International Journal of New Technology and Research (IJNTR)
ISSN:2454-4116, Volume-4, Issue-4, April 2018 Pages 70-72

Smart Gadget Product Label Reading Using OCR Algorithm & TTS Engine

R.Abhishek, K.Nanda Kumar, R.Karthik, R.Pushkal, S.Arun kumar

***Abstract*—Assistive technologies are being developed for visually impaired people in order to live confidently. This project work proposes a camera-based assistive text reading framework to help blind persons read text labels and product packaging from hand-held objects in their daily lives. The project work is framed into three stages. First, Image capturing Using a mini camera, the text which the user need to read get captured as an image and have to send to the image processing Platform. Secondly, Text recognition Using text recognition algorithm, the text will get filtered from the image. Finally, Speech output. A filtered text will be passed into this system to get an audio output. This project work can be able to insist the blind people in their daily life. The entire application will run on Raspberry Pi.**

***Index Terms*— Optical Character Recognition (OCR), Image Processing, OpenCV, Tesseract Tool, Text-To-Speech(TTS).**

I. INTRODUCTION

Although a number of reading assistants have been designed specifically for the visually impaired, to our knowledge, no existing reading assistant can read text from the kinds of challenging patterns and backgrounds found on many everyday commercial products. such text information can appear in multiple scales, fonts, colors, and orientations. To assist blind persons to read text from these kinds of hand-held objects, we have conceived of a camera-based assistive text reading framework to track the object of interest within the camera view and extract print text information from the object. Our proposed algorithm can effectively handle complex background and multiple patterns, and extract text information from both hand-held objects and nearby signage. In assistive reading systems for blind persons, it is very challenging for users to position the object of interest within the center of the camera's view. As of now, there are still no acceptable solutions. We approach the problem in stages. To make sure the hand-held object appears in the camera view, we use a camera with sufficiently wide angle to accommodate users with only approximate aim. This may often result in other text objects appearing in the camera's view (for example, while shopping at a supermarket). To extract the hand-held object from the camera image, we develop a motion-based method to obtain a region of interest (ROI) of the object. Then, we perform text recognition only in this ROI. It is a challenging problem to automatically localize objects and text ROIs from captured images with complex backgrounds, because text in captured images is most likely surrounded by various background outlier "noise," and text characters usually appear in multiple scales, fonts, and colors. For the text orientations, this paper assumes that text strings in scene images keep approximately horizontal alignment. Many algorithms have been developed for localization of text regions in scene images. We divide them into two categories: rule-based and learning-based Rule-based algorithms apply pixel-level image processing to extract text information from predefined text layouts such as character size, aspect ratio, edge density, character structure, color uniformity of text string. In solving the task at hand, to extract text information from complex backgrounds with multiple and variable text patterns, we here use Optical character recognition algorithm that combines rule-based layout analysis and learning-based text classifier training, which define novel feature maps based on stroke orientations and edge distributions.

II. EXISTING SYSTEM

An end-to-end real time text localization and recognition method was proposed, which does not rely on any prior knowledge of words to be detected. Its real time performance is achieved by posing the character detection and segmentation problem as an efficient sequential selection from the set of External Regions. In the first stage, the probability of each ER being a character was estimated using the features. Only the ERs with locally maximal probability were selected for the second stage, where the classification accuracy was improved using more computationally expensive features. A highly efficient clustering algorithm was used, which groups ERs into text lines and an OCR classifier trained on synthetic fonts is exploited to label character regions.

III. PROPOSED SYSTEM

We propose a camera-based label reader to help blind persons to read label name of the shop and the products. Camera acts as main vision in detecting the label image of the product or board then image is processed internally. And separates label from image and finally identifies the text and identified text is given as voice output. In order to make the process much faster the proposed system makes use of Open CV platform which helps to reduce the time in order to achieve the real time performance. OCR algorithm is used for the text recognition phase.

R.Abhishek, Computer Science, SRM University, Chennai, India
K.Nandha Kumar, Computer Science, SRM University, Chennai, India
R.Karthik, Computer Science, SRM University, Chennai, India
R.Pushkal, Computer Science, SRM University, Chennai, India
S.Arun kumar, Computer Science, SRM University, Chennai, India

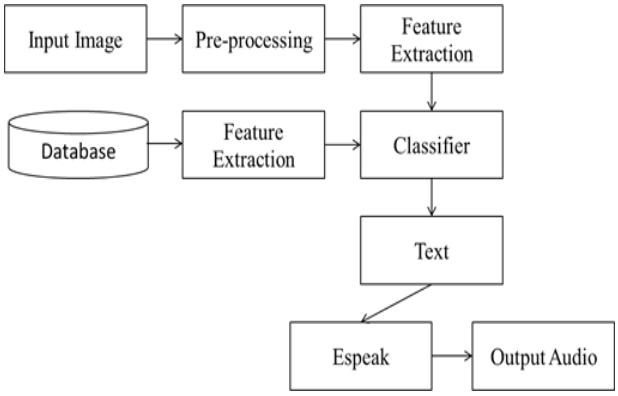

Figure 1. Basic system architecture.

A. *Optical Character Recognition*

Optical character recognition (also optical, character, reader, OCR) is the mechanical or electronic conversion f images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo or from subtitle text superimposed on an image. It is a common method of digitizing printed texts so that they can be electronically edited, searched, stored more compactly, displayed on-line, and used in machine processes such as cognitive computing, machine translation, (extracted)text-to-speech, key data and text mining. OCR is a field of research in pattern recognition, artificial intelligence and computer vision.

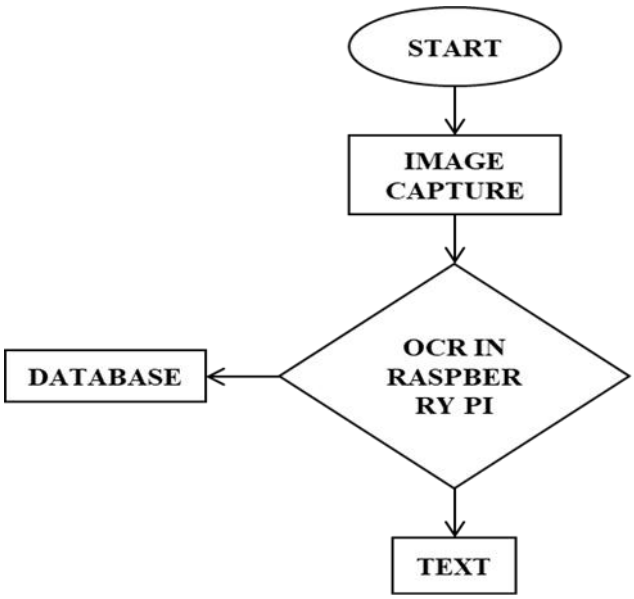

Figure 2. Working Of OCR Algorithm.

A frame sequence *V* is captured by a camera worn by blind users, containing their hand-held objects and cluttered background. To extract text information from the objects, motion-based object detection is first applied to determine the user's object of interest *S* by shaking the object while recording video

$$S = \frac{1}{|V|}\sum_{i} \mathcal{R}(V_i, B)$$

Where *Vi* denotes the *i*th frame in the captured sequence, |*V* |denotes the number of frames, *B* denotes the estimated background from motion-based object detection, and *R* represents the calculated foreground object at each frame.

Text localization algorithm is applied to the object of interest to extract text regions. At first, candidate text regions are generated by layout analysis of color uniformity and horizontal alignment

$$X^C = \text{argmax}_{s \in S} L(s)$$

Where L(s) denotes the suitability responses of text layout and XC denotes the candidate text regions from object of interest S using stroke orientations and edge distributions of text characters.

Open CV: It is a library of programming functions mainly aimed at real-time computer vision. Originally developed by Intel. It has C++, C, Python and Java interfaces and supports Windows, Linux, Mac OS, iOS & Android.

Tesseract Tool: It is a tool which uses the OCR algorithm to find the character from a paper or image and converts the character into text format.

B. *Data Collection*

Camera: When a product or product label has to be read then the camera module is used to capture the image of the product label to process and get the output.

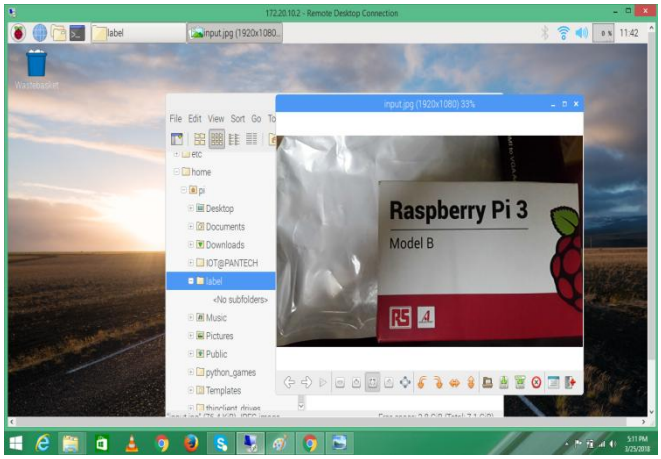

Figure.3 Image Captured Using A Camera.

Image Processing: Loading any image format (bmp, jpg, png) from given source. Then convert the image to grayscale and binarize it using the threshold value. Detecting image features like resolution and inversion. So that we can finally convert it to a straightened image for further processing. Lines detection and removing. This step is required to improve page layout analysis, to achieve better recognition quality text. Page layout analysis. In this step I am trying to identify the text zones present in the image. So that only that portion is used for recognition and rest of the region is left out. Recognition of characters. This is the main algorithm of OCR; an image of every character must be converted to appropriate character code. Sometimes this algorithm produces. several. Character codes for uncertain images. For instance, recognition of the image of "I" character can produce "I", "|" "1", "l" codes and the final character code will

be selected later. Saving results to selected output format, for instance, searchable .TXT. It is important to save original page layout: columns, fonts, colors, pictures, background and so on.

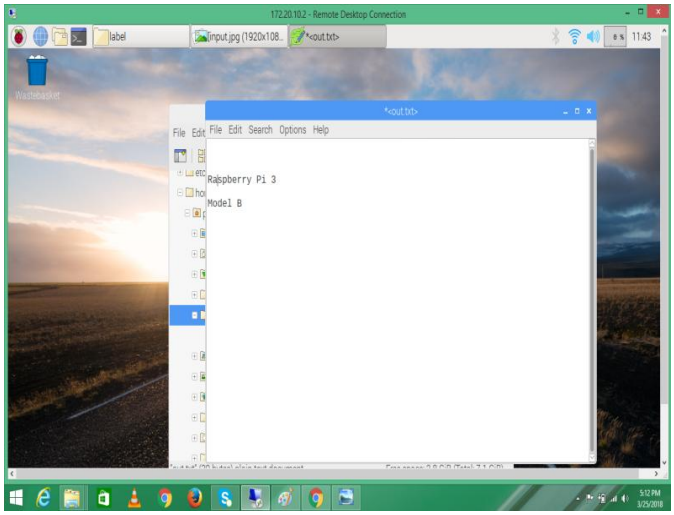

Figure 4. Converted Image Into Text

Text-To-Speech Engine: Espeak is a compact open source software speech synthesizer which converts the text which is extracted by the OCR and gives us output in the audio format. The audio can be listened by speaker or headphones.

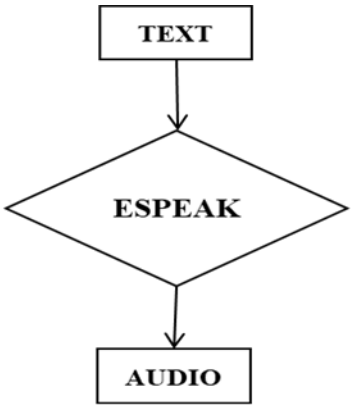

Figure.5 Working of Espeak.

IV. CONCLUSION

In this paper successfully presents the text detection and recognition with Audio output system was successfully demonstrated on Open CV platform. This system is very handy and useful for the visually challenged persons. Compared with a PC platform, the Open CV platform is portable and more convenient to use. This system will be helpful for visually challenged persons to access information in written form and in the surrounding. It is useful to understand the written text messages, warnings, and traffic direction in voice form by converting it from Text to voice. It is found that this system is capable of converting the sign boards and other text into speech.

ACKNOWLEDGMENT

I would like to express my special thanks of gratitude to my Guide Mr. S. Arun Kumar (M.Tech) as well as our Head Of Our Department Dr.J.Jagadeesan who gave me the golden opportunity to do this wonderful project on the topic Product Label Reading Using OCR Algorithm & TTS Engine, which also helped me in doing a lot of Research and i came to know about so many new things am really thankful to them. Secondly i would also like to thank my parents and friends who helped me a lot in finishing this project within the limited time.

REFERENCES

[1] Ms.Rupali, D.Dharmale, Dr. P.V. Ingole "text detection and recognition with speech output for visually challenged person“ ,vol.5, issue 1, January 2016.
[2] Nagaraja , Et al. "Vision based text recognition using raspberry pi." National conference on power systems, industrial automation (ncpsia 2015).
[3] Rajkumar N,Anand M.G, Barathiraja.N,“Portable camera based product label reading for blind people“ .Ieee vol. 10 number 11 –apr-2014.
[4] Boris Epstein, Eyalofek, Yonatan Wexler, "Detecting text in natural scenes with stroke width transform.
[5] Ezaki, Nobuo, “Improved text-detection methods for a camera-based text reading system for visually impaired persons."Eighth international conference on document analysis and recognition (icdar’05). Ieee, 2005.
[6] Ray Smith, “an overview of the tesseract OCR engine”.
[7] World Health Organization. (2009). 10 facts about blindness and visual impairment blindness/blindness_facts/en/index.html
[8] International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2005, 2007, 2009, 2011).