RESEARCH ARTICLE OPEN ACCESS

Handwritten Tamil Character Recognition Using Deep Neural Networks

Khanaghavalle G R [1], Arvind Nachiappan L [2], Bharath Vyas S [3], Chaitanya M [4]

[1] Assistant Professor, Dept. Of Computer Science and Engineering, Sri Venkateswara College of Engineering,

[2],[3],[4] B. E Students, Dept. Of Computer Science and Engineering, Sri Venkateswara College of Engineering - Sri Perumbudur

ABSTRACT

Handwritten Tamil Character Recognition plays a huge role in digitization and documentation of official documents which are handwritten. Tamil Handwritten character recognition is still in its primitive stages. It is observed that this setback is due to lack of robust datasets. Unconstrained Tamil Handwritten Character Database (uTHCD) a relatively new dataset which pushes the scopes of Tamil character recognition technology to a significant extent. The dataset has overcome class imbalance and is unconstrained as it has both online and offline samples having real variations of handwritten letters such as discontinuity and varied strokes. The paper projects the attempt to generate a deep learning model to perform handwritten character recognition of a single character provided using Unconstrained Tamil Handwritten Character Database (uTHCD) dataset. There are 156 unique sets of classes, each class a unique Tamil character of itself. Convolutional Neural Network is employed in this process as it plays a huge role in recognizing patterns in images. The model using which the Convolution Neural Network (CNN) model developed is Deep Neural Network. A deep neural network (DNN) is an artificial neural network (ANN) with multiple layers between the input and output layers. Deep Neural Networks can model complex non-linear relationships.

***Keywords:* - Deep** Neural Network, Unconstrained Tamil Handwritten Character Database (uTHCD)

I. INTRODUCTION

Machine learning is a subset of Artificial Intelligence involving the application of computer algorithms that employs computation processes that can improve automatically through experience by the use of data. Machine learning algorithms build a model based on input data, known as "training data", in order to make predictions or decisions without being explicitly programmed. Although Machine Learning (ML) predominantly involves applied statistics, The scope of machine learning is beyond them. Machine learning algorithms are used in cases and situations where the rules for evaluation are not discrete and definite.

Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks. Deep Learning (DL) involves replicating human neurons by establishing a system of nodes which act as a single unit of neuron, Each node having assigned a weight bias is adjusted each time when it is trained using training data . Deep-learning architectures have been applied to fields including computer vision, speech recognition, natural language processing, machine translation, bioinformatics, drug design, medical image analysis and material inspection etc.

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image,assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. Convolutional networks were inspired by the working of neurons in the brain where each neuron acts as a node which transmits an electrical pulse via which the brain functions. A CNN uses a system much like a multilayer perceptron that has been designed for reduced processing requirements. The layers of a CNN consist of an input layer, an output layer and a hidden layer that includes multiple convolutional layers, pooling layers, fully connected layers and normalization layers. Convolutional Neural Network is predominantly used in image recognition since images can be converted to a sequence of matrices and passed as an input.

Optical Character Recognition (OCR) is defined as the computation process of identifying handwritten analog characters. The characters are scanned through an electronic medium and converted to a digital format which is converted to the respective characters using suitable computer algorithms. Optical Characters Recognition are used in places where there is a significant necessity of digital format of a document or data.

II. RELATED WORKS

A. URDU-TEXT DETECTION AND RECOGNITION IN NATURAL SCENE IMAGES USING DEEP LEARNING

Urdu texts are considered as a sequence of primary characters in Urdu script with 32 basic isolated characters. Recognition involves rate sequence matching was calculated for given ligature images. Urdu text poses a challenge for detection/localization from natural scene images, and consequently recognition of individual ligatures in scene images. Methodology is proposed that covers detection, orientation prediction, and recognition of Urdu ligatures in outdoor images. The proposed methodology covers the essential phases of PhotoOCR which is presented and evaluated using various sub-architectures. These phases include the detection of text, Orientation determination of text, and finally recognizing the written text in outdoor images. Using Custom Faster RCNN builder, A CNN model-selection module and the other is the training module, which trains Faster RCNN using the features from the selected CNN model.

B. A NEW HYBRID CONVOLUTIONAL NEURAL NETWORK AND EXTREME GRADIENT BOOSTING CLASSIFIER FOR RECOGNIZING HANDWRITTEN ETHIOPIAN CHARACTERS

CNN models are to be used in three different ways. Training CNN from scratch, Transfer learning strategy to leverage features from a pre-trained model on a larger dataset and Keeping the transfer learning strategy and fine-tuning the weights of CNN architecture. A novel hybrid CNN–XGBoost model is proposed to solve the handwritten scripts recognition problem. In this integrated model, CNN works as a trainable automatic feature extractor from the raw images, whereas XGBoost performs the recognition part. CNN trained using backpropagation in order to extract features. The extracted features are given to the XGBoost for classification.

C. TWO-STEP CNN FRAMEWORK FOR TEXT LINE RECOGNITION IN CAMERA-CAPTURED IMAGES

Optical text recognition using mobile devices can be classified into two groups: client-server solutions, which transfer images to a ‘‘cloud’’ and require internet connection, and ‘‘on the device’’ methods that perform the recognition process without data transmission. The paper proposes a ‘‘on the device’’ text line recognition framework considering per-character segmentation as a language-independent problem and individual character recognition as a language-dependent one. Experimentation with the classic MNIST dataset and acquire the results comparable with the state-of-the-art ones, Based on two separate artificial neural networks (ANN) and dynamic programming instead of employing image processing methods for the segmentation step or end-to-end ANN. The primary purpose of our framework is the recognition of low-quality images of identity documents with complex backgrounds and a variety of languages and fonts.

D. HMM-BASED LEXICON-DRIVEN AND LEXICON-FREE WORD RECOGNITION FOR ONLINE HANDWRITTEN INDIC SCRIPTS

An overview of an approach to developing a largely data-driven and script-independent online handwritten word recognition system for Tamil Words based on HMMs. In contrast to previous approaches, the techniques we propose are largely data driven and script independent. We propose two different techniques for word recognition based on Hidden Markov Models (HMM).The various stages in implementing our approach such as symbol set definition and data set creation, and steps in recognition such as preprocessing and feature extraction. The lexicon-driven technique models each word in the lexicon as a sequence of symbol HMMs according to a standard symbol writing order derived from the phonetic representation. The lexicon-free technique uses a novel Bag-of-Symbols representation of the handwritten word that is independent of symbol order and allows rapid pruning of the lexicon.

E. TEXT DETECTION AND RECOGNITION FOR IMAGES OF MEDICAL LABORATORY REPORTS WITH A DEEP LEARNING APPROACH

Digitization of medical reports involves challenges such as low quality scanning ,tarnished images and noises scanned along with the image .Given an image of a medical laboratory report, first, a patch-based training strategy is applied to a detector that outputs a set of bounding boxes containing texts. Then a concatenation structure is inserted into a recognizer, which takes the areas of bounding boxes in source image as inputs and outputs recognized texts. The approach consists of two modules: text detection and recognition. In text detection, a patch-based training strategy is applied, which can achieve the recall of 99.5% in the experiments. For text recognition, a concatenation structure is designed to combine the features from both shallow and deep layers in neural networks which can improve the accuracy of multilingual text recognition.

F. RMAF: RELU-MEMRISTOR-LIKE ACTIVATION FUNCTION FOR DEEP LEARNING

Activation functions facilitate deep neural networks by introducing non- linearity to the learning process. The non-linearity feature gives the neural network the ability to learn complex patterns. To apply the effectiveness and performance of the proposed RMAF transfer function by comparing with state-of-the art activation functions such as sigmoid, Tanh, ReLU, ELU, SELU, PReLU and Swish. The focus is finding a scalar activation function, which takes in scalar input and output a scalar. This is because scalar activation functions can replace the ReLU function without changing the architectural network. The Methodology used here is a cost function called ReLU which is one among many cost functions available used for increasing the performance of our Convolution Neural Network. The RMAF function can be any point utilized when fitting data and giving knowledge to the basic forms dependable on its dynamics.

III. METHODOLOGY

Handwritten character recognition technology in Tamil language is underworked. This is mainly due to lack of availability of robust dataset. Traditional datasets which are used to train the model for Tamil character recognition are relatively primitive and have some significant shortcomings regarding implementation as they are constrained format. The constrained dataset lacks strokes of variable width and discontinuous pattern that are practically encountered during the optical character recognition of handwritten characters. Such limitations in the primitive datasets are mainly due to their mode of collection, that is they are collected through online means using digital pens. The project is based on the unconstrained Tamil Handwritten Character Database (uTHCD) dataset which has overcome the previously specified shortcomings by collecting the samples from both online and offline mode. Objective is to develop a Deep Learning model using Deep Neural Network that can recognize the handwritten character in Tamil and turn to digital characters.

G. Architecture

The model consists of 2 convolutional layers and 2 max pooling layers. In the first convolutional layer, the dimensions are 62x62x32. The max pooling layer takes a kernel 2, and a stride 2 which halves the dimensions to 31x31x32. In the next convolutional layer and max pool layers, the dimensions become 14x14x32 successfully halving the image size. This is then connected to a fully connected layer and then the output layer where the number of neurons matches the number of classes (156).

The model using which the dataset is trained is a seven layer convolutional neural network which consists of two convolution layers, two max pooling layers, one flatten layer and two dense layers. The image that is to be trained is taken from the Unconstrained Tamil Handwritten Character Database (uTHCD) which consist of 156 classes of unique Tamil characters. Each image is in bitmap format. The bitmap format is converted to a dataset of arrays using data augmentation where the image data is converted to arrays of ones and zeros. The model is trained for ten epochs.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 62, 62, 32)	896
max_pooling2d (MaxPooling2D)	(None, 31, 31, 32)	0
conv2d_1 (Conv2D)	(None, 29, 29, 32)	9248
max_pooling2d_1 (MaxPooling2	(None, 14, 14, 32)	0
flatten (Flatten)	(None, 6272)	0
dense (Dense)	(None, 128)	802944
dense_1 (Dense)	(None, 156)	20124

Total params: 833,212
Trainable params: 833,212
Non-trainable params: 0

Fig. 1 Architecture summary of the proposed model

Activation functions facilitate deep neural networks by introducing non-linearity to the learning process. The non-linearity feature gives the neural network the ability to learn complex patterns. To apply the effectiveness and performance of the proposed RMAF transfer function by comparing with state-of-the art activation functions such as sigmoid, Tanh, ReLU, ELU, SELU, PReLU and Swish. The focus is finding a scalar activation function, which takes in scalar input and output a scalar. This is because scalar activation functions can replace the ReLU function without changing the architectural network.

The Methodology used here is a cost function called relu which is one among many cost functions available used for increasing the performance of our Convolution Neural Network. The RMAF function can be any point utilized when fitting data and giving knowledge to the basic forms dependable on its dynamics

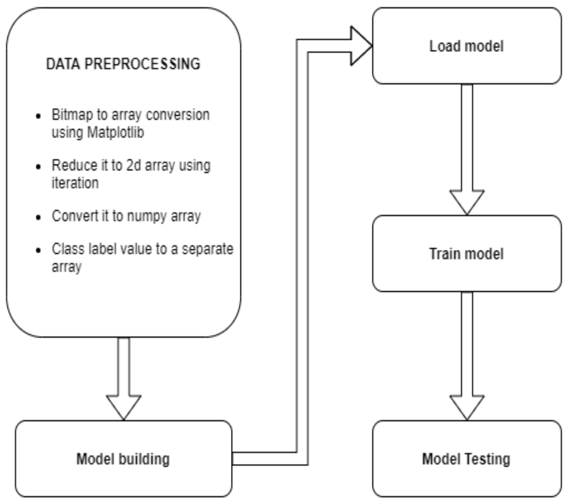

Fig. 2 Architecture of the proposed model.

H. Data visualization

The dataset consists of a collection of bitmap images of 156 classes. Each class represents a unique tamil character. With the requirement of understanding the dataset's meta characteristics, The dataset is subjected to required visualisation. The visualisation involves understanding the frequencies of the classes and general comparison between the classes through a pie chart . The pie chart represents the number of classes present in each frequency, which validates the stability of the dataset.

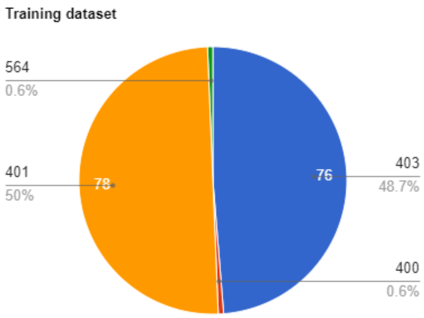

Fig. 3 Pie chart of the varied data frequency.

I. Model training

Model training is a process where we feed our chosen algorithm with the training data and make it learn from the samples. The training model is used to run the input data through the algorithm to correlate the processed output against the sample output. The dataset is splitted to a 70-30 ratio where 70 percent of data is used for training the model.

J. Model testing

Model testing is a process in which a fully trained model is tested and validated on a test set. This is a process to verify the performance of the model using data which is not part of the training dataset.

K. Dataset

This project is based on the unconstrained Tamil Handwritten Character Database (uTHCD) .Unconstrained Tamil Handwritten Character Database (uTHCD) has substantial benefits, mainly from OCR use-case perspectives, Unlike the traditional dataset, the uTHCD database is a unified collection of both offline and online samples, the offline samples capture useful characteristics of scanned handwritten samples such as distortion due to over writing, variable thickness of strokes, and stroke discontinuity and Modern deep learning algorithms for OCR need a considerable amount of samples to develop a robust model . The uHTCD database has 90950 samples with approximately 600 samples each from 156 classes with the provision to enhance more in the coming years.

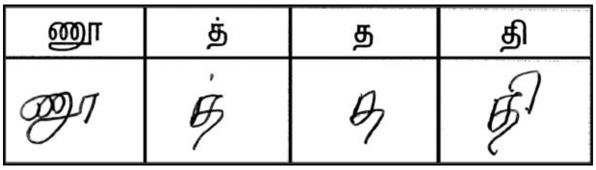

Fig. 4 Sample of the label class and the dataset.

IV. RESULT

The output is displayed as a comparison table having columns Input, Output and Prediction.Input has the name of the class from the test dataset and Output contains the name of the class predicted by the model and the Prediction displays whether the predicted data is true or false.

```
Input   :  Output   :  Prediction
class 000   :   class 000   :   True
class 081   :   class 081   :   True
class 082   :   class 082   :   True
class 083   :   class 083   :   True
class 084   :   class 084   :   True
class 085   :   class 085   :   True
class 086   :   class 086   :   True
class 087   :   class 087   :   True
class 088   :   class 088   :   True
class 089   :   class 089   :   True
class 090   :   class 090   :   True
class 091   :   class 091   :   True
class 092   :   class 092   :   True
class 093   :   class 093   :   True
class 094   :   class 094   :   True
class 095   :   class 095   :   True
class 096   :   class 096   :   True
```

Fig. 4 Test dataset prediction and comparison.

The Number of inputs, prediction of true and false along with the test accuracy is displayed in a separate space using custom code.

```
Total Number of Inputs : 9050
True Predictions : 8744
False Predictions : 306
accuracy : 96.61878453038673
```

Fig. 5 Test accuracy

V. CONCLUSION

In this digital age where everything is digitized , It is an essential need for a character recognition model for Tamil language as Tamil plays an important role on official scale in cases such as government documentation and petition which are mostly still handwritten. Henceforth this project can push some significant extensions of Tamil handwritten character recognition technology.

The future scope is to perform Tamil word detection by customising our single character recognition model. We will be adding additional layers to our architecture and make sure it gives the best accuracy it could provide. It will be more useful to detect words which are tormented, not visible easily and hard to read in Tamil languages. The result of this project will be the best OCR model for word detection for Tamil language which is easily understandable by everyone.

REFERENCES

[1] SYED YASSER ARAFAT AND MUHAMMAD JAVED IQBAL (2020) "Urdu-Text Detection and Recognition in Natural Scene Images Using Deep Learning", IEEE Access (Volume: 8) , 96787 - 96803.

[2] Halefom Tekle Weldegebriel, Han Liu(Member IEEE), Anwar Ul Haq, Emmanuel Bugingo and Def Zhang. (2019) "A New Hybrid Convolutional Neural Network and eXtreme Gradient Boosting Classifier for Recognizing Handwritten Ethiopian Characters", IEEE Access , Vol.8, pp. 17804 - 17818.

[3] Yulia S. Chernyshova; Alexander V. Sheshkus and Vladimir V. Arlazarov (2020) "Two-Step CNN Framework for Text Line Recognition in Camera-Captured Images", IEEE Access , Vol.8, pp. 32587 - 32600.

[4] A. Bharath and Sriganesh Madhvanath, (2011) "HMM-Based Lexicon-Driven and Lexicon-Free Word Recognition for Online Handwritten Indic Scripts",IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 34, Issue: 4), pp. 670 - 682.

[5] WENYUAN XUE, QINGYONG LI AND QIYUAN XUE (2019) "Text Detection and Recognition for Images of Medical Laboratory Reports With a Deep Learning Approach",IEEE Access, Vol.8, pp. 407 – 416.

[6] YONGBIN YU, KWABENA ADU, NYIMA TASHI, PATRICK ANOKYE, XIANGXIANG WANG, AND MIGHTY ABRA AYIDZOE (2020) "RMAF: Relu-Memristor-Like Activation Function for Deep Learning", IEEE Access (Volume: 8), pp. 72727 - 72741.