Improved Real time Signboard Detection and Translation Using FRCNN

Dona Lal[1], Jothika Mohan[2], RenjishaT S[3], Saviya Davis[4], Divya Mohan[5]
[1]Department of Computer Science and Engineering, Albertian Institute of Science and Technology, Kalamassery, Kerala, India, donalal28@gmail.com
[2]Department of Computer Science and Engineering, Albertian Institute of Science and Technology, Kalamassery, Kerala, India, jothikamohan98@gmail.com
[3]Department of Computer Science and Engineering, Albertian Institute of Science and Technology, Kalamassery, Kerala, India, renjishasivaraman@gmail.com
[4]Department of Computer Science and Engineering, Albertian Institute of Science and Technology, Kalamassery, Kerala, India, saviyadavis03@gmail.com
[5]Assistant Professor atDepartment of Computer Science and Engineering, Albertian Institute of Science and Technology,Kalamassery, Kerala, India, divyamohan@aisat.ac.in

ABSTRACT

In India, majority of people are speaking the language Hindi, but a major portion of signboards are in English. On a business or pleasure trip, the travelers get confused by the various sign boards written in English. As smartphones becomes most popular in recent years, they can rely on smartphone for the same. This paper explains the work intended to build a mobile application that can recognize the English content and sign present on the signboard image, detect and translate the content and symbols from English to Hindi and display the translated Hindi text back to the screen of the phone. The system uses pre-trained faster regional convolutional neural network using pre-trained CNN for object detection, tesseract OCR for text extraction and English-to-Hindi dictionary for translation.

Key words: Convolutional Neural Network(CNN), Deep Learning(DL), Faster Regional Convolutional Neural Network(FRCNN),Optical Character Recognition(OCR),Regional Proposal Network(RPN).

1. INTRODUCTION

Nowadays, people have more opportunities to travel across various parts of India. Language barrier is a major problem. It is also important to understand the information on signboards. The travelers face difficulty in understanding the signboard which are written in English. This often spoils the experience of visiting a new place and the travelers goes back with not so-fond memories. So, it is necessary to have a mobile-applications for detecting signs and translating simple texts on signboards. We think of mobile phones first, since it become most popular and people carry them almost everywhere they go. The gaining momentum of portable mobile devices and the inclusion of cameras on them makes a way for faster and comfortable translation.

With the arrival of deep learning, the machine learning systems are able to recognize and classify the objects of interest in an image. Deep learning is a form of machine learning where the computer is able to identify and correct its own mistakes. When compared to other algorithms, deep learning algorithms uses neural networks to support the computer to use logic to draw conclusions in a way similar to human brain. Therefore, the issues with traditional method can be overcome by using deep learning algorithms as it is far more capable form of machine intelligence.

The main purpose of this project is sign detection, English text extraction from signboard image and translating the text to Hindi. When compared to traditional optical character recognition, text extraction from natural scene images is more challenging because of complex backgrounds and light conditions. Therefore, text detection and extraction is an essential step in extracting text from camera captured images as it computes the sub-regions of the images containing text. When the image is capture, the image goes through various processes to detect the sign and English text within the image, extract them and translates it into Hindi text. The different phases of the system are shown in the Figure1.
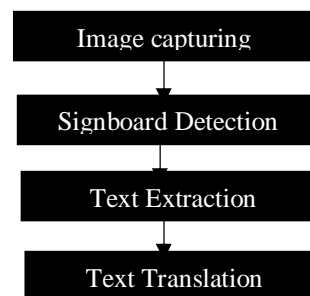

Figure 1. Phases

Capturing image of the signboard using the mobile phone camera is the very first step. The next phase is the signboard detection using faster regional convolutional neural network. Signboard detection is a challenging task because of having different layout with different angles and shapes. Text extraction is the next phase and it is done using Tesseract OCR. During last phase, i.e text translation, the extracted English text is translated to Hindi.

2. LITERATURE SURVEY

In the evolution of a human machine interactive software application named as "textract"[1] which can be used for text extraction and recognition in images captures via digital devices and mobile. Here a mobile phone act as a portable language translator where the text is later translated into another language. Challenges faced by this work is that there is a variety of text: font, size, skew angle, shape of the object which text lies on, etc. Textract better results in recognizing large texts. If the text are too thin then certain parts of the text tends to fade out. The processing time cane be reduced if mobile phone processor speed and the memory size is increased. Considering the nature of mobile devices, the proposed solution gives the best possible choices to balance between processing speed and recognition accuracy.

Another work introduces a frame work to extract text from natural scenes[9]. For extracting information from the natural scenes, both machine learning algorithm and pattern recognition plays a vital role. Machine learning is one of the difficult topics in IT. The text recognition is divided into several levels. Firstly, using a smart device on image is captured from the outside environment, followed by detection of the edges of the signboard. In the next level, detection of text and recognition of the text into two languages like Urdu and English is done. Finally, ANN is used for classification and recognition of text extracted from the outside environment.

To identify and recognize a traffic sign board in various lighting and background condition from static digital images[3]. It is possible to prevent accidents by using an automatic road sign recognition system traffic information to the driver. In future this processing can give information to a theoretical autonomous vehicle, heads up display or another driver assistance device. Here from the image the software uses color processing techniques to isolate relevant color data. A different variety of image processing techniques are used to detect edges, threshold, filter and further process the image. Morphological processing methods are used to remove non pertinent data and isolate signboards. Shape detection is applied to determine whether a sign is present in image, if present it will be resized and classified. Finally, the recognized traffic signboard is set as voice and text message.

A methodology that covers detection, orientation prediction and recognition of Urdu in images [11]. Firstly, Faster-RCNN algorithm is used, conjunction with CNNs like Resnet50,Resnet18,Google-net, Squeeze-net for localization and detection of images with size 320*240 pixels. For ligature

Orientation they used RRNN (Regression Residual Neural Network) on random oriented ligatures in the dataset. Two Stream Deep Neural Network (TSDNN) was used for the recognition of ligatures. Finally, all four was used to compare and evaluate their ability to detect and localize. Primary challenge they had to face was the detection of Urdu-text due to different types of scripts, large variance of text patterns, scale, orientation, color, background etc. These issues make it difficult for the photo-OCR system to detect and recognize with maximum accuracy.

A work was proposed for the travelers [6] who are facing troubles in understanding the signboards which are written in regional language. This project intended to build a web application which can recognize the English text on signboard captured using a mobile phone and translate the English text to Telugu and will display the translated Telugu text on the screen. The extracted text in English will be recognized by the system and it will verify the dictionary and the corresponding Telugu text will be displayed on the screen.

A FRCNN work was proposed to detect and localize potholes and road markers on the bird's eye view images of the road. It is mainly for the safety of driver's; it provides critical information for traffic guidance. A Point Grey Blackfly camera was mounted on the roof of a car to collect the data. To transform the image into a bird's eye view perspective an Inverse Perspective Mapping (IPM) algorithm is used. Here FRCNN is used due to its ability to identify and localize lanes, arrows, and potholes based on the test image automatically.

3. INDIAN SIGNBOARD TRANSLATOR

The proposed system captures the image of the signboard using a mobile phone. The image is then pre-processed to extract the text. The extracted text is then converted to Hindi language. The symbol in the image is also recognized and is delivered to the users. The proposed system follows these steps: (1) Image capturing, (2) Signboard detection, (3) Text extraction, (4) Text translation.

1. Image capturing

A mobile camera is used to capture the image of the signboard. The captured image may contain complex background and be noisy. So, the images should be pre-processed so that the system can extract the text easily.

2. Signboard detection

Faster R-CNN is used for signboard detection. Here, the input image is passed to the conVNet which returns a feature maps for the image. Object proposals are obtained by applying Region Proposal Network (RPN) on these feature maps. In order to bring all the proposals to the same size, ROI pooling layer is applied. Finally, these proposals are passed to a fully connected layer to predict the bounding boxes for the image.
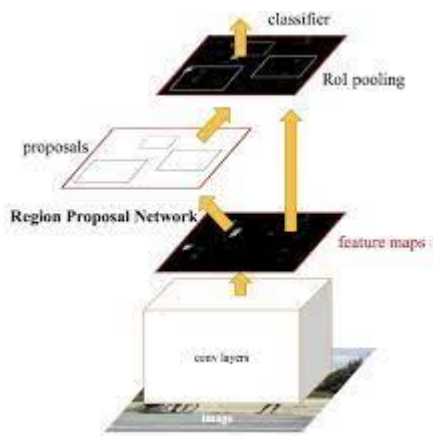
Figure 2. FRCNN

3. Text extraction

Optical character recognition (OCR) is the process of extraction of text from the captured image. In this paper, Tesseract OCR engine which is an open-source OCR engine having the ability to recognize many languages is used for text extraction. The OCR process is divided into following phases: pre-processing, segmentation, feature extraction and classification and post processing.
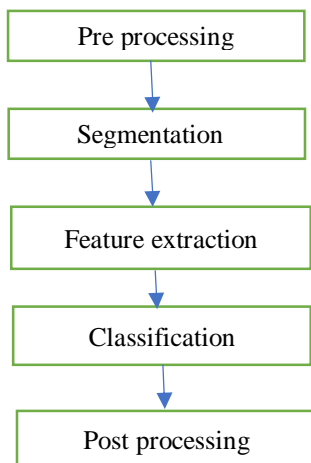

Figure 3. OCR phases

3.1 Pre-processing

Pre-processing phase involves image contrast enhancement, noise removal, binarization and smoothing. To enhance the contrast the image is subjected to adaptive histogram equalization which increase the contrast and reduced the imperfections in the image. This image is processed with median filter to put down the noise content which is subjected to locally adaptive threshold to obtain binary image. Finally smoothing which eliminates the small breaks, gaps, holes in the digitized characters and reduces the width in the lines by filling and thinning.

3.2 Segmentation

Segmentation is a process which helps to divide each character from a word present in the image. The objective of segmentation is to extract each character from the image. After performingsegmentation, the characters of string are separated and used for further processing.

3.3 Feature extraction and Classification

In this stage, the features of the characters which are crucial for classifying them at recognition stage are extracted and classification is the process of identifying each character and assigning to it the correct character.

3.4 Postprocessing

Postprocessing involves grouping of symbols and error handling. The process of performing the association of symbols into strings is commonly referred to as grouping. It is based on the symbol's location in the document. Symbols that are found to be sufficiently close are grouped together.

4. Text translation

To do the translation of extracted text from the signboard image, we created a dictionary of warning or hazards in English which generally appear on the signboards along with their meaning in Hindi. Whenever an English text is recognized by the system, it will verify the dictionary and the corresponding Hindi text is retrieved. This Hindi text is displayed over the mobile phone display.

4. EXPERIMENT AND RESULT

The image of signboard containing sign is given as input and it is shown in figure 4. Figure 5 shows the sign detected with accuracy 93% and the sign detection and translation are shown in figure 6.
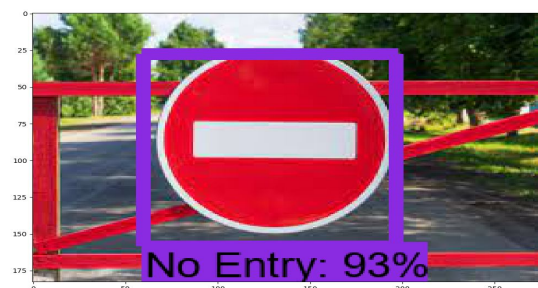

Figure 4. Input

Figure 5. Image Detected

3

RESULT OBJECT DETECTION

No Entry

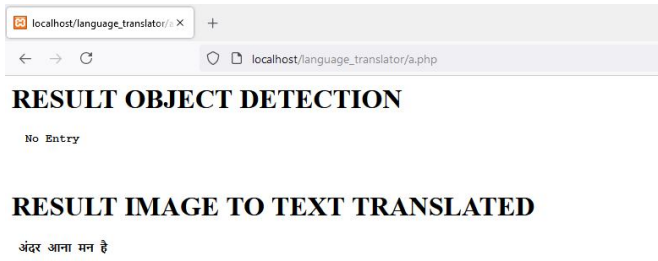**RESULT IMAGE TO TEXT TRANSLATED**

अंदर आना मन है

Figure 6. Result

To evaluate text extraction and translation performed by the system, a signboard image containing English text is given as the input to the system as shown in figure 8. The figure 9 depicts the text extraction andthe translated text.

Figure 7. Signboard image with English text

RESULT IMAGE TO TEXT

CAUTION
AREA UNDER
CONSTRUCTION

RESULT IMAGE TO TEXT TRANSLATED
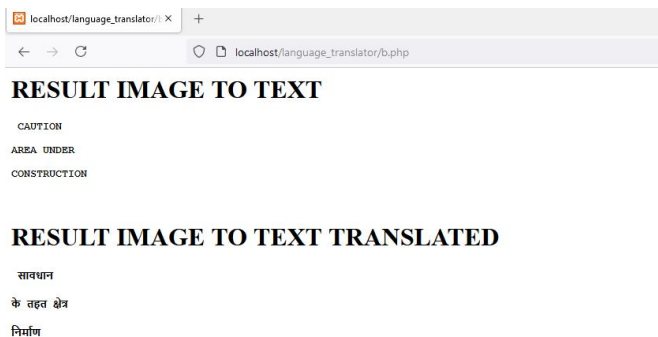
सावधान
के तहत क्षेत्र
निर्माण

Figure 8. English text extracted and translated into Hindi

The accuracy of the system has been measured by containing the number of correctly recognized images. Figure 9 depicts the recognition along with training samples. The accuracy of the system is shown in figure 10. Table 1 shows the accuracy of the system concerning its steps.
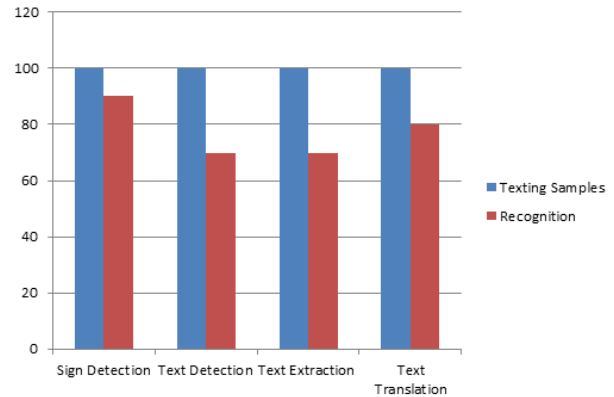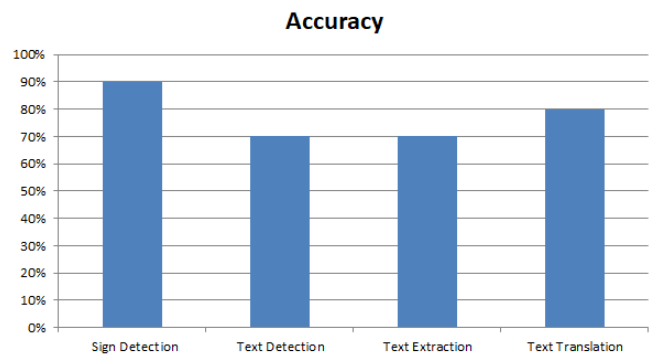
Figure 9. Recognition

Figure 10. Accuracy

Overall steps	Testing Samples	Recognition	Accuracy
Sign Detection	100	90	90
Text Detection	100	70	70
Text Extraction	100	70	70
Text Translation	100	80	80

Table 1. Accuracy with steps.

5. CONCLUSION

This system is developed to make it easier for travelers who travel across various parts of India to translate the signboards in English which will be difficult for them to understand. The system translates signboard images captured using mobile phone from English to Hindi. The purpose of this project is to remove the language barrier that people are facing nowadays. The proposed system is able to detect the signs, translate the texts in different font, color, light etc.

REFERENCES

[1] Jian Yuan, Yi Zhang, KokKiong Tan Tong Heng Lee.**Text extraction from images captured via mobile and digital devices**, 2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics Suntech Convention and Exhibition Center Singapore, July 14-17, 2009.

[2] Aparna A. Dalve, Sankriti S. Shiravale Department of Computer Engineering, MMCOE. Pune, Maharashtra, India. **Real time traffic signboard detection and recognition from street level imagery for smart vehicle**. International Journal of Computer Applications (0975-8887) Volume 135-No.1,February 2016.

[3] G. Revathi, Dr.G.Balakrishnan. **Indian sign board recognition using image processing techniques.** International Journal of Advanced Research in Biology Engineering Science and Technology(IJARBEST) Vol 2, Special Issue15, March 2016.

[4] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, **Faster R-CNN: Towards real time object detection with region proposal networks**, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016.

[5] G.R.Hemalakshni, M.Sakthimanimala, J.Salai Ani Muthu, Assistant Professor, U.G. Student Department ofComputer Science And Engineering, National Engineering College, TamilNadu, **Extraction of text from an image and its language translation using OCR**, International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Volume 4, April 2017.

[6] IlaiahKavati, G Kiran Kumar, Sarika Kesagani and K Srinivasa Rao. **Signboard text translator: A guide to tourist**, International Journal of Electrical and Computer Engineering (IJECE) Vol. 7, No. 5, October 2017.

[7] Deepti Kaushik and Vivek Singh Verma, **Review on text recognition in natural scene images**. Springer Nature Singapore Pte Ltd, B. Panda et al. (eds.), Innovations in Computational Intelligence, 2017.

[8] Syed Mazhar Abbas, Dr. Shailendra Narayan Singh, Amity University Uttar Pradesh, International Conference on Computational Intelligence and CommunicationTechnology, **Region based object detection and classificationusing Faster R-CNN**, 2018.

[9] Muhammad A. Panhwar, Kamran A.Memon, Adeel Abro, Deng Zhongliang, SaleemullahMemon,Sijad A .Khuro, **Signboard detection and recognition using artificial neural network**, 2019 IEEE.

[10] NazrahZulkifli, Muhammad Amirul IdhamLaili, AzleeZAbidi, **Fast region convolutional neural network lane and road defect detection for autonomous vehicle application**. International Journal of Advanced Trends in Computer Science and Engineering. Vol 8,No.1.3, 2019

[11] Syed Yasser Arafat and MuhammdJaved Iqbal, **Urdutext detection and recognition in natural scene images using deep learning**, Received April 21, 2020, accepted May 3.2020, date of publication May 12, 2020.