DESIGN CREDIT INDICPHOTOOCR

Evaluation report

By -
Jivitesh M S (b24me1039)
Prem Vikram Mokal (b24bb1028)

DATA COLLECTION

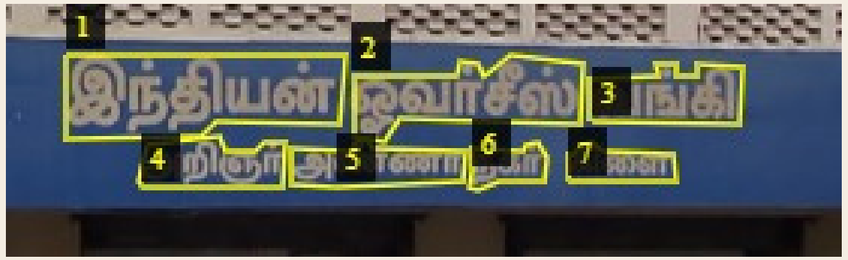
Tamil

Marathi

DATA ANNOTATION

The images were annotated using VGG image annotator and the data was stored in a JSON object in the format specified.

Tamil



Marathi

MODEL LOADING

We loaded the IndicPhotoOCR model using Anaconda. Then we used the model to recognize the text in the images. We stored the output in a JSON file in the following format for testing.

```
"10.png": {
    "இந்தியன்": "tamil",
    "ஓவர்சீஸ்": "tamil",
    "வங்கி": "tamil",
    "அறிஞர்": "tamil",
    "நகர்": "tamil",
    "கிளை": "tamil",
    "அண்ணா": "tamil"
},
```

```
"img1.jpg": {
    "शुभारंभ": "marathi",
    "होलसेल": "marathi",
    "डेपो": "marathi",
    "साडी": "hindi",
    "ॲन्ड": "marathi",
    "रेडिमेड": "marathi"
},
```

MODEL TESTING

To test the model on our dataset, we defined metrics like word match accuracy and language match accuracy.

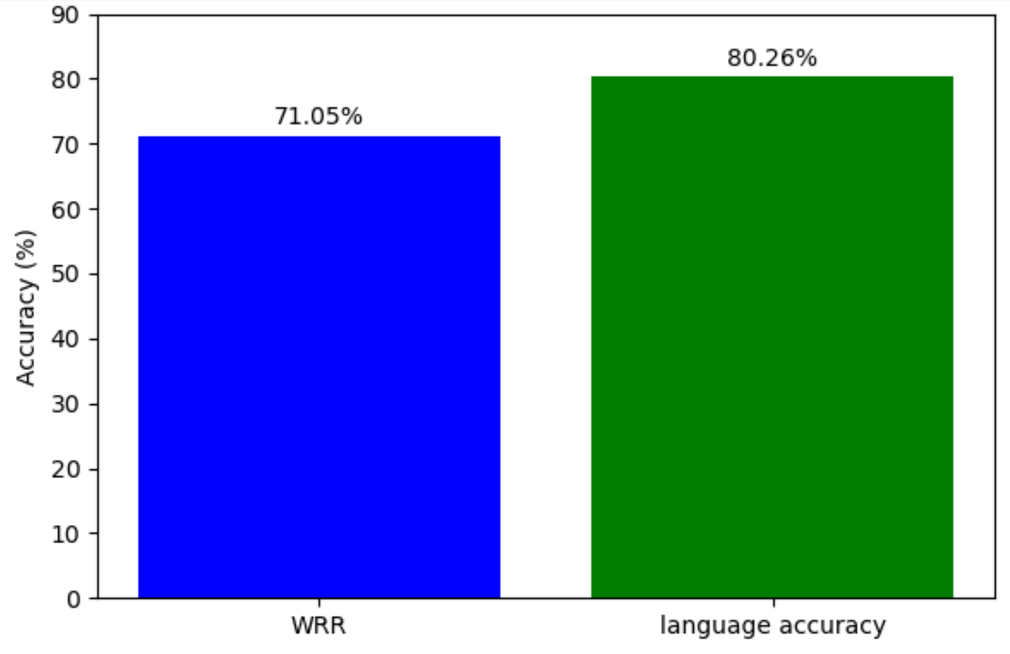
$$\text{word match accuracy} = \frac{\text{words recognized correctly by model}}{\text{no. of words detected by model}} \times 100$$

$$\text{language match accuracy} = \frac{\text{no.of times model detected language correctly}}{\text{no. of words detected by model}} \times 100$$

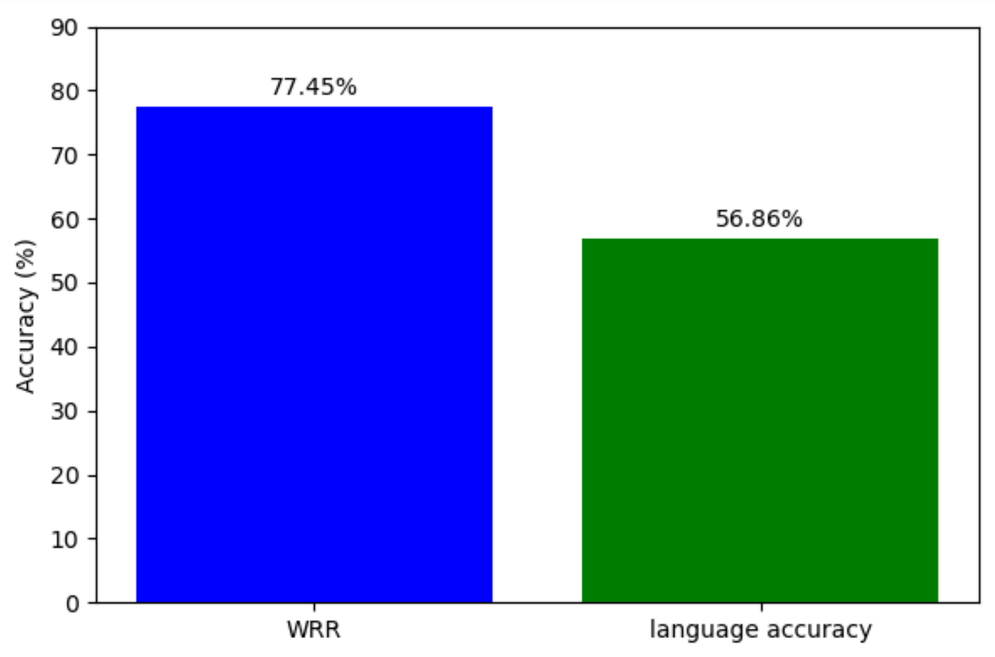
TEST RESULTS

We obtained the following results by testing the model on the metrics.

Tamil



Marathi



THANK YOU