**Sign Board translation from Vernacular languages**

**TEAM MEMBERS:**

**Sruthi Srinivasan-19BAI1046**

**A.Prasad-19BAI1061**

**Vishnu Prasad - 19BAI1127**

**Abstract**

India is a very diverse country as a result of which numerous vernacular languages are spoken around the country. For an interstate traveller it would be highly tedious to understand and decipher every language and read various sign boards written in different languages along his journey.

Traditionally, people would use several translation tools in order to overcome this language barrier. However, this would prove to be inefficient since text in different languages often cannot be translated directly. In order to facilitate translation of written text we chose to build this application which can directly translate text written in several indian vernacular languages into any other language of the user’s choice merely by the input of an image of the signboard clicked by the user. Subsequently, the user can obtain the translated text as an audio message as well.

When an input image is fed, the text in the image is firstly detected, then recognised and finally extracted from the image. This type of text recognition can be done using OCR or optical character recognition. In our project, the EasyOCR software is used for the purpose of multilingual text recognition. The text can be translated into any language of the user’s choice. This is carried out using the google translate module in python and finally the audio conversion is done using a module in python named pyttsx3 used for text to speech conversion.

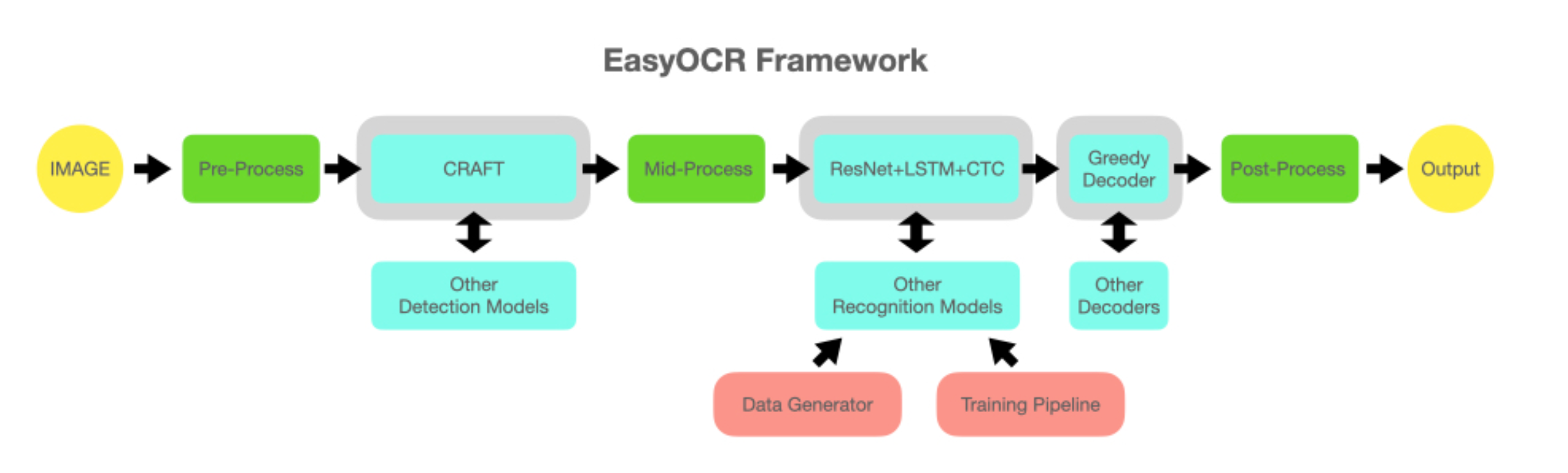
The dataset in consideration consists of images of several words in multiple languages and these are trained to form the EasyOCR module.

**Introduction**

Text detection is the premise and guarantee of text recognition. Multi-oriented text detection is the current research hotspot. Due to the variability in size, spatial layout, color and the arrangement direction of natural scene text, natural scene text detection is still very challenging. In order to carry out text detection and recognition we are using the EasyOCR module.

This module was built by JaidedAI for the purpose of multilingual text detection and recognition. Most text detection and recognition techniques are applicable only to english. EasyOCR however is an OCR that performs text detection and recognition for several indian vernacular languages which is the requirement in this project. This is done by incorporating all the characters present in that language supported by nearly 30000 to almost 50000 words if the language is popular.

EasyOCR uses a generative adversarial network(GAN) to generate a realistic handwritten dataset to improve handwritten text recognition.



EasyOCR uses the CRAFT text detector for text detection. Text recognition consists of ResNet and LSTM with Connectionist temporal classification(CTC) for neural network output and scoring.

**Related Works**

1. This work[1] presents an image operator that seeks to find the value of stroke width for each image pixel, and demonstrate its use on the task of text detection in natural images. The suggested operator is local and data dependent, which makes it fast and robust enough to eliminate the need for multi-scale computation or scanning windows. Extensive testing shows that the suggested scheme outperforms the latest published algorithms. Its simplicity allows the algorithm to detect texts in many fonts and languages.
2. This work[2] proposes a simple and fast multi-oriented text detection method. This method first optimizes the regression branch by designing a diagonal adjustment factor to make the position regression more accurate, Secondly, this method adds an attention module to the model, which improves the accuracy of detecting small text regions.Then, it introduces DR Loss to solve the problem of positive and negative sample imbalance.
3. This method[3] proposes an attention mechanism which roughly identifies text regions via an automatically learned attentional map. This substantially suppresses background interference in the convolutional features, which is the key to producing accurate inference of words, particularly at extremely small sizes.
4. This method[4] proposes an end-to-end method for scene text localization and recognition. It introduces features such as a departure from a strict feed-forward pipeline that is replaced by a hypotheses-verification framework simultaneously processing multiple text line hypotheses, the use of synthetic fonts to train the algorithm eliminating the need for time-consuming acquisition and labeling of real-world training data and the use of MSERs which provides robustness to geometric and illumination conditions.

**Methodologies:**

The user needs to upload the image to the software from their device. After uploading the image the user needs to select the language present in the image and the language that the user desires to be translated. Once choosing the languages the user needs to press the Process button.

Once the process button is pressed the image chosen by the user is sent to the EasyOCR function.

The EasyOCR model initially converts the images to gray scale and then to binary images made up of only 2 colors (black and white). Once the images are pre-processed the image is sent to the CRAFT model(Character Region Awareness For Text detection), here is where the model will find the bounding box of where the text is present. Once the bounding box is formed, the image is cropped based on the coordinates of the bounding box.

This OCR uses the pattern recognition approach which works by identifying the character as a whole we can identify a line of text by looking for rows of white pixels with rows of black pixels in between in the same way we can identify what an individual character begins and ends next we convert the image of the character into a binary matrix or white pixels or zeros and black pixels or ones then by using the distance formula we can find the distance from the centre of the matrix to the farthest one we then create a circle of that radius and split it up into more granular sections at this point the algorithm will compare every single subsection and can send database of matrices representing characters with various fonts to find a character it statistically has the most in common with doing this for every line in every character makes it easy to bring printed media into the digital world.

Once the text from the image is recognised the image is sent to the Google Translator which translates the given text to the desired language specified by the user and this text is displayed on the screen along with the speech output using Pyttsx3.

**Experiments and results:**

|  |  |  |
| --- | --- | --- |
| **ASSAMESE** | Pytessseract | EasyOCR |
|  | নি হে মম ৰা  হ্ণ্ঞ ১) "৮. |  ু (১ \_ঢ  ₹) ৰ ন,  ৰ , ৰ ৰ  ট্|  ন ঢাপ;  [যা ৰ | সহায় কবক  {0হ:ো৫:]  5=হফূ=5  ^৬ঔঋ৯৪\_&^7 ও০ৃ৯হড়)ন ৪হ)ঊঋ  ভং5হঋ{ঞ৩  }৩=৫ ৪৯খ&  প্রতীক  মযদাব  গমাব  সগণত  গঞ৫চঢ৩ট{] |
|  | কি ৷ }};  লম ব্ত্া ৰ  সি} ভক চু  ১ ৰ- তত্‌ স্মি  ৷ ভাৰ] | ভাহাশহদসেশন শিনচ৭  মাসা হামীব ন হিলব  3||/3|5//4[{[ 50//{|!;|[|[? |
|  | | ৰ ন  ৮৬৮ ,  | '}]) ই;  [ত ৰ সু জে | আদবিছোঁ  খ্াকপেরতূয় মগব  ৮৬  দবঃ |
|  | ঢ় | গছ  মানুহ  থাকিলেহে  থাকিব |

|  |  |  |
| --- | --- | --- |
| **MARATHI** | Pytesseract | EasyOCR |
|  | द घ्  वभ र  र्न न."  ना  -  ) 7  ठर गं  । नि  11 आ, नि. डर झट  [टा  | वि टी: | महानगर पालिका भवन  डेक्न  शनिवार  बाडा  शिवाजीनगर |
|  | आ  न  टा | पॉडर्नटीडेपो  ३त३५ 5टँ [ँ2ऐ२ |
|  | का ् पा  8: 8) । | )  प.  व 841 | माऊल |
|  | ३... स.  >. चस  ळा र  न ळे!  क) र | साइन्स  ईल  मरीत लाईन्स  "{ड  २/६6 |

|  |  |  |
| --- | --- | --- |
| **ENGLISH** | PyTesseract | EasyOCR |
|  | 13103 [h ]  2,01 . | EMERGENCY  EXIT |
|  | o @ =  Oakland |[Berkeley  enevrl |\ SIS\ 4  3 ’.‘ 'g o) | MEST  24)  13)  Wa  Oakland  Berkeley  LEFT EYIT |
|  | I ‘/AIE  o1 8 | Plecse be  mindful  cf  distoncing  soaidl |
|  | 1  L {[TR N S  B =N AU  b 1IN A  Tt M, m"?’!‘!!m e L - ‘ \_::  Ll WARNING! G  \_ e p—  S / M \_ Do not proceed until "  3 . ® bariers fully raised.  N , No Talgating ————  // \ . | WARNING!  STOP  Do not proceed until  barrier is  raised.  No  0  fully  Tailgating  SEGR  eHL{  8JS655[  "0#5 |

|  |  |  |
| --- | --- | --- |
| **Tamil** | pytesseract | Easy OCR |
|  | துப்ப  டூவண்டஈம்    30.00 (=…    10.00 ௦… | துப்ப  &  8  வேண்டாம்  30.00 ணஸ |
|  | \_... 3060(12ாள) க \_  வு) ம  லவா 911011 | ஞ  ச2ூம ம  30 ஸ 12 3றந^)  ஏ5=ௌ}ொ  அனுமதி இல்லை  $:9"=ௌ} |
|  | … குப்னப  … டூபஈடஈடூத    மொ பாவோகௌ | குப்பை  போடாதே  வ0 \_:{1}3^9 |
|  | தியப ம வத பயன்ப ப்ப  லவ்வ அ ஆட.  பஹ்ல் க வல்லவ வவததல்‌ ப  த்ய க க த ன்‌  க ல்லை  ன ணை ந்‌  ன்‌ ணட. ்‌ 7  ம்‌ ்‌ ன ளா! | நெல்லை  மன|ை  ஈ4 |

|  |  |  |
| --- | --- | --- |
| **Urdu** | pytesseract | Easy OCR |
|  | ‘، ۱١۱ 1٠۱ | |[!  کيس/ۓراۓراهىکژرشـغ  ې كوتٴمحـےې | کیٹکےسامنے گاڑی کڈُی زنا  خت منع ے  پالنگت مقلقکای اداخظهعے |
|  | او 23 ۴ت | تاج آفش |
|  | سف  - ہے 7ں  سے ‎٤‏  ‏۳ 7 یڈ | میہ لوائنٹ |
|  | 1  ھنہڈ)  ۱ عم ِ لق  ا  چپ ۱ و  ۱  . ۶ ےکک  - 8× وچ ےر ہے۔- و رےھل'  0 ےج ہے سے جم  5 سے-- 2-7  1 1ا ا ا ا ا ا ا ا 7 ار ا زا  وی 2 نے و اے : کے چ..> ۰ | سکندرآباد  ل]ا  @\*عة |

|  |  |  |
| --- | --- | --- |
| **Bengali** | Pytesseract | Easy OCR |
|  | ক  সা দূর | করোনাভাইরাস  সামাজিক দূরত্ব |
|  | নু ৬.গগনা.\*১// 4... / .ন্‌ ১৬৮৬২ নে  নর | গগনা  /%,)2]'%<]5^][8 |
|  | ৮6145517106 51. 5.2...জজ যা সা -.সে [নে ৬০ ০ আন  আচে শারযার নীট[সস আর ক মত তি ।মে মি তে | @চচঙঞঋ|চ 31: !@  [চেশায়ার ঠট |
|  | বর্ণপরিচখুঁ | বর্ণপরিচয় |

|  |  |  |
| --- | --- | --- |
| **Hindi** | Pytesseract | EasyOCR |
|  | न ह ही की ४ कै कल्व्यी  | ड - ज क्र  1  ः « हि,  दी  ! ष्थ्ि  ७ है को [0  जी ह | रूकिये |
|  | कर दिस का, 2, 2 “सु क की ९»...  भ्शट | ल हि जि | है हे न  के 2८ अर  हे. की न ४4  अं, लि... |  बाल हि के कि धाा 00९0९ पर. ७ उख  न जा ॥॥||||| त लय घद , [जिला  जा 0 है फककन्नननत  - उब| हे  ) 4 ही:  रे न |  भा के हि दि जिस | नरसिंहपुर  क़"$|ढ:लणर |
|  | 'फोटो सींचना  मनाहै।  शिण 0५९१४  15 शश०ााए) | फोटो खींचना  मना हे|  क़=07कक4शख}  $ शब0=|8च7६२ |
|  | | बटचलर)  \_ . | बैट्री चार्जर  रूम |

|  |  |  |
| --- | --- | --- |
| **Telugu** | Pytesseract | EasyOCR |
|  | క గ "గా న  స వక క్‌ న  ? 1 | తా | గు  టుపన్లి |
|  | కే లక్ష్మీ విలాస్‌ బ్యాంక్‌  క ద.  1...  ట్ట - మాడ్రా క ౦  మా... జడ | లక్షీవిలాస్  పంబగుట వాధ  ఘ^కోృగఇ ఎ14$ ట4ఊక  లక  బ్యాంక్ |
|  | గ్గ ౫  గ క... క్‌  తే 93106 శ  "౬ ఉంగుతిశ0య అయ  "౫ ల | టచ@0డథచమ $ొ0  సపూ3జశ9ంఎ 1ంఊవ  0ల-న్నకన  ప3న9పఎడ\* |
|  | 0121212211 0210003  1-౨ పన  ము య... |  చం... | 5డ 'ప`ఎ:"  ద-చ] 05"  ఫలకొనుమా రైల్వే స్టేషన్  రఇఇాన] #శఇ  ~^ఎలగ<^ న>ల. $+శ+చంన  29థ |

|  |  |  |
| --- | --- | --- |
| **Kannada** | Pytesseract | EasyOCR |
|  | ಜ. ಶಿ | ಎ ಬ ಟಕ  138.2 | % " |(ಫ  " ೀ"6೦ ಚ ಚ^ )  4  '(1 51  ಐಾಂಟಲ  [ಥ.ಪ] ೨:೪ಷ[ಶ  [ತಚ:^೯  "( )  ಥ್ಞಾಗತ್  [ಸಸೈಾಪ್ |
|  | ತ ದ್ದ.  ''ಷತಾಷಾಗ ಷ್ಟ) | ಎತ  ಉಆಯ  ೨"ದ  19  ೪19]] |
|  | ಸ್‌ ಸ್ರ 1  ೫ 11.1.1...  ಹ ಕ (1!  ಇ ಬೂ  ೫ ೫ 8.126 | ಕರ್ನಾಟಕ ಉಚ್ಯ ನ್ಯಾಯಾಲಯ  }3ಾ (0=೪1 05 ^4೪ೈ^[4^^  ~6452.53 |
|  | . ' ನಂಜನಗೂಡು 3  ಅ  ೫.1 8)876ರ | ನಂಜನಗೂಡು  ]ಸಮ್ಾೊತ  3ಇ್ಊಇ/9-]  ಇ.ಞಖ 5[ಚ- 06-.೦ |

**Comparative study:**

Accuracy calculation:

Single image accuracy=Correctly Predicted Characters/Total no of actual characters

Accuracy of the OCR for a language is = sum of all single image accuracies/total images

EasyOCR vs Pytesseract:

|  |  |  |  |
| --- | --- | --- | --- |
| Language | Pytesseract | EasyOCR | Average |
| Marathi | 3/26,0/6,0/3,1/7 | 24/26,5/6,2/3,6/7 | 6.45%,82% |
| English | 0/13,15/29,3/33,52/65 | 13/13,27/29,27/33,  64/65 | 35.2%,93.34% |
| Assamese | 4/24,5/26,3/18,0/14 | 9/24,12/26,14/18,14/14 | 13.14%,65.35% |
| Tamil | 7/9,1/8,6/11,3/5 | 9/9,8/8,11/11,5/5 | 51.2%,100% |
| Urdu | 10/30,0/4,3/7,2/6 | 19/30,4/4,5/7,6/6 | 27.38%,83.69% |
| Bengali | 3/15,3/3,2/7,5/6 | 15/15,3/3,6/7,6/6 | 57.97%,96.42% |
| Hindi | 1/3,2/6,7/8,2/7 | 3/3,6/6,7/8,7/7 | 45.68%,96.87% |
| Telugu | 2/5,9/9,7/20,2/10 | 4/5,9/9,9/20,10/10 | 48.75%,81.25% |
| Kannada | 3/7,2/6,3/11,7/7 | 2/7,3/6,11/11,7/7 | 50.86%,69.64% |

**References:**

1. Epshtein, B., Ofek, E., & Wexler, Y. (2010, June). Detecting text in natural scenes with stroke width transform. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 2963-2970). IEEE.
2. He, P., Huang, W., He, T., Zhu, Q., Qiao, Y., & Li, X. (2017). Single shot text detector with regional attention. In *Proceedings of the IEEE international conference on computer vision* (pp. 3047-3055).
3. Cao, Y., Ma, S., & Pan, H. (2020). FDTA: Fully Convolutional Scene Text Detection With Text Attention. *IEEE Access*, *8*, 155441-155449.
4. Neumann, L., & Matas, J. (2010, November). A method for text localization and recognition in real-world images. In *Asian conference on computer vision* (pp. 770-783). Springer, Berlin, Heidelberg.
5. Mathew, M., Jain, M., & Jawahar, C. V. (2017, November). Benchmarking scene text recognition in devanagari, telugu and malayalam. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)* (Vol. 7, pp. 42-46). IEEE.
6. Ye, Q., & Doermann, D. (2014). Text detection and recognition in imagery: A survey. *IEEE transactions on pattern analysis and machine intelligence*, *37*(7), 1480-1500.
7. <https://ai4bharat.org/articles/sign-board>
8. <https://github.com/JaidedAI/EasyOCR>
9. https://pypi.org/project/pyttsx3/