Text and label reading using Raspberry Pi and Open CV

Shakthi Vel
shakthijambuvel@gmail.com
SRM Institute of Science and Technology, Chennai,
Tamil Nadu

W. Rohan
wrc971@gmail.com
SRM Institute of Science and Technology, Chennai,
Tamil Nadu

R. R. Raghavan
rrragh28@gmail.com
SRM Institute of Science and Technology, Chennai,
Tamil Nadu

Somdeep Choudhury
somdeepchoudhury@gmail.com
SRM Institute of Science and Technology, Chennai,
Tamil Nadu

ABSTRACT

In this digital Age with the help of our cognitive abilities, we can see, hear and sense various new technologies and can communicate with them but this is not the case for visually impaired people In order for them to communicate with the world they need assistive technology and adaptive devices. The key idea of the project involves helping the sight challenged people. The Proposal of the project involves capturing an image with the help of a camera, recognizing it and extracting the content in the image using various data algorithmic techniques. With the help of this model people with blindness can be more independent with confidence. The major step in making this project a reality is to obtain images from the real world and recognize them with little to no latency and convert them into audio files that can be played. This can be executed out with the help of Open CV and Raspberry pi the main usage of the latter is its portability and compatibility which can be achieved with the use of battery backup functionality which can also be used for forthcoming endeavours. The size of the Raspberry pi SOC permits the customer to hoist it all over and wield it.

Keywords— Optical Character Recognition (OCR), Open Computer Vision (Open CV), Raspberry Pi, Image recognizing, Portability

1. INTRODUCTION

Visually challenged of us have to handle a barrage of challenges to make sense of visual content with matching labels by using the systems available at their disposal now, merging concerns along midway, settlement, correctness, adaptability and vivacity. On board a shrewd structure that aids the sight challenged which adequately and flourishingly inspect through lithographed content. The overall aim of the plan in action involves a camcorder positioned utilitarian accessory which is effectively used by singleton to look through lithographed article pieces. The point is to execute a picture catching strategy in an introduced system in perspective on the Raspberry Pi board. The arrangement is encouraged by uniting with sight-impaired people, which enables an increasingly accommodating movement with the little framework. The camera present in this system is used as the input component to capture the picture for digitization [2]. The Optical Character Recognition algorithm (OCR) is used to scan and recognize the images.

The Open source Computer Vision (Open CV) libraries are used for acquiescing the character from the given scanned testimony. A greater piece of advancement worked for individuals with sight challenged and deprived of vision rely upon the two quintessential structure modules known as the Optical Character Recognition (OCR) programming and Text-to-Speech Engine (TTS) [1]. With the aid of the above technique mentioned above, we can use it for filtering report of lithographed content into apparatus ciphered measure of symbols in alphabetic order. The above-mentioned procedure that is Optical Character Recognition is a process in which it interfaces with a representative significance with components consisting of (alphabetical signs, portraits, and figures) along with the interpretation of a figure. The above-cited design is often portrayed in a mode of technique for deciphering examined archives or pictures of printed content into a format which can be accessed by a computer. The sight challenged who could not read a document, but need to connect with the contents of those documents can be helped with a useful but productive technology called OCR. Here with the help of the null electronic framework the photos and converted messages are paired up with zeros and ones in a computerized format. The process of converting an image from one format to zeros and one's format that is digital, here it correspondingly diminishes supplementary place. We all know that modifying and replicating a piece of content on a document which is lithographed on a sheet can be monotonous as well as industry raised. We know the fundamental purpose of the Optical Character Recognition is to search the alpha symbols on a ciphered sheet, but it is also predominantly employed for the transition of volumes along with a summary in the direction of through to an electronic

transcript. Apparatus elucidation, substance for talk and content burrowing of an occupied/analyzed sheet is only feasible only with the guidance of OCR in the administer systems. The final outcome of the content report is inclined to the yield gadgets which presents an alternative to the customer. The ultimate turnout of the contraption can be understood with either a handset which has been established with the Raspberry Pi 3 model 3 board or an audio yielding device where the amount of the content can be recorded boisterously.

2. EXISTING SYSTEM

A screen per user is a PC application that engages a sight hindered PC customer to perceive what's on the screen through talk. But the main drawback is that it can quintessentially scrutinize the content or text or a message on a screen but not the whole depiction of an image. Everything considered, as of now there is not an absolute approach for the structure to describe a drawing figure is conceivable to detect a portion on an image. With the employment of the above procedure, the device examines by taking advantage of the shock passkeys along with the above-mentioned apparatus which can unravel the main achievement of the gadget through the verse appearing on the screen. All this can be achieved by a tiny procedure of pressing an enter [4] which is a vital step in order for the customer to interact with the above-voiced interface with a peculiar art of alphanumeric symbols. Also if the customer when asserting every set of alphanumeric symbols came to know about that his confirmation of the depicted pattern was erroneous it would have a pulverizing effect on him and he would ultimately eradicate his entire work for the proper execution of the apparatus. But we know that being a customer is just a single instance of the points of interest that a deprived of sight PC habitué get from a screen per user.

Here the deprived of sight people simply check out a screen per user scrutinizing the substance appearing in front of a screen, so they don't know for certain whether the given outcome of the apparatus is entirely accurate and authenticate so, in other words, the client who employs this apparatus doesn't know the exact logography of a peculiar tête-à-tête notably when the discussion is not ordinary identical to the restorative conditions, etc. [2]. So without raising a question, they are able to cause the client to scrutinize the alphanumeric symbol by symbol but there is a humongous repercussion in it that is when a customer hears every sentence or a word through the gear it is not possible for them to even have the faintest thought regarding the logography. As expected it turns out to be a tiresome and gruelling experience. Screen per users use a PC resonating tone and a couple of individuals find this awfully debilitating. As of now there a couple of associations which are performing the finest job making the synthesizers, this reflects a mortal on perceiving a pronouncement [6]; which sounds like an ideal tone, but this is not the case here as we have some colossal overhauls which date many years back from now, these gadgets are yet from achieving their ultimate goals as they fell short before accomplishing them. Some of the existing methods are mentioned below

Table 1: Comparison of methods used for Text Detection and Recognition

Methods	Features of methods
Stepwise Methods	1. Separate detection and recognition modules 2. Divided into localization, classification, segmentation, and recognition stages 3. Suitable for detection of a large number of words 4. Less computation cost, but complexity increases due to more steps
Integrated Methods	1. Detection and recognition modules are not separate 2. Can avoid segmentation or replace it with word recognition 3. Suitable for identifying specific words from an image that is small lexicon 4. Increase in lexicon size makes recognition difficult

3. PROPOSED SYSTEM

So as To prevail with regards to managing the issues in the current framework we have built up a venture for visually impaired People utilizing OCR in Open CV. The proposed framework is to help dazzle people to peruse content from testing example and foundation to read the report. The principal focus of our framework is to dissect the content in the archives. Right off the bat, the item picture is caught by utilizing a camera module which is implanted inside Raspberry Pi and is trailed by the picture preparing [4]. To deal with a mechanized framework, which looks at an archive and peruse its texture to the person on a snap of a catch. The lyric is conveyed with the aid of mouthpiece which would push the person to readout the substance in the sweep archive. Our framework causes the visually impaired individuals to read without expanding on multiple parameters which can be used in the process. The proposed framework is expected to address the issues featured in the issue proclamation. This new framework will contain Raspberry Pi alongside a camera module and headphones or some other speaker framework as its yield while its camera will fill in as its info. This is blend with programming that with the assistance of Tesseract OCR will most likely proselyte any picture that it is being pointed to an ongoing comprehend the content that is available in the picture and converts that content which it finds into sound. Furthermore, this sound will be given to the client by the utilization of headphones that the client will have and hence help individuals
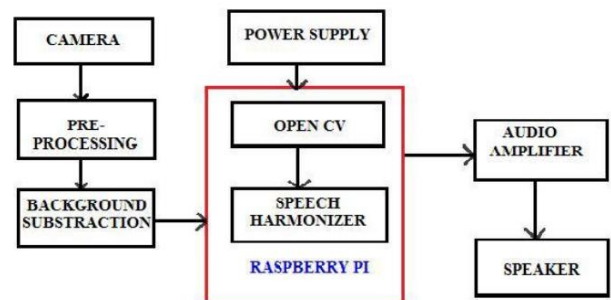

Fig. 1: Proposed system

This can be executed out with the assistance of Open CV and Raspberry pi the principal use of the last is its movability and similarity which can be accomplished with the utilization of battery reinforcement usefulness which can likewise be utilized for future undertakings. The measure of the Raspberry pi soc grants the client to convey it anyplace and use it by the consolidated intensity of these equipment and programming segments the pointed result of a minimized machine which is going to enable the outwardly moved move to advance in their regular day to day existence without breaking a sweat and more certainty.

4. DIGITAL IMAGE PROCESSING

The ID of items in a picture would presumably begin with picture preparing systems, for example, noise explosion, trailed by include extraction to find an edge, commune and potentially regions with the specific exterior. The main objective is to remove accumulations of these shapes as single items, for

example, vehicles cm a road, boxes on transit lines or harmful objects on a magnifying lens slide. Another problem is an Artificial intelligence issue is that an article can seem generally differ when viewing from various edges or under various lighting. Another issue is choosing what highlights have a place with what item and which are infrastructure or shades and so on. The human visual system plays out the assignments for maximum part unwittingly yet a PC requires capable programming and bunches of handling capacity to approach human execution. Controlling information as a picture through a few conceivable methods. A picture is nominally translated as a 2-D exhibit of splendour esteems and is larger recognizably spoken 10 10 by if, 4.h examples as those or a pictorial print. Drift. The TV screen, or motion picture cover. A picture can be prepared -visually or carefully with a PC.

To carefully measure a picture_ it 15; important to diminish a picture to a progression of 1 1uies that cart is controlled by the PC. Each number speaking to the brilliance estimation of the - picture. At a specific area is known as an image component. Or pixel. A regular pixeled picture may have 512*512 pixels, albeit a lot bigger pictures are getting to be normal. When the picture has been pixeled, there are three fundamental activities that are executed on it in the PC. For a marked activity, pixel esteem in the yield picture relies upon solitary pixel esteem in the information picture. For nearby tasks_ a few neighbouring pixels in the information picture decide the estimation of an iced l picture pixel. In a worldwide activity, the majority of the information picture pixels add to a yield picture pixel esteem. These activities, taken separately or in blend_ are the methods by which the picture is upgraded, reestablished. Or packed.

A picture is improved when it is adjusted so the data it contains is all the more unmistakably obvious, yet upgrade can likewise incorporate making the picture all the more outwardly engaging. A model is clamour smoothing. To smooth an uproarious picture, middle separating can be connected with a 3 x 3-pixel window. This implies the estimation of each pixel in the loud picture is recorded, alongside the estimations of its closest eight neighbours. These time numbers arc then arranged by size, and the middle is chosen as the incentive for the pixel in the new picture. As the 3 3 windows are moved one pixel at any given moment over the boisterous picture, the sifted picture is shaped. Another case of the upgrade is differentiated control. Where every pixel's an incentive in the new picture depends entirely on that pixel's an incentive in the old picture: at the end of the day. This is a point activity.
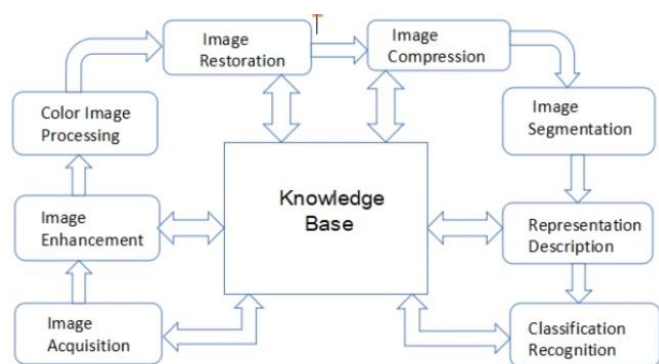

Fig. 2: DIP Architecture

4.1 Acquiring input picture
The underlying stage where a gadget is carried over the lithographed sheet was an innate camcorder catches the photos of the substance. The True idea of the image caught will be

immensely huge so as to have brisk and explicit acknowledgement as a result of the superior quality of the camera. Computerized imaging or digital picture acquisition is the production of a digitally ciphered representation of the ocular attributes of an item. Computerized imaging was created during the 1970s, to a great extent to maintain a strategic distance from the viable shortcomings of film cameras, for logical, experimental and military missions. As computerized innovation wound up less expensive in later decades, it substituted the old film strategies for some reasons.

4.2 Pre-Processing
In this model there is the Slant Correction, Linearization and expulsion of cacophony were completed where the snatched depicted drawing is reviewed for skewing. We know that there is some conceivable outcome of the skewed depiction being reviewed by either right or left initiation [5]. Here the image is bright and paired with zeros one. The skew's unmistakable confirmation audits for a point of presentation in the middle of plus or minus fifteen angles of intensity and at whatever point apparent then a little picture change angle is fulfilled till the boundaries compose with the certifiable dimension focus point, which passes on a skew amended image[9]. The typical uproar amidst getting or due with meagre pixel nature of the side must be tidied up ahead of extra planning.
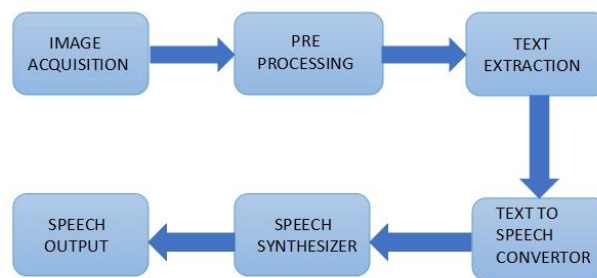

Fig. 3: Flow of process

4.3 Segmentation
The disorientation free depiction is after the pre-planning where the image is ceased into to the Segmentation compose. Where a system that confines a filtered picture of accumulation to hieroglyph in the direction of through to auxiliary images of respective picture (signs). The assessed picture is digitized and paired with zeros and ones together with bury boundary scattering is checked. Later the cover edge spaces are seen, by then the picture is separated into sets of segments over the cover holes. The boundaries in the segments are broke down for matched space spot over passing on a dimension plane concerning establishment. The scatter diagram of the checked picture is employed to see the birth of the dimension boundary. By then the boundaries are checked straight up for erect space get together. Here the scatter diagrams are handled to see the thickness of the letters. By then the letters are isolated into unique aspects taking advantage of tone girth figuring.

4.4 Feature Extraction
In this process, the features of the depicted image glyph are extracted where the picture is examined over a respective personal. In this depicted illustration the thickness and stature of the symbol, the horizontal and vertical boundaries which are in current form (both brief and protracted), the figure of oriented arcs with both vertical and horizontal shapes, numbers of spheres present, the picture is centred, location of features are diversified and finally the total number of pixels in distinct sectors are delineated by a character glyph through various consecutive facets.
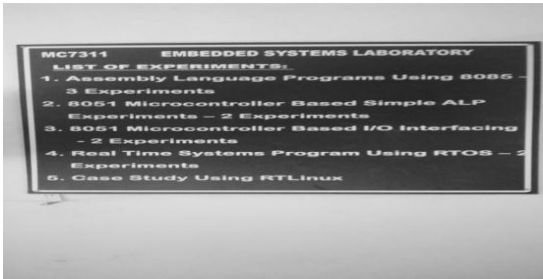
Fig. 4: Input snatched a picture
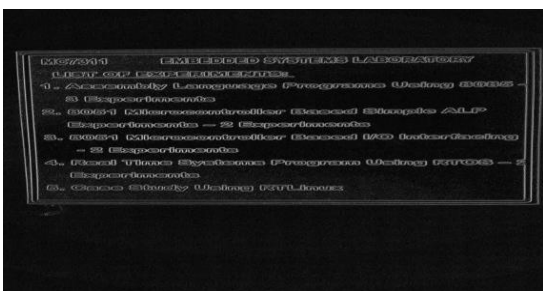

Fig. 5: Picture in the gray scale format

Fig. 6: Threshold image

4.5 Text Extraction

The current moment we anticipate text of letters is exhibited in the depicted picture where the image is examined with the help of OCR engines. Thus by employing the OCR engine identified symbols of letters is isolated. As we can see the text extraction can be only performed with the assistance of Optical Character Recognition. This Optical Character Recognition is often called or abbreviated into OCR, which is a form of modifying the depicted picture in automated or computerized form.it can also convert a lithographed passage into an instrument ciphered passage and also changeover of a portrait of an archive. The essential usage of OCR is for info access of engraved sheet with data records. It can modify diversified items which involve various day to day utilities such as the passport, pan card, bank passbook, demand drafts, voter id, driving license and other static data in the form of a printout or in a document. OCR is used in a number of research such as pattern recognition, A.I and computer vision.

Fig. 7: OCR Output

Fig. 8: Spell Corrected Output

5. SOFTWARE SPECIFICATION

The operating system which for the executed project is called the method of design is created by taking advantage of the computer-aided learning a language in python. The design method known as the algorithm has a variety of OpenCV library functions in it. As we know that just like its name OpenCV is an open sourced library in computer vision. The OpenCV library is formed with the use of C and C++ which runs under the hood of operating systems such as Kali Linux, various versions of Windows and Mac OS X. OpenCV was coined and proposed for the sole purpose of Image Processing with better computational efficiency and also with an effective spotlight on ceaseless applications.

Open CV is created in improved C and can likewise exploit its multi-centre processors. The Open CV library contains in an overabundance of five hundred capacities navigate various locales in vision, including producing plant thing examination, User Interface, change in the camcorder, image restoration, the conception of stereo vision and automated machinery. Since PC vision and AI much of the time go as an indistinguishable unit, Open CV in like manner contains a full, Machine Learning Library (MLL) in a full around valuable to assist the optical character recognition and text to speech engine .the most important aspect of these features is we need to present them predefined libraries.

6. HARDWARE SPECIFICATION

The Model of Raspberry utilized is the latest and more recent third gen Raspberry pi called the raspberry pi 3. It is an amazing pocket charge card estimated framework on a chip which can be employed for a lot of computer functions and usurp is the first gen model known as Raspberry Pi model B [12]. While keeping up the fame load up organization the current third-gen Raspberry Pi 3 Model B delivers a more dominant processor with multi-core option which is multitude time quicker and powerful than the first gen original Raspberry Pi. Furthermore, it likewise has a remote computer network (LAN) and Bluetooth availability for associating gadgets hence structure a system, making it the perfect answer for amazing associated plans.

7. RESULTS

The final outcomes of the given methodology depicted over are shown with the help of images underneath. Fig.4 In the particular framework the picture of the source is the Red Green Blue picture that may be provided by the clients for acquiring the shape picture and also the vehicles included in yield screen. Accompanying code may be utilized to size the yield screen automatically, exhibits the information picture which found using the camera, figure 5 the grayscale picture can be utilized to show the articles in the configuration of highly contrasting. In this framework the yield will be shown by the grayscale picture in the wake of getting the source picture just, in the light of the fact the source of the picture can be changed over into the grayscale picture format and gives the dim scale changed over picture Fig.6 demonstrates the edge distinguished pre-prepared sweep picture which is fed into Optical Character Recognition motor in order to evacuate the substance found in the image. The limit splendour or differentiation of the grayscale picture. In this Framework, we can change the greyscale image into edge detection picture. With the help of this process, we can create a binary image. Fig.7 shows the yield of the Optical Character Recognition motor. Fig 8 exhibits the return of the spell corrector which is used to adjust the inaccurately augured letters from the OCR motor.
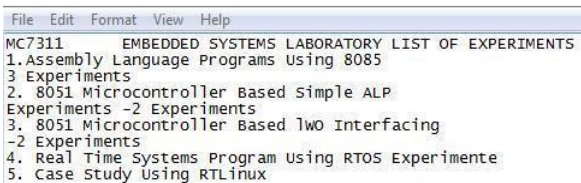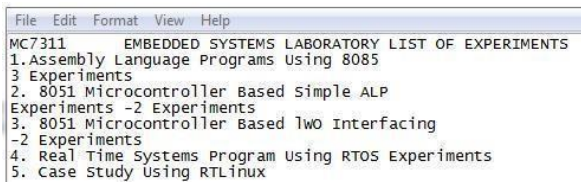
8. CONCLUSION AND FUTURE WORK

Our framework makes a point to scrutinize content present in the picture hence by helping sight tested people. Next is the Pre-preparing part which guarantees a productive forefront extraction, which has the required measure of the content district which is to be examined. In any case, the framework couldn't ready to separate the closer view of the picture when it has an intricate foundation. An Enhanced calculation for foundation subtraction can chop down the impacts of complex foundations where the separated content is then sent yet yield is defective. Subsequent to getting the revised yield we send it to the OCR motor which gives a discourse yield. Along these lines, by giving a battery reinforcement to the raspberry pi we can achieve the primary focus of the ideal undertaking of convey ability. The future work will be focused on advancing towards a proficient convenient item that can extricate the content from any picture in this manner by engaging the sight tested to peruse content present on the items, standards, books and so on.

9. REFERENCES

[1] Faiz Alotaibi, Muhamad Taufik Abdullah, "Optical Character Recognition for Quranic image similarity matching," IEEE, 2017.

[2] Shalini Sonth, Jagadish S kallimai, "OCR based facilitator for the visually challenged,"IEEE, 2017.

[3] Michael D.kim and Jun Ueda, "Dynamics-based motion de-blurring improves the performance of Optical Character Recognition during fast scanning of a robotic eye,"IEEE, 2017.

[4] Sanjana, Monisha, "OCR based automatic book reader for the visually impaired using raspberry pi,"IEEE, 2016.

[5] Velumurugan D. et al, "Hardware Implementation of Smart Reader for Visually Impaired using Raspberry Pi", IJAREEIE, March 2016.

[6] N. Venkata Rao et al, "Optical Character Technique Recognition Algorithms, Journal of Theoretical and Applied Information Technology", 20th January 2016.

[7] K. Nirmala Kumari et al, "Image Text to Speech Conversion using OCR Technique in Raspberry Pi", IJAREEIE, May 2016.

[8] Seiichi Uchida, Koichi Kise, "OCR fonts revisited for camera- based character recognition," IEEE, 2016.

[9] D. Velmurugan et al, "A Smart Reader for Visually Impaired People Using Raspberry PI", IJSEC, Vol 6, pp 2997-3001.

[10] Nagaraja L. et al, "Vision Based Text Recognition using Raspberry Pi, International Journal of Computer Applications", NCPSIA 2015.

[11] K. Kwang in Kim, Keechul Jung, "Texture-based approach for text detection in images," IEEE, 2013.

[12] https://raspberry.piaustralia.com.au/raspberry-pi- 3-model-b

[13] Study and Implementation of Various Morphology Based Image Contrast Enhancement Techniques. Suman Thapar, Shevani Garg, CSE department, Ludhiana College of Engineering and Technology, Punjab.

[14] Image Compression Using the Discrete Cosine Transform. Andrew B. Watson, NASA Ames research Center.

[15] Short-term connection problems might affect your updatescode.google.com