Recognition of Tablet using Blister Strip for Visually Impaired using SIFT Algorithm

R Shashidhar1*, V Sahana2, Sudeshna Chakraborty3, S B Puneeth4, M Roopa5

1 Assistant Professor, Department of Electronics and Communication Engineering, JSS Science and Technology University, Sri Jayachamarajendra College of Engineering, Mysuru, India
2 Assistant Professor, Department of Information Science and Engineering, JSS Academy of Technical Education, Bengaluru, India
3 Associate Professor, Department of Computer Science and Engineering, School of Engineering and Technology, Sharda University, Greater Noida, Uttar Pradesh, India
4 Assistant Professor, Department of Electronics and Communication Engineering, Presidency University, Bengaluru, India
5 Professor, Department of Electronics and Communication Engineering, Dayananda Sagar College of Engineering, Bengaluru, India

Abstract

Objectives: To create a tablet recognition system that recognizes and gives an audio output of the name of the tablet so that the visually impaired person may recognize it. Methods: In this, to create a user friendly system, such that a blind person does not require the help of another person to use it. These systems give medical assistive instrument to take the drug at the accurate time as per the doctor prescription and removes dependency on others for the visually challenged people. Various models were technologically advanced to compact with misidentification of remedies but are incompetent of formative exact pill picked by the person. Findings: The name of the medicine is obtained from the image. Then the name of the medicine is converted from text to speech using Google text to speech (gTTS). The proposed system will be helpful for visually challenged people, SQLite Database management system has been used along with Scale Invariant Feature Transform algorithm (SIFT) to identify and pronounce local features in the image. The image of the medicine given by the user is matched with the images in the database. Novelty: Developed a user-friendly voice based tablet recognition system for visually impaired people and this system is fast and accurate.

Keywords: SIFT algorithm; medicine identification; visually impaired people; gTTS; Tablet Recognition

1 Introduction

Due to high percentage of elderly people in the society, the regular age of the population is rising, human body functions are dwindling, and average visual acuity is deteriorating year after year. Affording to a statistic published in October 2017, there are nearly 285 million visually disabled people in the world, with 140 million of them being aged people over the age of 50 and 110 million of them suffering from multiple chronic
diseases. It is well understood that a person's physiology deteriorates as they get older. These 110 million vulnerable visually disabled aged people are more likely to take the wrong medications or fail to take their medications if they take several medications. The computer vision group is currently working on developing assistive devices for visually impaired people. Such a device offers a medical assistive service that allows visually disabled persons to take the right medication at the right time, as recommended by their doctor, and to live safely in their daily activities.

In this paper, the authors have proposed a device that will be capable to recognize a tablet by a image of the tablet's blister pack and the name of the tablet is voiced out. SQLite Database management system has been employed in which database of 10 medicines that are commonly used by patients was saved the same was used for testing. SQLite is a client-server database system, not a client-server database engine. Rather, it is incorporated into the final product. It can be modified according to the wishes of the patient. Later image preprocessing was performed, with the goal of removing unwanted distortions or enhancing those image features that are necessary for subsequent processing. To give them a smoother look, picture sharpening was used. The SIFT algorithm is then used to detect and identify local image features. Feature point identification, feature point localization, orientation assignment, and feature descriptor generation are the four stages of the SIFT algorithm.

The image of the medicine given by the user is matched with the pictures in the database with the help of this algorithm. The name of the medicine is obtained from the image. Then the name of the medicine is converted from text into speech using Google text to speech (gTTS). GTTS is a text to speech API that can convert text into speech. By the help of this GTTS API that is available as open source in python library, this was used to spell the name of the medicine and it can be voiced out in 30+ languages. Google uses WaveNet, a deep generative network for generating the raw audio. The output will be the audio of the name of the medicine. Main objective of the work is to create a tablet recognition system that recognizes and gives an audio output of the name of the tablet so that the visually impaired person may recognize it. To create a user friendly system such that a blind person does not require the help of another person to use it. To create a fast and accurate system that recognizes pills with speed and accuracy, regardless of shape size and color. Suntronsuk et al proposed an Ostu's threshold method for binarisation of the imprint area of the pill image. Optical character recognition method used to extract the text from images (1). Bose, Gitika Pabari & Tejit et al approach to identifying pills using primarily opencv algorithms. To determine the shape of a pill, they used k mean clustering and then found the contours in the image. With this they were able to figure out the shape of the image. To determine the name of the corresponding colors, k-nearest neighbors was employed by comparing the distance between the cluster centers and every RGB value observed for each color and choosing the majority of the 5 closest colors (2). According to Tanjina piash Proma et al, data sets are divided based on the amount of colors and imprinted texts on the pills. Color information was extracted from the pill image by segmenting the pill area, and then some statistical measurements for likelihood scatterings produced from the image histograms were calculated. The possible text area is spotted for error-free in text recognition. The orientation images from the NLM RxIMAGE database were used to create high-quality image data. (3)

The aim of this work is to reduce the issue of patients forgetting about their drugs and not knowing how much to take. It's a hybrid of physical and digital reminders that can benefit people of all ages, but it's particularly useful for seniors who forget to take their medications. One disadvantage was that they did not allow for the storage of the original prescription. (4). To assistance chronic victim in correctly taking many drugs and avoiding taking the incorrect medications, and to include other drug-related functionalities such as medication reminders, medication details, and so on. An Android app, a deep learning training server and a cloud based organization framework make up the scheme. The proposed system can currently identify eight different medicines (5). For visually disabled individuals, the survey suggests a deep learning based wearable medicine gratitude system. A couple of wearable smart spectacles, a wearable waist-mounted drug pill recognition kit, a mobile App and a cloud based organization framework makes up the planned system. To stop taking the wrong medications, the proposed method employs deep learning technology to recognize opioid tablets. (6)

The number of processes which starts with the thresholding applied to the input query pill image for extraction of the shape feature vector and generation of mask images. The extracted shape feature vector is used for shape recognition through a trained neural network. For pill imprint extraction, a modified stroke width transforms (MSWT) and two-step sampling is applied (7). Huynh et al worked on Using the polar transform and neural networks, a new imprinted tablet recognition algorithm has been created. This algorithm determines whether or not examined tablets have the same imprinted mark as a reference tablet. The algorithm is divided into two parts: neural network training and imprinted tablet recognition (8). S. Kang et al investigated the low accuracy of tablet detection due to the effect of illumination caused by the source of light. As a result, removing the effect of illumination will increase pill recognition efficiency even more. They suggested an algorithm for removing lighting effects from pill and tablet recognition that is appropriate for preprocessing. Since the context of the captured pill images was so broad, they used the class activation map (CAM), a weakly controlled localization technique, to approximate the direction of illumination and find the location of the pills (9).
Y. ou et al. suggested that deep learning and Convolution Neural Network (CNN) can be used to replace repetitive screening work and increase drug screening reliability. The R-CNN series has a lot of success in classification and identification tasks. R-CNN offers Selective Search, which generates 1000-2000 proposal candidates and uses a deep network to extract features. Feature Pyramid Networks (FPN) employs CNN’s feature hierarchy to construct a feature pyramid that combines semantics from high to low. For object detection, various pyramid levels are used, each with a different scale and ratio.

Suntronsuk and his colleagues worked on Imprints on pills typically provide vital details such as brand names and dosages. It would be advantageous to be capable to remove texts from imprints automatically. The results can be used to add pill information to existing pill databases or to scan for it using pill photos. Two phase Sampling Distance Sets are proposed as an imprint descriptor in this study (TSDS). This definition was described as a vector, which included details about the pill’s shape, color, and text imprints. They used K-means clustering with K=3 on the picture since most pills only have two foremost colors. The backdrop was made up of the clusters with the greatest pixels on the border. The context cluster was stripped of all pixels. Further, clever edge detection was performed to locate the pill’s edges. Edge chart was used to find a bounding box that encompassed the major outline, which was the pill. They used tessera for OCR after that. For segmenting the pill image, Wang et al used the Modified Stroke Width and TSDS for describing the shape of the pill. Imprint features on the gradient magnitude picture include the Scale Invariant Feature Transform and Multi-scale Local Binary Pattern. Both of these approaches, on the other hand, were tested on pill images with a consistent context and noise levels, allowing for perfect pill segmentation. Top N classification accuracy and Mean Average Precision (MAP) score were used to evaluate the proposed classifiers’ results.

Krutarth Majithia and colleagues suggested a camera grounded assistive text analysis frame to assist blind people in reading add text labels and product wrapping from handheld items in their everyday lives. In an Ad boost model, they projected a novel text localization algorithm based on learning gradient features of stroke orientations and distributions of edge pixels. Optical character recognition programmed recognizes the text characters in the localized text regions as binaries. ICDAR-2003 and ICDAR-2011 Robust Reading Data sets, the proposed text localization algorithm is quantitatively evaluated. To identify if there are pills in the section, Htsai et al use an infrared photo interrupter. Each compartment has three sets of infrared emission LEDs and photo-detector receiver sensing circuits to improve sensitivity. Serial peripheral interface (SPI) is the communication protocol among Arduino and marginal devices. Serial clock, master in slave out, master out slave in CS, DC, and RESET are the six signal lines that make up SPI. The hardware will analyses the pill and shows the pill’s name on a 2.4-inch display.

Nijiya jabin Najeeb et al suggested smart pill box, which a smart box is built on the internet of things that keeps track of a patient’s wellbeing and medicine. Since this is a smart package, it can accommodate more than one medication at a time, obviating the need for separate bottles for each medicine. This box costs less than 1500 dollars, making the machine more accessible to the general public. The survey revealed that they have also created app for the tablet identification and keeping count about the number of tablets taken per day and such things. This app can be provided to patient or caretaker for monitoring tablet intake. This app also provides facility of automatically booking doctor appointment on scheduled date. Soyenong Lee et al approaches a new method in this method needs a picture to be taken, but a label is also printed. A picture of the tablet that has to be recognized in this case is taken with a built-in camera. Furthermore, without a smartphone, a minimum price braille embosser that connects to a smartphone via Bluetooth will print the classification results as a braille mark for future identification. Therefore, there is no dependency on the device all the time. As most pharmaceutical companies do not have braille printed labels, this project will be useful for blind people as a more permanent solution for recognizing their tablets. To initially recognize the tablet, a CNN based model is used.

Mi Zhang et al proposed system and method is provided that utilizes deep learning to identify subject objects such as unknown pills. An image of a pill may be captured and subsequently processed using deep learning models to identify the pill. The deep learning models is optimized to have a small footprint (in terms of computational and memory re- sources). By using Deep learning, they were able to overcome difficulty faced in increasing the accuracy of the system (different angle, blur, etc). Wan Jung Chang et al focuses more on the elderly people. To solve the problem of the elderly people is failing to recognize the pills, instead of having a user take a photo, or have a built-in camera take a photo, glasses are being using instead. This product is aimed at the elderly, particularly above the age of 50. The accuracy of this device is around 95%. The major drawback of the aforementioned approach is, the complexity of the circuit involved in the tablet’s identification. As well the images taken at different angle affected the result, adding the discomfort made by the smart glass is an additional shortcoming. Snigdha kesh et al senses and understands text in the environment and converts it into speech to assist blind people with daily tasks. All of this is done by a mobile phone and Android program. It was design to start the program automatically when the blind person shakes the phone. This method takes a picture of something and extracts the text from it. The written text information will be converted to a voice message by the system. This program can also be used to recognize pills. The application will notify a blind person when medication is available. The system would take a picture of the drug and mark the name of it. Following that, the machine will issue a voice command containing the name of the drug.
The approach taken by Somchart Chokchaitam et al, as seen in this survey, is to look at pill color. Pill color is used as one of the key features for pill identification; however, the majority of pills are white or cream in color. Since pill color is sensitive to luminance strength, it's hard to classify two alike looking pills. When the luminance intensity is increased, however, it may cause problems. To compensate, this paper proposes RGB reimbursement constructed on context shadow subtraction. This method of recognition is particularly useful for distinguishing between pills of the same color. Mateus A. Vieira Neto suggested that, in addition to color, the pill's form be taken into account. To distinguish pills based on shape and color, a pill feature extractor was proposed (CoforDes). It was possible to achieve accuracy of over 99%. It was also discovered that, in addition to having extremely high precision, it also had extremely high processing speeds, making it suitable for real-time applications. This method converts images to HSV, extracts shape and color attributes, and compares them to a database. Palenychka and colleagues Pill recognition can also be done through computer vision. The main goal of this work is to remove or reduce the limitations of current computer vision-based drug verification methods. Another goal is to provide a solid algorithmic foundation for a low-cost assistive device that can be used in small pharmacies, long term care facilities, or by individuals. The following components make up the proposed framework: a pill scanner, a pill confirmation server, a database of drug pill information, a patient organization method, and a nurse's graphical interface with a touch screen display. The accuracy obtained would be around 93%. In the recent past, IoT is used extensively for monitoring in most of the systems, this approach helps the maintenance by reducing the need of hardcopies of the report and it can be effectively used by the community with challenge in eye sight.

The structure of the manuscript is as follows. Methodology of the proposed approach is discussed in Section II. Discussion of the results of the proposed approach, applications and comparison with previous work is mentioned in Section III. At last in Section IV, the proposed work is concluded with the future scope.

2 Methodology

To implement this, it was necessary to first create a database to store the name and the image of the medicine. This is done with the help of SQ Lite Database management System. SQ Lite is not a client server database engine. Rather, it is embedded into the end program. Currently, information of 10 medicines has been stored for testing. It can be changed based on the user's prescription.

The medicine (that must be identified) is first preprocessed by sharpening the image. To sharpen the input image unsharpened filter was used. The unsharp filter is a simple sharpening mask that improves edges by subtracting an unsharp, variety of an image from the original image. For christening edges, the unsharp filtering method is widely used in the photographic and printing productions. Here, unsharpening was achieved with Open CV library. The main goal of this pre-processing is to eliminate unnecessary distortions or to improve certain image features that are critical for subsequent processing. To find the key points in the image, the SIFT Algorithm is applied to the preprocessed input image. SIFT is an image feature detection algorithm that senses and defines local features. The attributes of the input image are then mapped against all of the images in the database. The following equations can be used to describe SIFT.

\[
D(x, y, \sigma) = (x, y, k_i \sigma) - L(x, y, k_j \sigma)
\]

\[
L(x, y, k\sigma) = G(x, y, K\sigma) * I(x, y)
\]

Where D is the (DoGs) difference of Gaussians and L is the convolution of the original image I. Key points of sift are taken as maxima/minima of the DoGs. If the input image and the images present in the database have similarities, additionally, number of good common points between them was checked. If the number of good points is above the threshold, then the name of the medicine is then sent to a Google Text to Speech API, which then converts the name of the medicine to audio signal. This audio signal is then sent to a speaker. This way the blind person can identify the name of the medicine.

2.1 Flowchart

Below is the flow chart of the proposed work. Figure 1 shows the flowchart.

- Step 1: Feed an input image (tablet image) to Preprocessing.
- Step 2: output of the Preprocessing feed to SIFT Algorithm.
- Step 3: Database data send to SIFT Algorithm.
- Step 4: SIFT algorithm check with the input image and database image and send a decision to decision maker.
Step 5: If decision is yes means the medicine is matched and algorithm decided that is correct medicine that data send to GTTS API and after that name of the medicine is in the form audio.

Step 6: If Not matched the input and database image then the decision is wrong. That decision is send to GTTS API and GTTS API gives the output in the form of audio.

3 Results and Discussion

In this section shows the result of our work, Figure 2 shows the input image of the medicine and sent this image for preprocessing. Figure 3 shows the output of the preprocessing. Figure 4 shows the SIFT algorithm applied and Figure 5 shows the comparison of the input image and database image. Figure 6 snapshot of the proposed work.

3.1. Preprocessing Step

3.2 Results after preprocessing

Our result also tells the name of the tablet using text to speech using gTTS and if there is no match for tablet in database, then a beep sound can be heard to notify to a visually challenged individual.
Fig 2. Input Image for preprocessing (sharpening)

Fig 3. Preprocessed image

Fig 4. Image given as input for SIFT algorithm to compare

Fig 5. Database base image
3.3 Comparison table

Here, authors compare proposed work with existing work (23). Compare the proposed work based on different parameters. Figure 7 shows the input image was fed to the algorithm in the proposed work. (23) is considered as the major base paper or as reference paper for the proposed work, as it was published in the near past.

| Parameters             | [24] Proposed work |
|------------------------|--------------------|
| Accuracy               | 82%                |
| Method used            | CNN                |
| Tablet recognition technique | OCR Shape of pill |

4 Conclusion

Due to lack of vision, blind people may take different medicine instead of taking intended one and also, they can mix up medicine during buying it. These kinds of problems should be reduced because intake of medicine without its need can cause side effects and because the individual fails to take the intended medicine, resulting in problems that might attention.

Therefore, the authors have come up with the proposed work that can solve aforementioned problems and by employing the proposed approach, accuracy of tablet blisters pack detection is increased when it is compared with the previously used method (OCR). In future work, tablets blisters database has to be increased to a greater extent. Also, a feature of adding prescription would be included, so that the prescribed tablet is taken accordingly and include feature to alert when the same tablet is detected within short interval of time, to notify them that tablet is already taken. As well efforts will be made to identify tablets out of blister packs so that the algorithm runs more effectively.
References

1) Suntronsuk S, Ratanotayanon S. Pill image binarization for detecting text imprints. 13th International Joint Conference on Computer Science and Software Engineering (ICCSSE). 2016;p. 1–6. Available from: 10.1109/ICCSSE.2016.7748859.
2) Bose G, Pabari T, editors. Pill Detection & Prescription Analysis. 2019. Available from: https://www.researchgate.net/publication/340528010_Pill_Detection_Prescription_Analysis.
3) Prama TP, Hossan MZ, Amin MA. Medicine Recognition from Colors and Text. In: Proceedings of the 2019 3rd International Conference on Graphics and Signal Processing (ICGSP ’19). Association for Computing Machinery. 2019;p. 39–43.
4) Zanjal SV, Talmage GR. Medicine Reminder and Monitoring System for Secure Health Using IOT. Procedia Computer Science. 2016;78:471–476. Available from: https://dx.doi.org/10.1016/j.procs.2016.02.090.
5) Chang WJ, Chen LB, Hsu CH, Lin CP, Yang TC. A Deep Learning-Based Intelligent Medicine Recognition System for Chronic Patients. IEEE Access. 2019;7:44441–44458. Available from: https://dx.doi.org/10.1109/access.2019.2908843.
6) Hsu WCYXYHCGZZYSKTHYCH, Chen LB, Chen MC. A Deep Learning Based Wearable Medicine Recognition System for Visually Impaired People. 2019 IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS). 2019;p. 207–208.
7) Yu J, Chen Z, ichiro Kamata S, Yang J. Accurate system for automatic pill recognition using imprint information. IET Image Processing. 2015;9. Available from: 10.1049/iet-ipr.2014.1007.
8) Ou YY, Tsai AC, Wang JF, Lin J. Automatic Drug Pills Detection based on Convolution Neural Network. Suntronsuk S, Ratanotayanon S. Pill image binarization for detecting text imprints.
9) Mateus A, Neto V, ao WMJ, Souza PPD, Filho R, De AW, et al. CoforDes: An Invariant Feature Extractor for the Drug Pill Identification. Kesh S, Ananthanagu U. Text Recognition and Medicine Identification by Visually Impaired People.
10) Ou YY, Tsai AC, Zhou XP, Wang JF. Automatic drug pills detection. IEEE Access;2020(1):9–17.