International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048
Volume 6, Issue 5, pp 220-224, May-2022
https://doi.org/10.47001/IRJIET/2022.605032

Text Identification from Input Images Using Support Vector Machine Algorithm and Its Method for Product Label Reading System

[1]K.M.Priyakumari, [2]N.Rupavathi, [3]Dr.K.Ramesh, [4]R.Bhuvaneswari, [5]S.Venkatesh

[1]PG Scholar, Applied Electronics, Jayam College of Engineering and Technology, Dharmapuri, Tamilnadu, India
[2]Associate Professor, Dept. of ECE, Jayam College of Engineering and Technology, Dharmapuri, Tamilnadu, India
[3]Professor, Dept. of ECE, Jayam College of Engineering and Technology, Dharmapuri, Tamilnadu, India
[4,5]Assistant Professor, Dept. of ECE, Jayam College of Engineering and Technology, Dharmapuri, Tamilnadu, India

***Abstract* - An assistive text reading framework to help blind persons read text labels and product packaging from hand-held object in their daily resides is proposed. To isolate the object from cluttered backgrounds or other surroundings objects in the camera view, we propose an efficient and effective motion based method to define a region of interest (ROI) in the image of product. In the extracted ROI, text localization and recognition are conducted to acquire text information. To automatically localize the text regions from the object ROI, we propose a novel text localization algorithm by learning gradient features of stroke orientations and distributions of edge pixels in an Ada-boost model. Text characters in the localized text regions are then binarized and recognized by off-the shelf optical character recognition software. The recognized text codes are output to blind users in speech. Nowadays printed text appears everywhere like product names, restaurant menus, instructions on bottles, signed boards etc. Thus blind people need some assistance to read this text .This paper presents a camera-based product information reader to help blind persons to read information of the products. Camera acts as main vision in detecting the label image of the product then image is processed internally and separates label from image by using MATLAB and finally identifies the product name and identified product information is pronounced through the optical character recognition (OCR). The OCR is used to convert the text from text regions and then converted to voice output.**

Keywords: optical character recognition, OCR, text identification, text to speech, product label.

I. INTRODUCTION

Recent developments in computer vision, digital cameras, and portable computers make it feasible to assist these individuals by developing camera-based products that combine computer vision technology with other existing commercial products such optical character recognition (OCR) systems. Reading is obviously essential in today's society. Printed text is everywhere in the form of reports, receipts, bank statements, restaurant menus, classroom handouts, product packages, medicine bottles, etc.

And while optical aids, video magnifiers, and screen readers can help blind users and those with low vision to access documents, there are few devices that can provide good access to common hand-held objects such as product packages, and object sprinted with text such as prescription medication bottles. The ability of people who are blind or have significant visual impairments to read printed labels and product packages will enhance independent living and foster economic and social self sufficiency and portable computers make it feasible to assist these individuals by developing camera-based products that combine computer vision technology with other existing commercial products such optical character recognition (OCR) systems. The ability of people who are blind or have significant visual impairments to read printed labels and product packages will enhance independent living and foster economic and social self sufficiency.

II. PROPOSED METHODOLOGY

We propose a ROI and SVM based helpful text reading framework to assist blind persons scan text labels and merchandise packaging from hand-held objects in their daily lives. 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 image of camera view. As of now, there are still no acceptable solutions. The most focus of our analysis is that the sand-blind person will get info concerning written text, text boards, scene text, hoardings, and directions on traffic sign boards in audio kind.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 6, Issue 5, pp 220-224, May-2022

https://doi.org/10.47001/IRJIET/2022.605032

Block Diagram

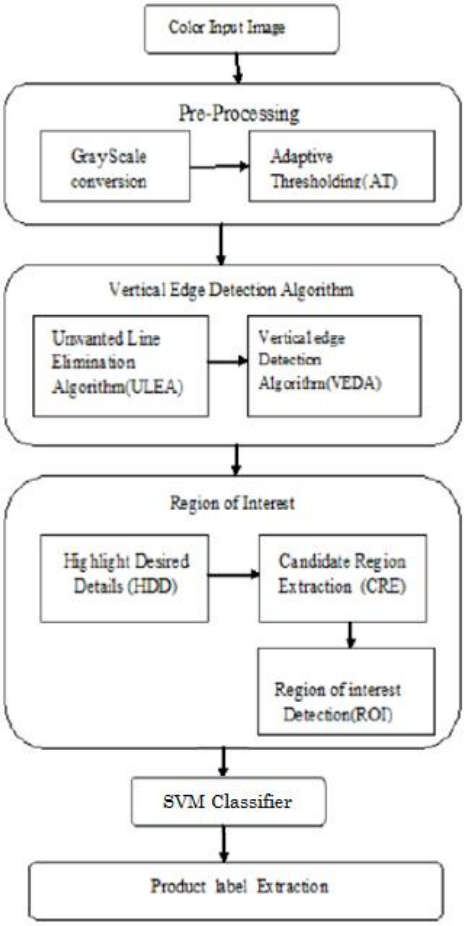

Figure 1: Proposed System block diagram

System Description

The proposed method consists of image acquisition, preprocessing the images with image-level segmentation, feature extraction, feature representation and classification using a Support Vector Machine (SVM).

We approach the problem in stages. For the text orientations, this system assumes that text strings in scene images keep approximately horizontal alignment. Many algorithms have been developed for localization of text regions in scene images. 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.

TTS Module A text-to-speech (TTS) system converts normal text into speech other systems render symbolic linguistic representations like phonetic transcription into speech. A text-to-speech system is used to read each word as the user`s finger passes over it, and distinctive audio and/or haptic cues can be used to signal other events, such as end of line, start of line etc.

III. PROPOSED ALGORITHMS AND SOFTWARE FLOW PROCESS

Adaptive Thresholding

After the color input image is converted to gray scale, an Adaptive Thresholding (AT) process is applied to constitute the binaries image. The idea used in algorithm is that the Pixel is compared with an average of neighboring pixels. Specifically, an approximate moving average of the last pixels seen is calculated while traversing the image. The value of current pixel is set to black if the value of current pixel is T percent lower than the average value, else it is set to white. This is very useful technique because comparing a pixel to the average of neighboring pixels will keep hard contrast lines and ignore soft gradient changes.

Vertical Edge Detection Algorithm

The Vertical Edge Detection Algorithm is to distinguish the region of interest, particularly the starting and the end of each character. Therefore, the text details will be easily detected, and the character recognition process will be faster. After thresholding and Unwanted Line Elimination Algorithm processes, the image will only have black and white regions, and the Vertical Edge Detection Algorithm is applied to these regions.

Region of Interest (ROI)

After applying the Vertical Edge Detection Algorithm, the next step is to highlight the desired details such as text and vertical edges in the image. Using text localization algorithm the region of interest is selected.

Text Recognition

The task of text recognition is performed by off-the-shelf OCR prior to output of informative words from the localized text region. A text region labels the minimum rectangular area for the accommodation of characters inside it, and border of the text region contacts the edge boundary of the text character. We propose to use Template matching algorithm for OCR. The output of the OCR is nothing but a text file containing the product label (its name) in textual form. Audio output component is to inform the blind user of recognize text code in the form of speech or Audio.

SVM Classification Process

A support vector machine (SVM) is a concept in statistics and computer science for a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis. The standard SVM takes a set of input data and predicts, for each given input,

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 6, Issue 5, pp 220-224, May-2022

https://doi.org/10.47001/IRJIET/2022.605032

which of two possible classes forms the input, making the SVM a non-probabilistic binary linear classifier. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible.

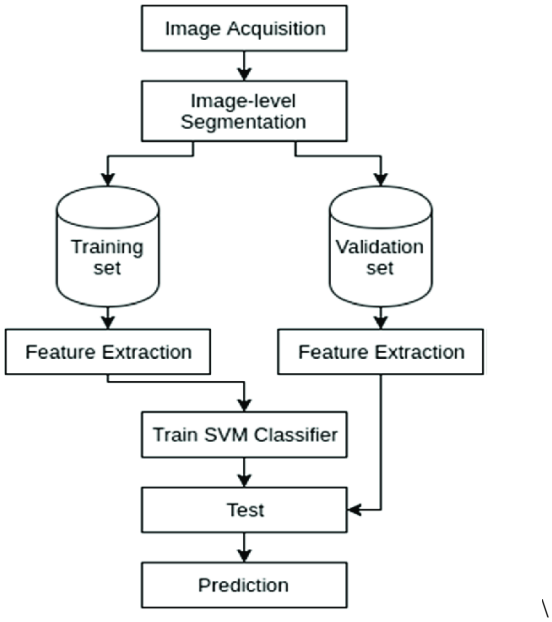

Figure 2: Block diagram of the SVM classification process

Optical Character Recognition (OCR)

OCR is optical character recognition module is that the mechanical or electronic conversion of pictures of written, handwritten or printed text into machine-encoded text. it's a standard technique of digitizing written text so it is utilized in machine method like text-to-speech. OCR is optical character recognition module is that the mechanical or electronic conversion of pictures of written, handwritten or printed text into machine-encoded text. The method of text extraction is dole out by matching with templates one by one and so forming a full word. The mentioned line or a word can be scan from the captured input text with a appropriate secret writing. Once matching with the templates and displays it as a text and reads it aurally.

Text Extraction Module

This module will recognize and extract the text. This will be achieved using OCR-Optical Character Recognition - is the Mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text. We will be using here MODI algorithm of OCR.

TTS

TTS Module A text-to-speech (TTS) system converts normal text into speech other systems render symbolic linguistic representations like phonetic transcription into speech. A text-to-speech system is used to read each word as the user`s finger passes over it, and distinctive audio and/or haptic cues can be used to signal other events, such as end of line, start of line etc.

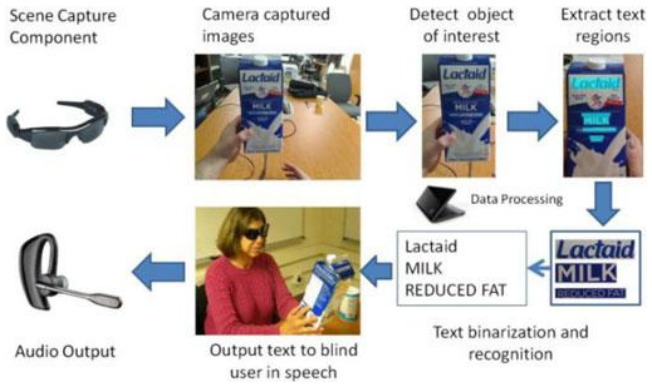

Figure 3: Flowchart of the proposed framework to read text from hand-held objects for blind users

III. RESULTS AND DISCUSSION

In this section the Adaptive Thresholding process is evaluated first. Then, the accuracy and the computation time of the VEDA are compared with that of the Sobel operator. Finally, the performance of the proposed Text detection method is evaluated. Firstly, Using Web camera the image of the product is captured. Fig 6(a) shows the original image. This image is then processed internally using Matlab. In fig 6(b) original color image is transformed into grey scale image for further processing of image. Grey image is calculated by taking average of each component in the color image.

Figure 4: Original image

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 6, Issue 5, pp 220-224, May-2022

https://doi.org/10.47001/IRJIET/2022.605032

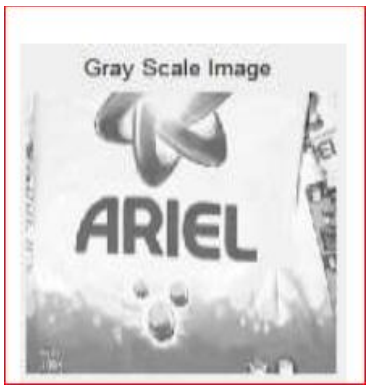

Figure 5: Gray scale image

Figure 6: Thresholded Image

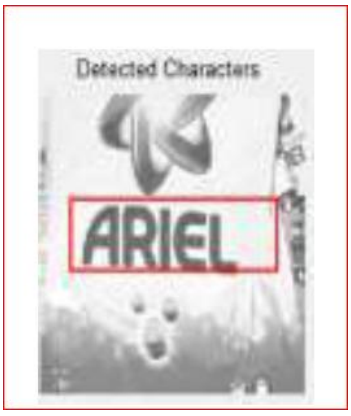

Figure 7: Candidate Region

Figure 8: Region of Interest

Figure 9: Binarized characters

Figure 10: Identified Product Name

Text Recognition and Audio Output

In the audio output part the blind user will get speech output of recognized text information. A Bluetooth earpiece or headphones is used for audio output. The task of text recognition is performed by off the shelf OCR to output of informative words of from text localized text regions. The text characters in the recognized text regions are accommodated in minimum rectangular area , hence the border of the text regions contacts the boundary of the text characters. The OCR performs better if text regions are aligned with proper margins and to segment text characters from clutter background. The height and width of each text localized region is enlarged by 10 pixels respectively. Then binarization of text region is performed by OTS's method, where text margin areas are considered as background. Then these recognized text codes are recorded as script file using Matlab software. Here blind user can adjust speech rate, tone, volume according to their performance.

IV. CONCLUSION

In this project, a prototype system framework to read printed text on hand held object for assisting blind person is described. This paper proposed a method for detecting texts in product images. First input image is captured by camera and it is converted to grey scale image. Then image enhancement technique is used to improve quality of an image. Using adaptive thresholding method the Binarized image is obtained. Then text region are detected by VEDA. After finding the Edges the feature of each text is calculated to recognize the text in the image. Then each character in the image was classified and correct product name is identified. Finally detected product name is pronounced the speech by using text

ISSN (online): 2581-3048
Volume 6, Issue 5, pp 220-224, May-2022
https://doi.org/10.47001/IRJIET/2022.605032

to speech converter in the matlab. Our future scope is to enhance these techniques for detecting text in different structures, fonts styles etc.

REFERENCES

[1] Yi, C., Tian, Y., & Arditi, A. (2014). Portable Camera-Based Assistive Text and Product Label Reading from Hand-Held Objects for Blind Persons. IEEE/ASME Transactions on Mechatronics, 19(3), 808–817.doi:10.1109/tmech.2013.2261083.

[2] R. U. Shekokar , A Survey On Camera Based Assistive Technologies For Visually Impaired Peoples, 2016.

[3] Ramesh Babu.Y, Vision Based Assistive System For Label Detection With Voice Output, 2017.

[4] Ch.Sridevi Reddy, Image Processing Based Assistive System for Label Detection with Voice Output For Blind People, IJRRECS/August 2016

[5] X. Chen, J. Yang, J. Zhang and A. Waibel, "Automatic detection and recognition of signs from natural scenes," In IEEE Transactions on image processing, Vol. 13, No. 1, pp. 87-99, 2004.

[6] D. Rajkumar and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: a survey," In IEEE Transactions on systems, man, and cybernetics, Vol. 40, No. 1, pp. 25-35, 2010.

[7] B. Epshtein, E. Ofek and Y. Wexler, "Detecting text in natural scenes with stroke width transform," In CVPR, pp. 2963-2970, 2010.

[8] Y. Freund and R. Schapire, "Experiments with a new boosting algorithm," In Int. Conf. on Machine Learning, pp.148–156, 1996.

[9] N. Giudice and G. Legge, Blind navigation and the role of technology, in The engineering handbook of smart technology for aging, disability, and independence, 2008, Wiley: Hoboken, N.J.

Citation of this Article:

K.M.Priyakumari, N.Rupavathi, Dr.K.Ramesh, R.Bhuvaneswari, S.Venkatesh, "Text Identification from Input Images Using Support Vector Machine Algorithm and Its Method for Product Label Reading System" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 6, Issue 5, pp 220-224, May 2022. Article DOI https://doi.org/10.47001/IRJIET/2022.605032

****

