Optical Character Recognition

Assignment

— Vedha Krishna Yarasuri

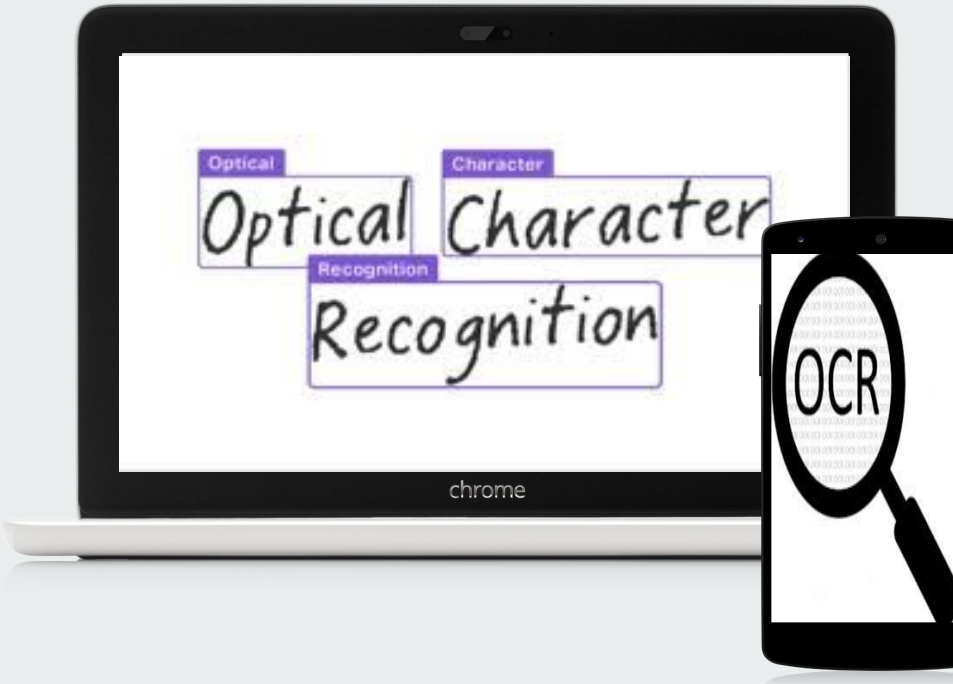
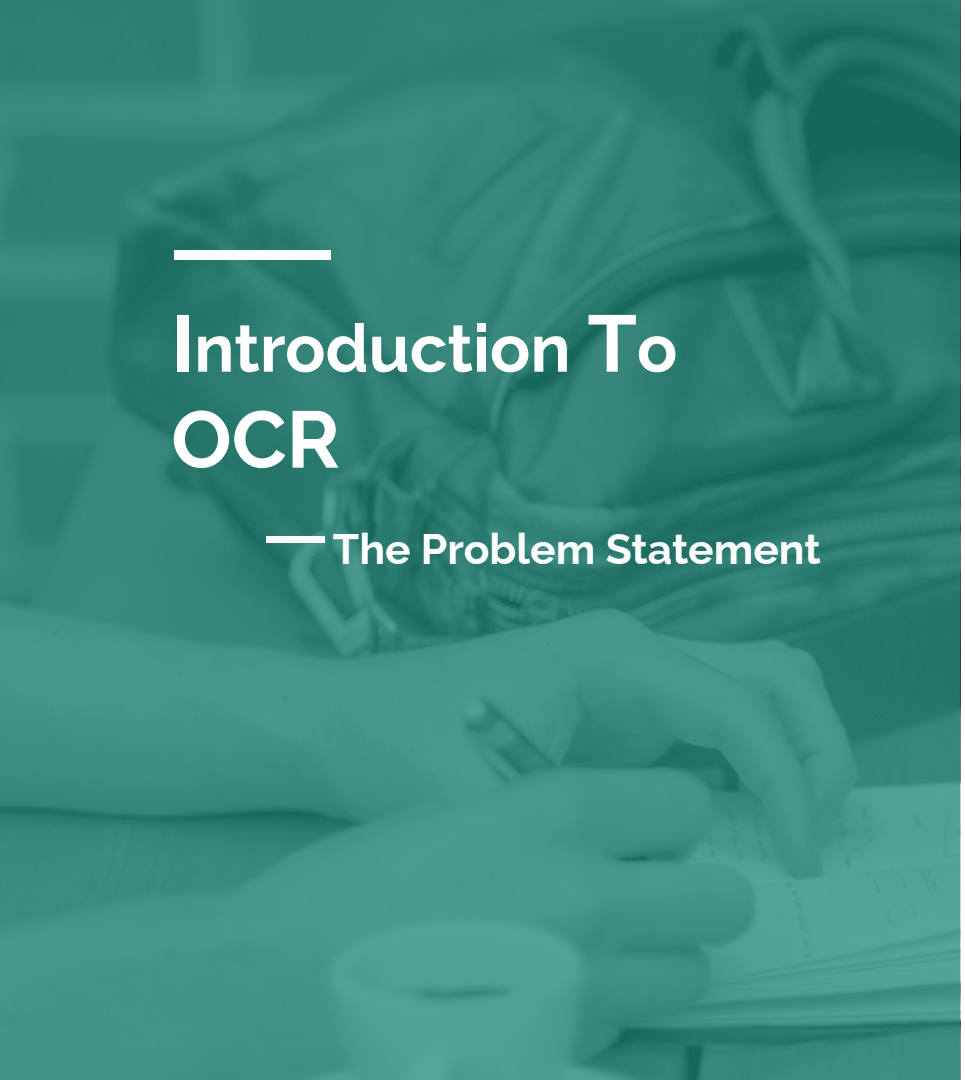


Table Of Contents



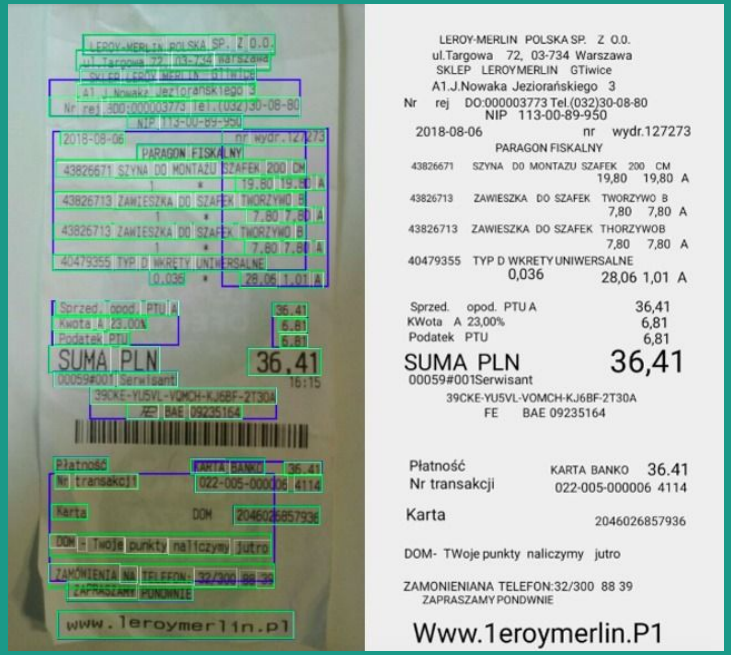
Introduction To OCR

The Problem Statement

Optical Character Recognition

Optical Character Recognition (OCR) systems mainly focus on extracting text from scanned documents.[1] Optical character recognition or optical character reader (OCR) is the electronic or mechanical conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo (for example the text on signs and billboards in a landscape photo)

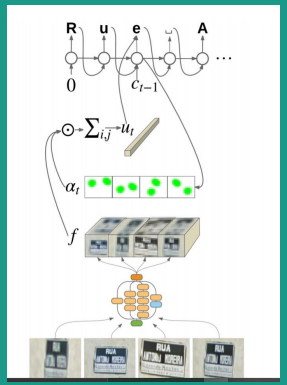
Introduction To OCR



Various Approaches

1 .The First approach is by using Attention Mechanism

Wojna Z et al has proposed a model with 84.2% accuracy for recognition of textual data from images using French street dataset. They have trained a model which takes various versions of an image and plots attention masks on the image where the text is available and predicts text in the image.



2 .The Second method for extracting text from an image is by using Connectionist temporal classification (CTC).

It can be used for tasks like on-line handwriting recognition or recognizing phonemes in speech audio. CTC refers to the outputs and scoring, and is independent of the underlying neural network structure.

3 .The Third method is by using image processing techniques

My Approach

The Attention Based Approach

Model Architecture

Feature Extraction The features of the given image are extracted using pre-Trained Inception Resnet V2 model. The features obtained are then stored using numpy.

Sequence Model

The features are given to an encoder (conv net) and then passed through attention model (neural net) to compute attention weights (context). These attention weights are sent to a decoder to decode the text embedded in the image.

The Dataset

IIIT Hyderabad's Dataset

The IIIT 5K dataset contains 5000 cropped word images from Scene Texts and born-digital images.

Few Instances of the images are shown in the following slide.

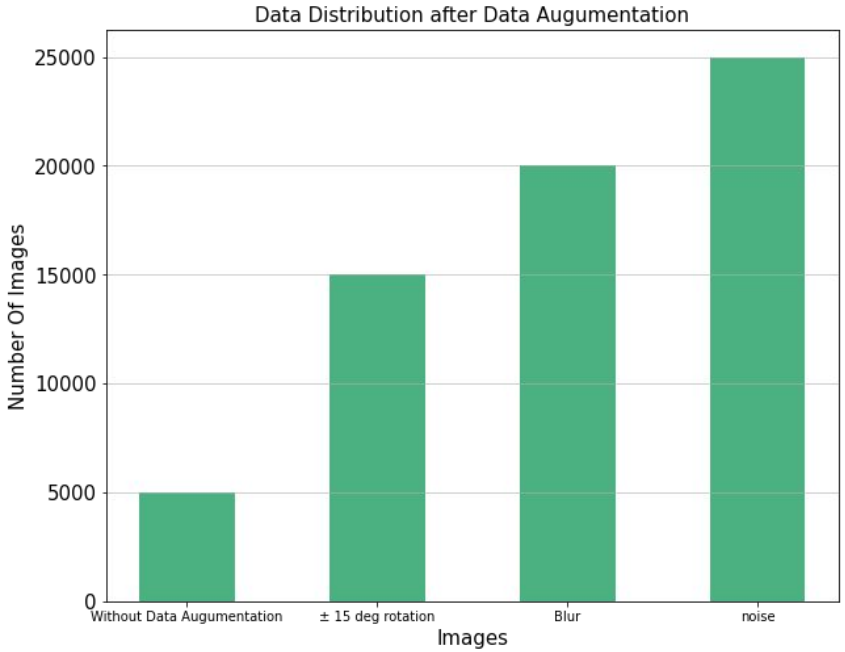
The Dataset


save
the
WWW
Internet
India
think)
State
D's
all
tea
Tues
STATE
RIVER
State
UFO
be
on
NOKIA
28
Bank
iPad
Air
Change
Loans
THAN
for
Zero-Emissions
A/C
WRATH
DELIVERY
of
TRUTANICH
AUGUST
Singapore

Data Augmentation

Data Augmentation Techniques :

1 . Rotating images ± 15 deg
2 . Image Blur using PIL
3 . Uniform Noise to the Images



After Data Augmentation

Column 1 : Normal Image

Column 2 : + 15 Deg Rotated Image

Column 3 : - 15 Deg Rotated Image

Column 4 : Image Blur

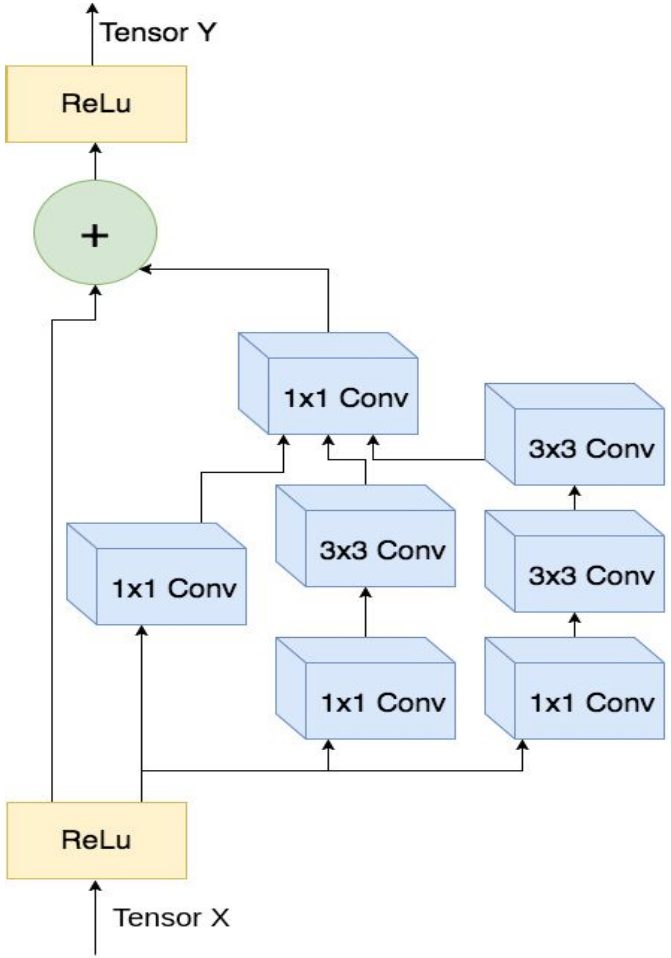
Column 5 : Image Noise

Feature Extractor

– Inception Resnet V2

• The inputs of the dataset are given to Inception Resnet model for extracting features at the layer “**block8_1_conv**” .

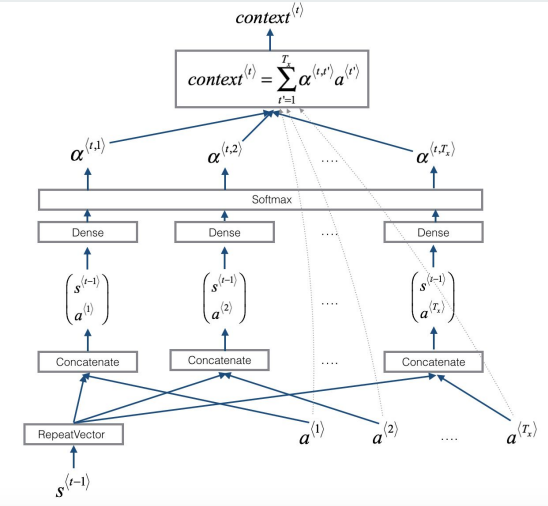
• The Layer “**mixed_7a**” and “**block8_1_conv**” nearly gives similar results.

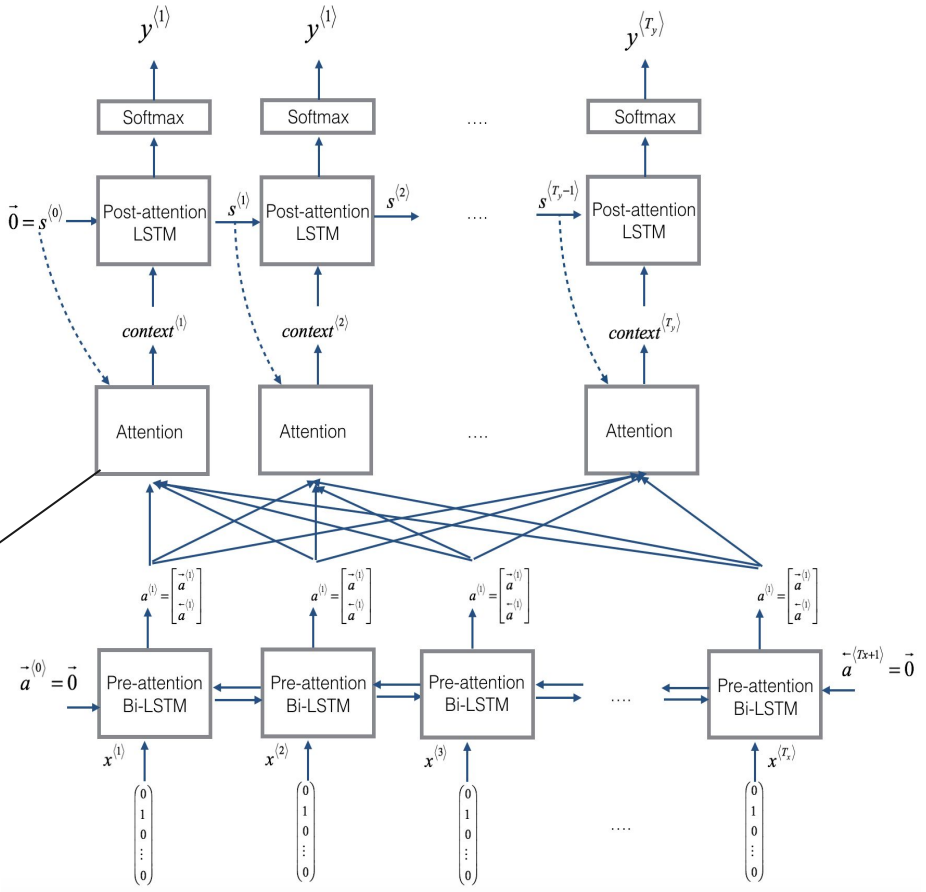


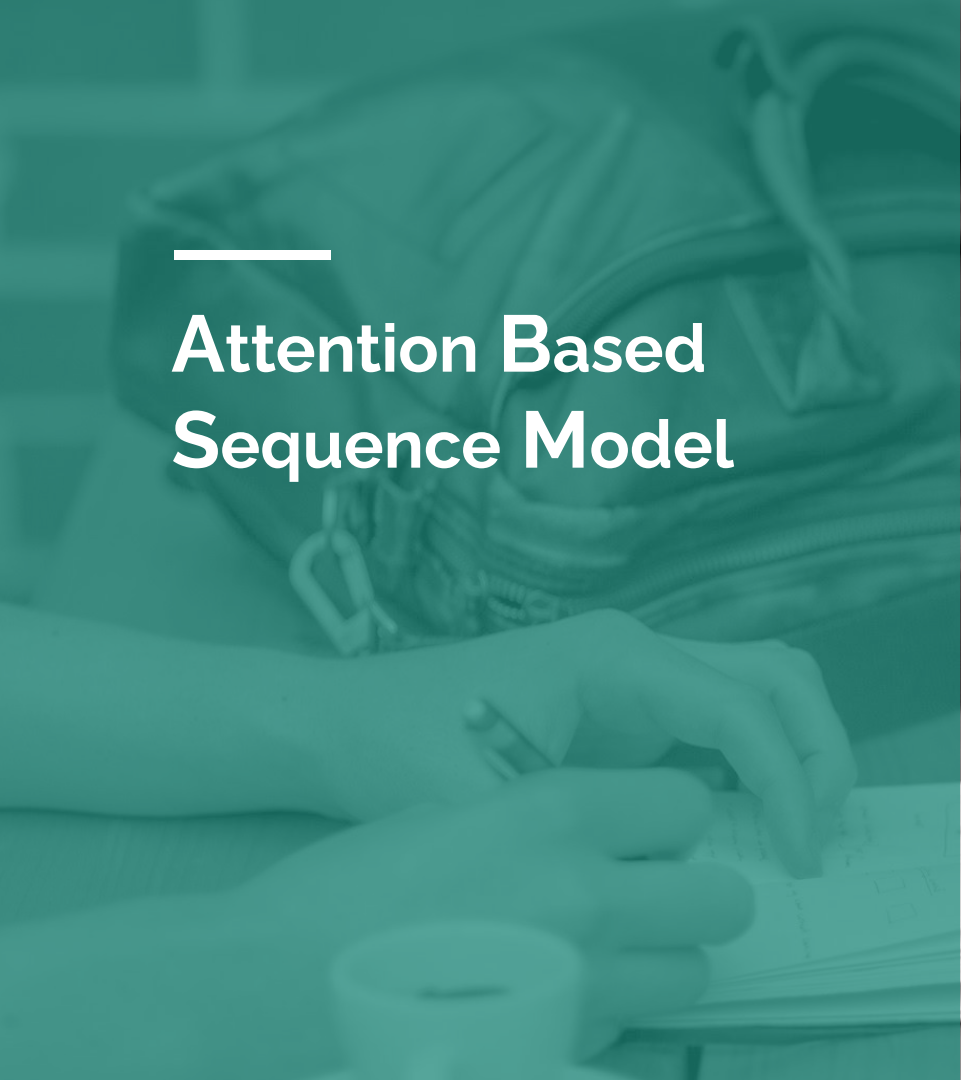
Features Extracted From “block8_1_conv”

Features Extracted from the image labelled “ **YOU** “

Attention Based Sequence Models







Attention Based Sequence Model

Model Architecture

Encoder

A Shallow Neural Network with 256 units, outputs the activations of the encoder.

Attention

A Deep neural network with 1 hidden layer sandwiched between input and output layers. Input and hidden layer contains 512 units.

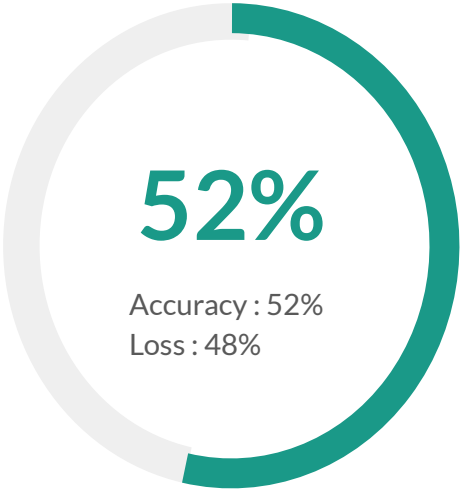
Decoder

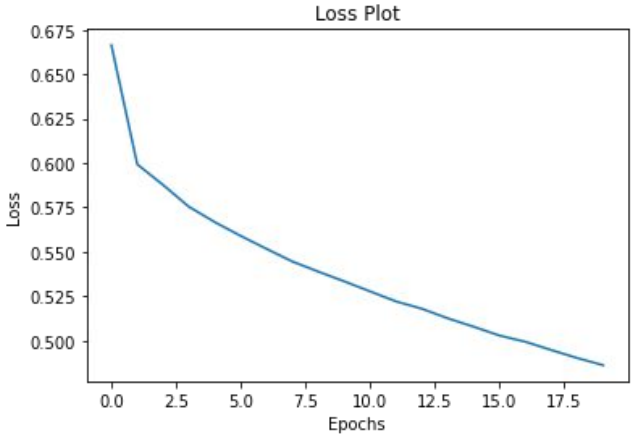
Contains an embedding layer, GRU layer and followed by 2 deep layers. GRU and 1st deep layer contains 512 units. The 2nd deep layer contains 41 (Vocab size) units.

Hyperparameters

Hyperparameter Name	Value
BATCH_SIZE	64
BUFFER_SIZE	1000
embedding_dim	256
units	512
vocab_size	41
num_steps	609
features_shape	2080
attention_features_shape	64

Training, Checkpoints, Saving The Weights

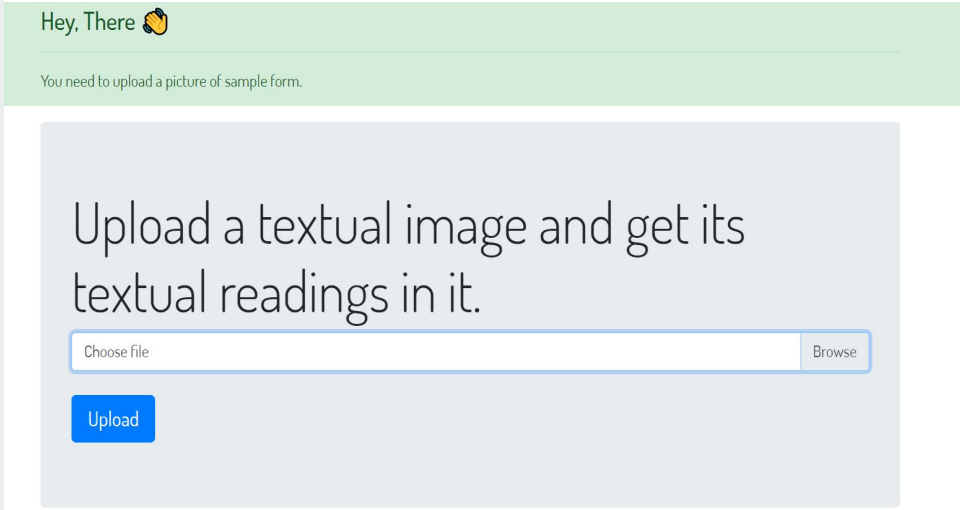




Web App Design

Web Application has been designed using :

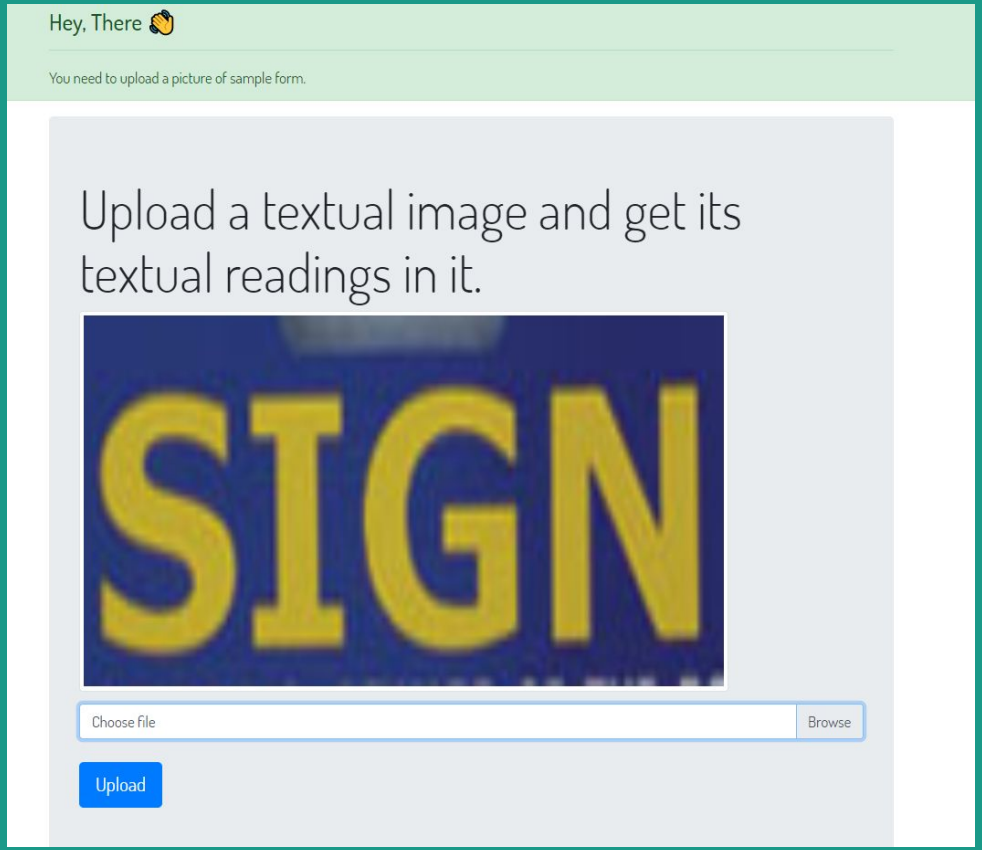
1 . Node.js
2 . Bootstrap



Working

When the user uploads then the image will be sent to the server using REST api and gets downloaded using multer.

Later The steps remains same.



Challenges Faced and Solutions

My Challenges and Solutions

1. Lack of Experience and Practice
   -- Time Management and Hard working
2. Designing Sequence Models
   -- TF Image Caption Example
3. Tensorflow.js -- Learning
4. Shell Scripting -- Learning

Thank You