DSAA PROJECT

INITIAL PROPOSAL

GROUP-11

Members: Vaishnavi NV, Syed Jahangir Peeran, V Tejkiran, V Sai Rathan, Srilekha N.

Aim

Extracting meaningful text from a given image using Optical Character recognition (OCR) along with machine learning and then converting the extracted text into speech.

Applications

(i) A software which reads out bedtime stories or as a textbook reader for students.

(ii) An OCR based app with access to camera and speech which communicates the street name/door number to a blind person (through speech).

(iii) A JEE aspirant vigorously preparing can use such a software to read out lines in the chemistry book, so that he/she can remember the chemical formulae and related stuff by reading as well as listening.

Challenges

(i) We will try to recognise even broken text in an image.

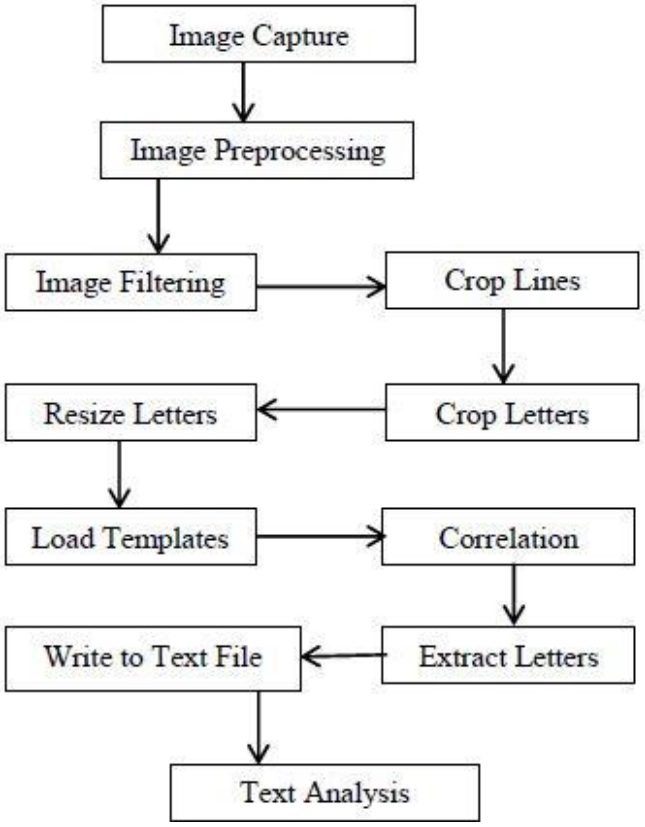
(ii) With the addition of using machine learning algorithm, the accuracy of OCR will be lifted up.

(iii) For better accuracy, we will run the converted text through grammar/spelling checker to predict the word/sentence better.

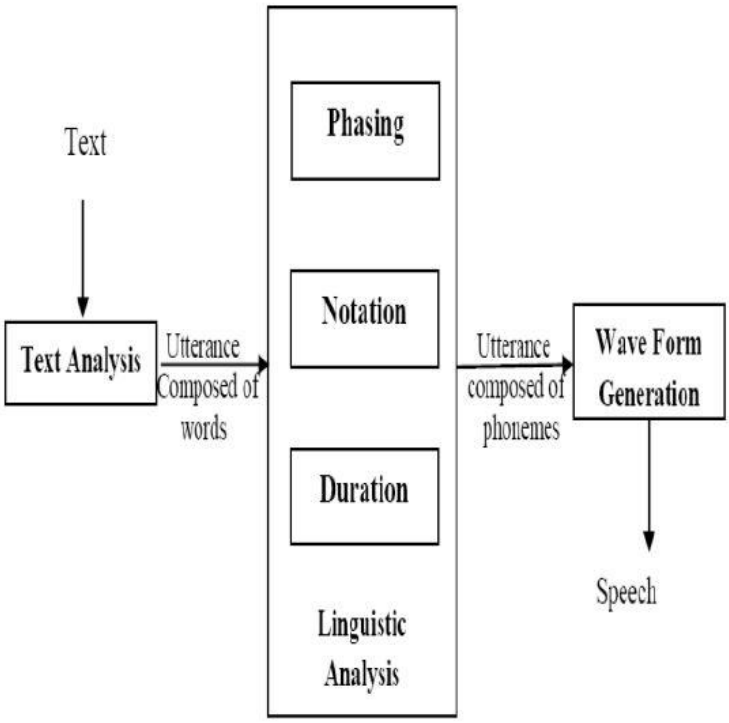
Input

A photo with text written on it, will be sent as the input for processing.

Processing



(i) Detect text regions alone from the image using sliding windows technique.
(ii) Binarization of Image.
(iii) Segmenting characters.
(iv) Classifying characters.
(v) Feature Extraction.
(vi) Recognition and Combining the letters and merging words to form sentences.
(vii) Checking its correctness.



(viii) Speech Synthesis is done using 'System.Speech' assembly in MATLAB.

(ix) A voice object is obtained from this synthesizer and voice is generated when speak function of this object is called.

A voice signal speaking out the text which was processed in the above step.

References

(i) Chen, Huizhong, et al. "Robust Text Detection in Natural Images with Edge-Enhanced Maximally Stable Extremal Regions." Image Processing (ICIP), 2011 18th IEEE International Conference on. IEEE, 2011.

(ii) http://www.voicerss.org/tts/

(iii) http://www.comsys.net/technology/speechframe/text-to-speech-tts.html

(iv) https://uk.mathworks.com/help/vision/ref/ocr.html

(v) https://www.coursera.org/learn/machine-learning