ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 6, Issue 11, November 2019

Text Detection and Recognition for Calorie-conscious Life style

[1] Mr. Aran Nash, [2] Ms. Swathika R, [3] Ms. N Radha
[1] Student, [2][3] Assistant Professor
[1][2][3] Department of Information Technology, SSN College of Engineering, Kalavakkam, Chennai, India

***Abstract*—Humans may not like tracking our food and the calories that every food item offers, as it can be more difficult than it sounds. We would need to do a lot of research and always keep a track of what we are eating every day. But, it can be one of the best things we can do to maintain a healthy weight and improve overall health. The amount of calorie we may intend to have depends on number of factors including our age, gender, daily schedule, weight and etc. Hence, in this paper we develop a system that recognizes the text on the restaurant bills and calculate corresponding calorie count thereby enlightening the user on the calories consumed. To detect text from the images, we use the help of a widespread technology – Optical Character Recognition (OCR). It is the mechanical or electronic conversion of images of handwritten or printed text into machine encoded text whether or not from a scanned document or photograph of a document. The user can be aware of the calorie consumed using the application developed. The objective of this system is to help users to keep track of calories on their smart phones, also helping their dietician prepare a calorie-based chart thereby improving their lifestyle.**

***Keywords*—Text detection, OCR, Text recognition, Calorie-chart, Smart phones**

I. INTRODUCTION

In our day to day lifestyle, keeping a journal or tracking our meals on our smart phone, make better choices. You become responsible to yourself, and you will track your progress finer. There is no more fooling yourself concerning what proportion you are or are not consuming once you write it down. Journaling and working your food decisions results in creating higher choices. We need some calories daily to support the normal functioning of the body. The amount of calorie we may take depends on many factors including our age, gender, daily schedule and weight along with many other factors. So, we would like to monitor these calories. Our work aims to utilize the Optical character recognition technique to detect the text on the restaurant bill accurately.

Text detection is the method of detecting the text exists in the image, followed by enclosing it with a rectangular bounding box. This can be achieved through image based techniques or frequency based techniques. In image based techniques, an image is divided into multiple segments. Each segment is a connected component of similar characteristics pixels. The statistical features of connected components are utilized to group them and form the text. Machine learning approaches are used to classify the components into text and non-text. In frequency based approaches, discrete Fourier transform (DFT) or discrete wavelet transform (DWT) are used to extract the high frequency coefficients. Here the assumption is that the text present in an image has high frequency components and selecting only the high frequency coefficients filters the text from the non-text regions in an image.

Text is the most excellent way of communicating information. It can be enclosed into documents, images and videos. The high performance mobile devices capable of acquisition and processing of images anytime, anywhere provides a way to recognize text in various environments. there have been numerous text related applications for both images and video, which includes page segmentation [1], [2], address block location [3], license plate location [4], [5], and content-based image/video indexing [6], [7].

Text detection and recognition in images with complex backgrounds, variations of text layout and fonts, and the existence of uneven illumination, low quality, degraded data and multilingual environments present a much greater challenge than well-formatted documents. So advanced computer vision and pattern recognition techniques are required to solve these problems. The performance of text detection and recognition in color imagery are examined and various technical challenges are discussed [8].

The aim of this paper is to develop a system that recognizes the text on the restaurant bills using optical character recognition techniques and to calculate corresponding calorie count thereby enlightening him/her on the calories consumed. We aim to provide this software on android platform which will ease the task of detecting text and computing calories.

The rest of this paper is organized as follows. Section

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 6, Issue 11, November 2019

II gives survey about the various text detection and recognition techniques from image. It also provides information about the OCR system. The proposed system is presented in Section III. Section IV reports about how the experiment is done and how the results are obtained. It also describes the challenges faced by the system and finally the conclusion and future enhancements are drawn in Section V.

II. REVIEW OF EXISTING WORK

Text comes in two categories: Graphic text and Scene text. The graphic text is machine printed text which is found in captions, subtitles and annotations in video and digital images on the web and in email [9]. The scene text refers to text on objects, includes text on signs, packages and clothing in natural scenes, and handwritten material [10]. Our proposed application uses graphic text i.e. printed text from restaurant bills.

The process of extracting text from mixed-type documents is a pre-processing and necessary step for several document related applications. A new method is proposed to automatically detect and extract text in mixed-type color documents based on a combination of an adaptive color reduction (ACR) technique and a page layout analysis (PLA) approach [11].

While the recent methods relies on visual cues only, a novel fine-grained object classification using recognized text in natural images is proposed. This approach combines textual as well as visual cues. Another new thing is the textual cue extraction. Unlike the state-of-the-art text detection methods, this method relies more on the background instead of text regions. Once text regions are found, they are further processed by two process to perform text recognition, i.e., OCR engine and a state-of-the-art character recognition algorithm [12]. A novel Text-Attentional Convolutional Neural Network (Text-CNN) system is developed for scene text detection. This method focuses on extracting deep text feature from an image component. Also it is able to discriminate ambiguous text from complicated background [13].

Edge based methods perform more efficiently in case of natural scene images having sharp edges. A method is presented in which candidate edge recombination and edge classification are used to stroke width transform are added. In candidate edge recombination, over segmentation and region merging is done. The edge of the image is segmented into small segment that separates the edge of text from the background. Then the merging of neighboring edge segment have been done the on the basis of similarity of color and stroke width. With this each character is explicated by a candidate boundary. The candidate boundaries are combined into text chains in the boundary classification process. The character based and chain based features are used in chain classification [14].

A real-time system which extracts text from natural scene images can prove to be very useful [15], [16]. These systems can be used to automatically read and understand addresses and landmarks from images, banners and posters on roads, and sign boards without much difficulty.

Although number of state-of-the-art approaches [17], [18], [19] have been proposed to detect and recognize text in scene imagery, there exists some research challenges like low detection rates and recognition rates [20]. By contrast, OCR achieves recognition rates higher than 99 percentage for scanned documents [21]. OCR is an electronic tool that converts scanned documents into text searchable form. A typical OCR system consists of the following basic components: Input scanned Image, Pre-processing, Feature Extraction, Classification, and Post-processing [22].

OCR [23] is a technology that converts different types of documents, such as scanned paper documents, PDF files or images captured by a digital camera into editable and searchable data. Possibly the most well-known use case for OCR is converting printed paper documents into machine-readable text documents. In OCR processing, the scanned-in image or bitmap is analyzed for light and dark areas in order to identify each alphabetic character or numeric digit. When a character is recognized, it is converted into an ASCII code. This technology has proven immensely useful in digitizing historic newspapers and texts that have now been converted into fully searchable formats and had made accessing those earlier texts easier and faster.

III. PROPOSED WORK

This application captures an image of a restaurant bill using the mobile phone camera. This image is then processed to find the text related to food items. The text should be detected irrespective of certain parameters like font type, font size and should be detected in low contrast images also. Figure 1 shows the system diagram which consists of four modules.

The first step is to scan the printed restaurant bills and extract the text from it. Next we refer the extracted text with food database and the text extracted is forwarded for calculating the corresponding calorie-count. The user is then enlightened of the calories consumed.

The system proposed in this paper is built using the OCR algorithm. Instead of creating a new OCR implementation, the open-source OCR engine called Tesseract [24] was used here. The engine is being developed and used by Google since 2006. However,

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 6, Issue 11, November 2019

several OCR engines use similar techniques to "read" the text in an image. The Tesseract supports various operating systems (Linux, Windows, and Mac OS) and several languages. It has been used for both handwritten as well as printed character recognition. The workflow and architecture can be found in [25].Tesseract uses a two-pass approach to character recognition. The second pass is known as adaptive recognition and uses the letter shapes recognised with high confidence on the first pass to recognise better the remaining letters on the second pass. This is advantageous for unusual fonts or low-quality scans where the font is blurred or faded.

A. Preprocessing

OCR software often "pre-processes" images to improve the chances of successful recognition. The first step of the processing will consist of different pre-processing tasks, such as converting a color image into grayscale image, remove positive and negative spots, smoothing edges and binarisation. The task of binarisation is performed as a simple way of separating the text or any other desired component of the image from the background.

B. Text Extraction

Once the images are captured, we now have this file that we can easily view from any computer system. The software that extracts text from image should always be equipped with three factors. First, it should be able to retain text and data formatting during the conversion step. Second aspect of the text extraction from image would be the extract data for content aggregation. Third and last, the text extraction from image should be able to preserve in full the text quality from the printed document.

C. Text Recognition

The OCR algorithm uses feature extraction or matrix matching technique to produce a ranked list of candidate characters. Matrix matching is the process of comparing an image to a stored glyph on a pixel-by-pixel basis. This technique works best with typewritten text. It does not work well when new fonts are encountered. Feature extraction segments glyphs into "features" like lines, closed loops, line direction, and line intersections. This reduces the dimensionality of the representation and hence makes the recognition process computationally efficient. These features are matched with an abstract vector-like representation of a character, which might reduce to one or more glyph prototypes. Nearest neighbour classifiers such as the k-nearest neighbours algorithm are used to compare image features with stored glyph features and choose the nearest match.

D. Calorie count generation

The text recognized is checked for availability in the database .If the text is available in the database, its corresponding calories is fetched, and displayed to the user. If the text is not available in the database, no calories are returned. Once all text in the image is checked, and their corresponding calorie count has been returned, their overall calorie count is generated. Finally the user is enlightened with the overall calorie count.

Here the preprocessing, text extraction and text recognition is done by the help of Google vision API. The Vision API can detect and extract text from any images. There are two annotation features that support optical character recognition. TEXT_DETECTION detects and extracts text from any image. DOCUMENT_TEXT_DETECTION also extracts text from an image, but the response is optimized for dense text and documents. The JSON includes page, block, paragraph, word, and break information.

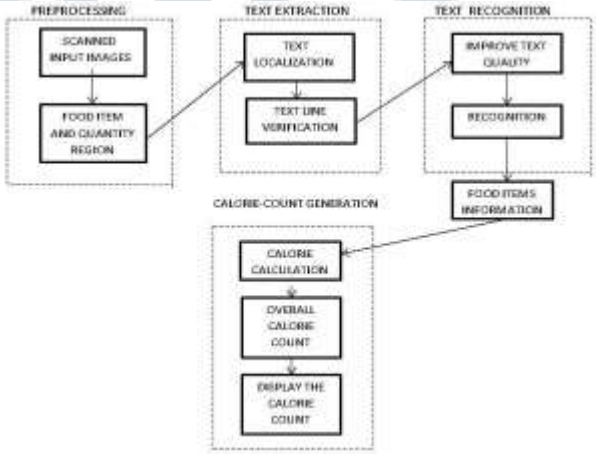

Fig.1 System design

IV. IMPLEMENTATION AND RESULTS

The proposed system is developed using Android studio and Firebase to detect text on the restaurant bills and show corresponding calories. Figure 2 shows the Application user interface. Also we have used Google vision API to perform the optical character recognition.

We need restaurant bills for recognizing the food items on the bill. For the purpose of evaluation, the bills from various restaurants have been collected and scanned using our application. One of such collected restaurant bill is shown in figure 3a. The bills are generally in the form of printed text. The color of the text in the bills will generally be in blue or black.

The first step is to create a database in Firebase console and to feed the different food items and their calories. The

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 6, Issue 11, November 2019

calories of various food items have been taken from Nutritionix [26] the world's largest verified nutrition database. Over 100+ health apps make 5M+ queries to nutritionix every month. We have also referenced MyFitnessPal [27] for calories.

The next step is to register the app with Firebase. The dependencies required for the system include both the Google Vision API as well as the Firebase. On opening the application, there is an icon, which when clicked takes the picture of the bill. The scanned text from the image is mapped with the Firebase and the corresponding calories of the food items are displayed in the same intent. Figure 3b shows the result of the sample restaurant bill.

On conducting the experiment on various restaurant bills we get accurate calorie count of the food consumed. The accuracy depends on the quality of the bill, the environment and the quality of the camera. The images of the restaurant bills that we collected for our experiment and the outcome of the experiment is shown in figure 4.

The proposed system faces certain difficulties while implementing the experiment. Some of the most important challenges faced by the system are:

Difficulties in the quality of the camera: Since our objective is to capture and detect text in the restaurant bill, the phone's camera must be in a condition such that it focuses the bill, and the text in it, thereby helping the Google vision.

Difficulties faced in the environment: Since the data on the restaurant bill is used to detect the calories, there rises a compulsion in setting the lights of the room an optimal one. Adequate lighting is required for the experiment to be done.

Difficulties in the Quality of the bill: Since data on the text is used to show the corresponding calories, the quality of the bill comes into play. Handwritten bills are seldom recognized by the API. Bad quality bills do not cooperate with extracting text.

Figure 5 and 6 shows the bills of different quality of the camera.

Fig.2 Application UI

(a) (b)

Fig.3 a) Restaurant bill b) Result

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 6, Issue 11, November 2019

Fig.4 Bills and Consumed Calories

(a) **(b)**

Fig.5 a) High clarity camera b) Low clarity camera

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 6, Issue 11, November 2019

(a) **(b)**

Fig 6 a) Faulty bill b) Ideal bill

V. CONCLUSION AND FUTURE WORK

The recent advancements in technology have always been beneficial to the mankind. This system takes a initiative in attainment of a goal that would help the people to live in healthy life style independently. In this paper, OCR is used to develop a system that recognizes the text on the restaurant bills .We have also calculated corresponding calorie count, thereby enlightening him/her on the calories consumed. Although we faced many challenges during our experiment, we can say with a reasonable level of confidence that above results helped to attain our proposed solution. Our future work is to include handwritten bills to detect and recognize the text on them and calculate the corresponding calories of the food consumed. We would also like to extend the functionality of detecting text on bills from videos.

REFERENCES

1. K. Jain, Y. Zhong Page segmentation using texture analysis Pattern Recognition, 29 (5), 1996, pp. 743-770
2. Y.Y. Tang, S.W. Lee, C.Y. Suen Automatic document processing: a survey Pattern Recognition, 29 (12), 1996, pp. 1931-1952
3. B. Yu, A.K. Jain, M. Mohiuddin, Address block location on complex mail pieces, Proceedings of International Conference on Document Analysis and Recognition, Vancouver, BC Canada, 1997, pp. 897–901.
4. Y. Cui, Q. Huang, Character extraction of license plates from video, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, 1997, pp. 502–507.
5. D.S. Kim, S.I. Chien, Automatic car license plate extraction using modified generalized symmetry transform and image warping, Proceedings of International Symposium on Industrial Electronics, Vol. 3, 2001, pp. 2022–2027.
6. H.J. Zhang, Y. Gong, S.W. Smoliar, S.Y. Tan, Automatic parsing of news video, Proceedings of IEEE Conference on Multimedia Computing and Systems, Boston, 1994, pp. 45–54.
7. J.C. Shim, C. Dorai, R. Bolle, Automatic text extraction from video for content-based annotation and retrieval, Proceedings of International Conference on Pattern Recognition, Vol. 1, Brisbane, Qld, Australia, 1998, pp. 618–620.
8. Ye Q, Doermann D. Text detection and recognition in imagery: A survey. IEEE transactions on pattern analysis and machine intelligence. 2015 ,37(7):1480-500.
9. Karatzas, S. Robles Mestre, J. Mas, F. Nourbakhsh, "ICDAR robust reading competition: Challenge 1: Reading text in born-digital images", Proc. IEEE Int. Conf. Doc. Anal. Recognit., 2011, pp. 1491-1496.
10. P. Shivakumara, A. Dutta, U. Pal, C. L. Tan, "A new method for handwritten scene text detection in video", Proc. IEEE Int. Conf. Front. Handwritten Recognit., 2010, pp. 387-392.
11. C. Strouthpoulos, N. Papamarkos, A.E. Atsalakis Text extraction in complex color document Pattern Recognition, 35 (8), 2002, pp. 1743-1758
12. Karaoglu S, Tao R, van Gemert JC, Gevers T. Con-text: Text detection for fine-grained object classification. IEEE transactions on image processing, 2017, 26(8), pp. 3965-80.
13. He T, Huang W, Qiao Y, Yao J. Text-attentional convolutional neural network for scene text detection. IEEE transactions on image processing, 2016, pp. 2529-41.

14. Chong Yu, Yonghong Song, Quan Meng, Yuanlin Zhang, Yang Liu, "Text detection and recognition in natural scene with edge analysis", Computer Vision IET9, 2015, no. 4, pp. 603-613.
15. Anand Mishra, Karteek Alahari, C.V. Jawahar, "An MRF model for Binarization of Natural Scene Texts", International Conference on Document Analysis and Recognition (ICDAR), 2011.
16. Udit Roy, Text Recognition and Retrieval in Natural Scene Images, Hyderabad:CVIT, International Institute of Information Technology, 2015.
17. L. Neumann, J. Matas, "Scene text localization and recognition with oriented stroke detection", Proc. IEEE Int. Conf. Comput. Vis., 2013, pp. 97-104.
18. J. L. Feild, "Improving text recognition in images of natural scenes", 2014.
19. J. J. Weinman, Z. Butler, D. Knoll, J. Feild, "Toward integrated scene text reading", IEEE Trans. Pattern Anal. Mach. Intell., vol. 3, no. 2, pp. 375-387, Feb. 2014.
20. X. C. Yin, X. Yin, K. Huang, H. Hao, "Robust text detection in natural scene images", IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, no. 5, pp. 970-983, May 2014.
21. J. J. Weinman, E. Learned-Miller, A. Hanson, "Scene text recognition using similarity and a lexicon with sparse belief propagation", IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 10, pp. 1733-1746, 2009.
22. Deshpande S, Shriram R. Real time text detection and recognition on hand held objects to assist blind people. In2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), 2016, pp. 1020-1024, IEEE.
23. How does OCR document scanning work?. Explain that stuff, 2012. http://www.explainthatstuff.com/how-ocr-works.html
24. R. Smith, "An Overview of the Tesseract OCR Engine," Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), vol. 2, pp. 629-633, 2007.
25. Tesseract-OCR, https://code.google.com/p/tesseract-ocr/
26. Nutritionix - Largest Verified Nutrition Database https://www.nutritionix.com/
27. MyFitnessPal - Calorie Chart, Nutrition Facts in Food https://www.myfitnesspal.com