RUNTIME TERRORS

INGENIOUS HACKATHON

AHMEDABAD UNIVERSITY

IDEATION

- As a background in Data Analysis and Image Processing, we wanted to contribute in the field related to it.
- Initially we thought of solving a agriculture problem but due to lack of depth in the field and dataset, we decided to let go it.

Proposed Problem Statement

- Handwriting Rating System
- Aims at providing a feedback to the user on how to improve his writings using CNN and Computer Vision.

APPROACH

Dataset

- We had a MNIST Dataset that consisted of 0-9 digits and A-Z dataset on kaggle.
- We combined both the datasets and form a combination that contained both digits and alphabets.
- We resized the images and normalized it.

Convolution Neural Network

- We build CNN classifier of our own for recognition of the character using 2 convolution layers and 2 dense layers.

APPROACH

Rating the Hand-Writing

- .Created the contours off a good hand-writing image consisting all the characters and stored those contours.
- Getting the contours of the testing image for rating and comparing them with previously saved contours of Good-handwriting using Structural Similarity Index Metric.
- Creating a threshold based on the basis of detected characters and classifying each character into good and bad using the threhold.

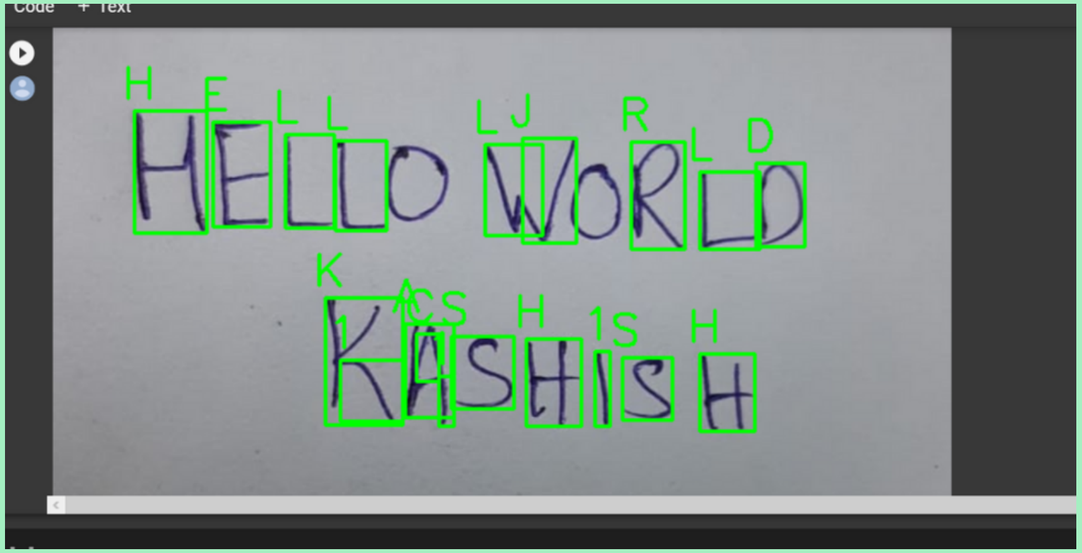
OTHER ATTEMPTS

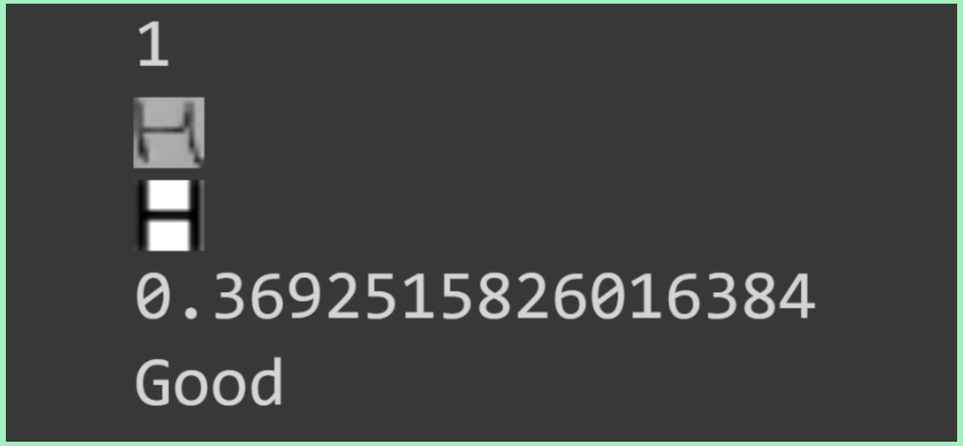
- We created another model and trained it with data of good handwriting characters. After training we predicted the output using both the models i.e the precious model which was trained with mnist and kaggle a-z dataset.
- We will classify the handwritting as good or bad based upon how the output from the previous model is similar to the output of model which is created with the good handwriting characters dataset.
- If a character is badly written then the new model will not recognize it as it is trained with good training set but the old model would have recognized it so the user has to improve writing for that character.
- This approach will generate better results as we are comparing results with the best set. But as we were not having a large dataset of good handwritting characters the model which was generated was not able to generalise well.

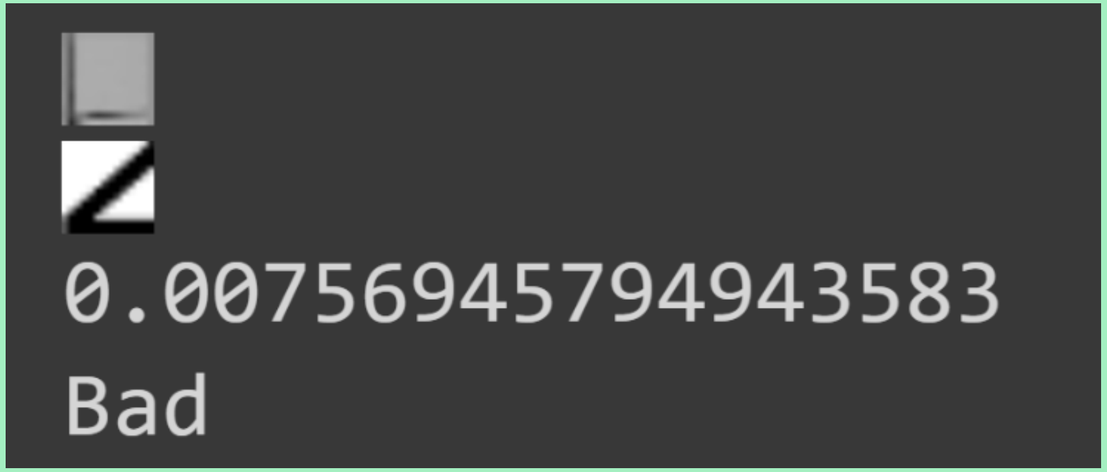
OTHER ATTEMPTS

- We also attempted on grammatical error detection post recognition of the data.
- The algorithm would be as follows:
- First the recognized data would be collected and tokenized into a set of 5 words.
- The tokenized words would be passed on to the google search engine.
- If the words chosen has some grammatical mistake, the google search will give you the "do you know" section which will be basically the corrected one.
- The corrected one would be selected and finally the data would be a kind of error free.

RESULTS


Code + Text
H E L L L J R L D
HELLO WORLD
K A S H H 1S H
KASHISH


1
0.3692515826016384
Good


0.007569457949435 83
Bad

THANK YOU