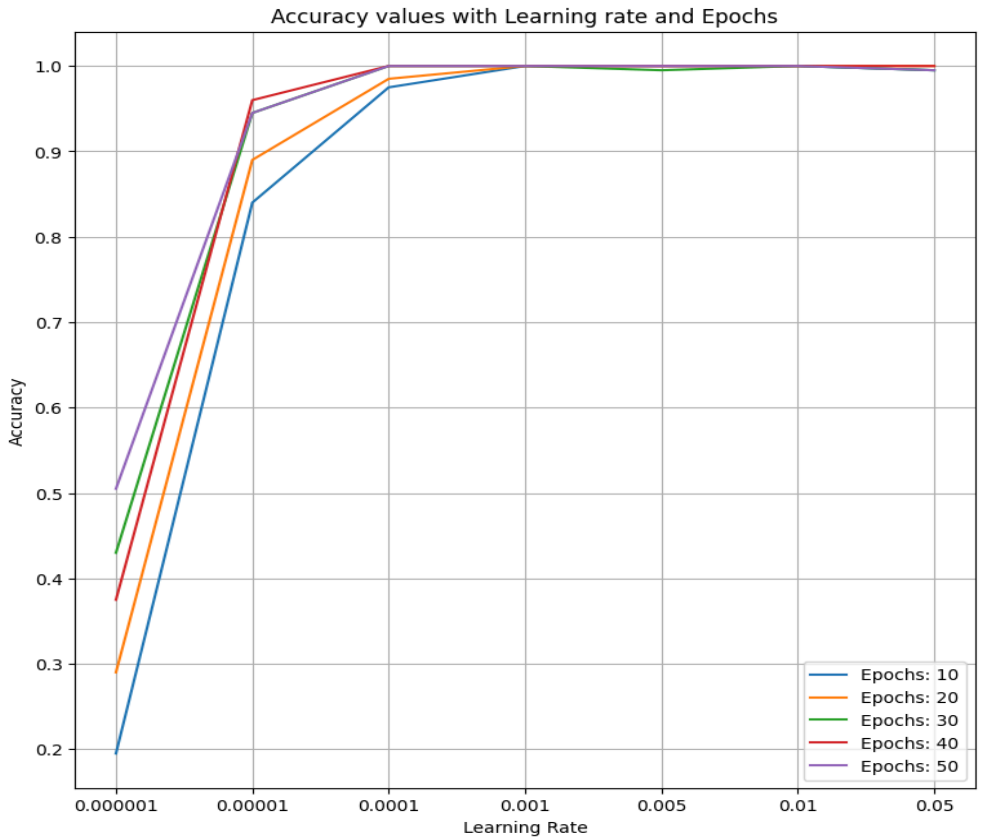
Using Keras deep learning API [4] a data model is created with 100 layers, with Pooling Layer GlobalAveragePooling2D, the final layer activated using “softmax” [5] and the rest of previous layers using “*relu”* [5] (Rectified Linear Unit) activation function.

With every combination of hyperparameters : epochs [10,20,30,40, 50] and learning rates [0.000001, 0.00001, 0.0001, 0.001, 0.005, 0.01, 0.05] accuracy values are noted down.

	Epochs/learning rate	0.000001	0.00001	0.0001	0.001	0.005	0.01	0.05
0	10	0.195	0.840	0.975	1.0	1.000	1.0	0.995
1	20	0.290	0.890	0.985	1.0	1.000	1.0	1.000
2	30	0.430	0.945	1.000	1.0	0.995	1.0	0.995
3	40	0.375	0.960	1.000	1.0	1.000	1.0	1.000
4	50	0.505	0.945	1.000	1.0	1.000	1.0	0.995

Graphs have been extracted using tensorflow [6] with "*adam*" [7] as optimizer. The graph below presents the accuracy values with learning rate and epochs.

It was observed that at learning rate 0.001, accuracy reaches 1 with whatever epoch applied. While presenting this paper, epoch(10) and learning rate (0.001) are chosen to develop an app using streamlit [8].

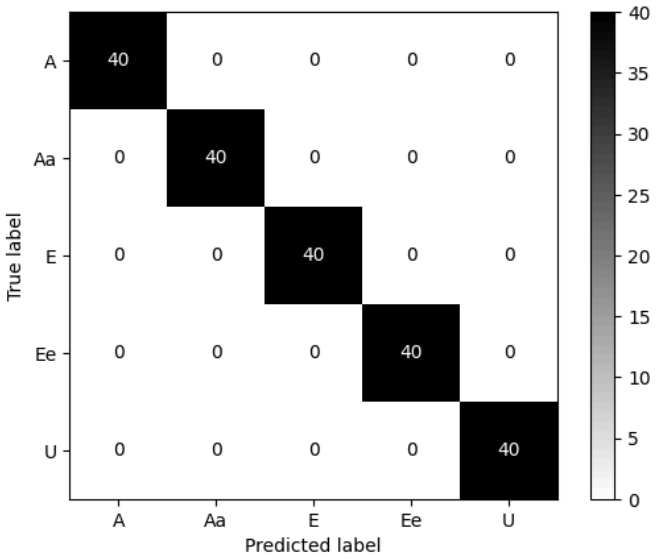


3.3 Results

All the experiments are performed over a standard dataset through Google Colab. The standard dataset contains 200 samples from each of the 5 categories in the Telugu character Dataset. Both the accuracy of the model for the training dataset and the accuracy of the trained model for testing data was shown 100%.

Following are the results for the dataset, hyperparameters chosen.

Confusion Matrix [6]:



Classification Report :	precision	recall	f1-score	support
0	1.00	1.00	1.00	40
1	1.00	1.00	1.00	40
2	1.00	1.00	1.00	40
3	1.00	1.00	1.00	40
4	1.00	1.00	1.00	40
accuracy			1.00	200
macro avg	1.00	1.00	1.00	200
weighted avg	1.00	1.00	1.00	200

3.4 Discussion

Keras, as a multi-framework API, can be used to develop modular components that are compatible with any framework such as TensorFlow, PyTorch, etc..

TensorFlow makes it easy to create ML models that can run in any environment. According to Kingma et al., 2014, the “adam” optimizer method is "*computationally efficient, has little memory requirement, invariant to diagonal rescaling of gradients, and is well suited for problems that are large in terms of data/parameters*".

4. Future Work

Implementing KerasTuner can easily configure search space with a define-by-run syntax, then leverage one of the available search algorithms to find the best hyperparameter values

for this model. KerasTuner is designed to be easy for researchers to extend in order to experiment with new search algorithms.

The present work is only with the first five alphabets which can be extended to the rest of them. Also the collected dataset mostly comprises handwritten alphabets by adults rather than children. This paper can help to build apps such as interactive learning apps for kids where the app is like a game asking the kid to draw an alphabet on screen then compares it with the deep machine learning knowledge for correctness. .

5. Conclusion

This paper presented a comprehensive and practical CNN system for Telugu Language. The proposed system is shape and edge dependent and requires pre-processing and feature extraction. The experimental result shows the performance characteristics of the Convolution Neural Network.

With Keras multi framework API, activation functions *relu* and *softmax*, and Training data - Data validation ratio 80:20, at learning rate 0.001, best accuracy can be achieved regardless of epoch rate. This attempt believes that this would extend to even more alphabets / classes added to the dataset.

Based on the output of the model, the dataset can be improved by including other alphabets, *vothulu* and *matralu* which are extensions to the Telugu alphabet. In future, this algorithm can be extended from character recognition to text recognition by creating our own dataset.

6.Reference

1. Guha R, Das N, Kundu M, et al. DevNet: An efficient CNN architecture for handwritten devanagari character recognition. International Journal of Pattern Recognition and Artificial Intelligence 2019; 34(12): 20052009. doi: 10.1142/S0218001420520096
2. K. C. Movva and Viswanath P, "A Moment-Based Representation for Online Telugu Handwritten Character Recognition, Data Analytics and Learning Conference, University of Mysore, November 2018.
3. DataSet Source: https://www.kaggle.com/datasets/syamkakarla/telugu-6-vowel-dataset
4. About Keras Model https://keras.io/about/
5. Activation Function for Keras Model https://keras.io/api/layers/activations/
6. About TensorFlow Graphs https://www.tensorflow.org/
7. Adam optimizer function https://www.tensorflow.org/api_docs/python/tf/keras/optimizers/Adam
8. https://streamlit.io/