DOCUMENTATION FOR THAMIZH HANDWRITTEN OPTICAL CHARACTER RECOGNITION (THAMIZH HOCR)

# resolving dependencies

The model and demo was tested on

OS: **Windows 10** and **Mint 18.**

Python Version: **3.6.6.**

**PIP** is needed to install dependencies.

After installing the above, run ***pip install –r requirements.txt*** from the root directory.

In the ***C:\Users\name\.keras*** directory open ***keras.json*** and edit the last line as shown below:

***"backend": "tensorflow"***

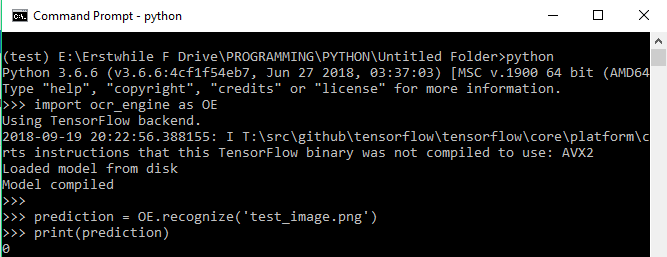
# MODEL FILES

The model directory contains two files ***model.json***and ***model.h5****.* These are output of a CNN model trained on **34,827** images and validated on **2,470** images. The validation accuracy is **91.6%.** These model files can be loaded and used to recognize handwritten tamil characters.

# ocr\_engine

The ***ocr\_engine.py*** loads the models and is ready to recognize. It has a method ***recognize(path)***. The path of the image is to be passed as the parameter to the method.

EXAMPLE:



The output of the recognize method will be an integer from ***0 to 246,*** representing the ***247*** letters of the tamil language. For applications, mappings are to be used to make Unicode representations easier. An example mapping is present in the web\_app\_demo directory: ***label\_map.csv***  which is also used for the demo.

# web\_app\_demo

A web app demo is also provided. Go to ***web\_app\_demo*** directory and type the command

***python app.py .*** A server starts running on the system host: ***0.0.0.0***, port: ***5000*** . If a phone is connected to the laptop through wifi find the ip and enter the address in a chrome browser. ***http://<ip\_address>:5000***

For best results test from a touchscreen enabled smartphone.

If the demo is opened using the browser on the same computer,

Open [***http://localhost:5000***](http://localhost:5000/)

On desktops resize the window to a rough square shape for better results.

Draw a tamil character and hit recognize, the result will be outputted on the right side of the recognize button.

