MALAYALAM HANDWRITTEN CHARACTER RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK

MAIN PROJECT REPORT

Submitted in partial fulfilment of the requirements for the
Award of the Degree of

MASTER OF COMPUTER APPLICATIONS

from

UNIVERSITY OF KERALA

By

JABIR ALI V

Register No: 15400023

DEPARTMENT OF COMPUTER APPLICATIONS

COLLEGE OF ENGINEERING TRIVANDRUM

SEPTEMBER 2018

DEPARTMENT OF COMPUTER APPLICATIONS
COLLEGE OF ENGINEERING
TRIVANDRUM

CERTIFICATE

*This is to certify that the project work entitled "**Malayalam Handwritten Character Recognition using Convolutional Neural Network** " is a bonafide record of the work done by **JABIR ALI V (14400023)**, student of College of Engineering, Trivandrum in partial fulfilment of the requirements for the award of the Degree of Master of Computer Applications from University of Kerala.*

Prof. Jose T Joseph   **Prof. Jose T Joseph**

Project Guide   Head of the Department

External Examiner

ACKNOWLEDGEMENT

If words are considered as symbols of approval and tokens of acknowledgement, then let words play the heralding role of expressing our gratitude.

I take this opportunity to express our sincere gratitude to the people who have been instrumental in the successful completion of our project.

First of all I thank God Almighty for bestowing us with wisdom, courage and perseverance which have helped us to complete our mini project successfully.

I would like to remember with gratitude Prof.Jose T Joseph, Head of the Department, Department of Computer Applications, College of Engineering, Trivandrum for the encouragement and guidance rendered. I would like to express my heartfelt gratitude to my project guide Prof.Jose T Joseph, Assistant Professor, Department of Computer Applications, College of Engineering Trivandrum, for his valuable guidance, support and advices that aided in the successful completion of my project.I also express my gratitude to Dr.Jiji C V, The Principal, College of Engineering, Trivandrum. I would be obliged for the helping hands of Prof. John Prakash Joseph, Prof. Baby Syla and all other staff members of the department.

Finally, I wish to express my sincere gratitude to all our friends, who directly or indirectly contributed in this venture.

JABIR ALI V

ABSTRACT

Optical character recognition is the mechanical or electronic conversion of images of typed, handwritten or printed text into a machine encoded format. Different methods are used in OCR for different languages. The main steps of optical character recognition are pre-processing, segmentation and recognition. Recognizing handwritten text is harder than recognizing printed text. Convolutional Neural Network has shown remarkable improvement in recognizing characters of other languages. But CNN implementation for Malayalam character recognition is done by only a few . The proposed system uses Convolutional neural network to extract features. This is method different from the conventional method that requires handcrafted features that needs to be used for finding features in the text

Convolutional Neural Networks (CNN) is a popular deep learning method and is state of the art for image recognition. CNN has achieved a breakthrough in the IMGENET challenge 2011. The CNN used in the challenge was Alexnet and gave an error rate of 16% in comparison to 25% in 2010. From then on it was CNN all the way. CNN. is very suitable to represent the image structure. The properties of CNN that makes this possible are the local connectivity strategy and the weight sharing strategy. Handwritten character recognition is a difficult task as the characters usually has various appearances according to different writer, writing style and noise. Researchers have been trying to increase the accuracy rate by designing better features, using different classifiers and combination of different classifiers. These attempts however are limited when compared to CNN. CNNs can give better accuracy rates but it has some problems that needs to be addressed. Malayalam characters are complex due to their curved nature and there are characters which are formed by the combination of two characters. These along with the presence of 'chillu' make recognizing Malayalam characters a challenging task. CNN has proved to be the state-of-the-art technique for other languages and hence provides the chance for giving higher accuracy rate for Malayalam characters too.