Domain Background

India being the most diversified country and got around more than 22 official languages, people face a problem in identifying the sign board Written in various languages. Local languages which people face problems in understanding the content written in the sign board.

Problem Statement

To build the model which would perform below tasks:

- An object detection engine to detect text in the signboard in the given image

- A CNN-RNN encoder decoder model to convert the cropped signboard into text (a sequence of characters)

- A sequence-to-sequence with attention model for transliterating the above sequence of characters to the desired language.

Datasets and Input

The datasets for each of these tasks is available at the below mentioned URLs:

- Bounding boxes detection which contains text in the images

Train Set: https://drive.google.com/open?id=1E5kI8CLoC-XffqQMTWwSpBIPp1Wb2tne

Test Set: https://drive.google.com/open?id=1Z6Qxr-q-F54iYB2G1AyoDymBh64f5REZ (428 real images with annotations)

- Identifying the text from the cropped section

Train Set: https://drive.google.com/open?id=1E5kI8CLoC-XffqQMTWwSpBIPp1Wb2tne

Test Set: https://drive.google.com/open?id=1C0-mc0WAIdssS5KJwOjghaWaqiImZZUr (1740 cropped word images from real pictures with annotations)

- Native Indian Language text to English

Link:
https://github.com/GokulNC/NLP-Exercises/tree/master/Transliteration-Indian-Languages/Original-NEWS2012-data

Solution Statement

Here we are building the object detection model which recognizes the text area in the given image of the sign board and this particular part of the image is cropped and passed into CNN-RNN encoder and decoder model which converts cropped image into text and here we will achieve in getting the sequence of characters, later this sequence of characters is passed into attention model for translating the above sequence of characters to the desired language.

Evaluation Metric

I will be considering following things as evaluation metric:

- Confidence score
- Intersection over Union, etc

Detection will be considered as true positive or false positive based on above mentioned techniques.

Project design

Step 1: To classify images which contain text/no text

Step 2: Bounding box location

Step 3: Extracting text from the bounding box

Step 4: Recognize extracted text

Step 5: Translate step 4 text into english language