Proceedings of the Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)
IEEE Xplore Part Number: CFP21OSV-ART; ISBN: 978-1-6654-2642-8

Smart Glasses for Visually Impaired Using Image Processing Techniques

P.N.Karthikayan
Veltech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology,
Avadi,Chennai
pnkarthikayan@veltech.edu.in

R.Pushpakumar
Veltech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology,
Avadi,Chennai
pushpakumar@veltech.edu.in

2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) | 978-1-6654-2642-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/I-SMAC52330.2021.9640715

***Abstract*— In this world, approximately 300 million are visually impaired, smart glasses helps the people who are Visually Impaired to enhance their life. With the help of this device they can navigate independently, feel convenient in the society, and also feel secure. They have many problems in their life, some of them are they cannot navigate independently, cheated in terms of money, Bus or Train number identification etc. In this system two features are implemented they are Currency Recognition and Traffic Light Detection. A Machine Learning (ML) Model is configured which consists of Convolution Neutral Network (CNN), image processing techniques, and Object detection model. Currency recognition extracts both visible and hided marks on the money for effective classification. Traffic light detection uses an image processing algorithm with standard object recognition methods. This system helps the Visually Impaired to cross the road without other's help. For the first feature i.e. Currency recognition, the output is the value of the currency that person has in his/her hand and for the second feature i.e. Traffic light detection, the output is whether the traffic light is red or green, if it is green then the number of seconds that person has to cross the road.**

Keywords*— *Smart Glass, Visually Impaired, Traffic Light Recognition, Currency Recognition, Machine Learning, Convolutional Neutral Network, Image processing, Object Detection.

I. INTRODUCTION

According to the survey conducted by the World Health Organization (WHO), in this world around 300 million are visual problems out of which 40 million are Visually Challenged. When it comes to Taiwan (Republic of China) around 1.1 million are Visually Impaired which is 4% of its total population, this shows the importance of this project regarding both engineering and commercial purposes [7]. Visually Impaired people are not able to feel free in society and socially convenient. They are facing many problems daily, which include finding reading materials, colour recognition; unable to navigate independently, easily being cheated when it comes to money. Every person wants to do things independently but that is not the case with Visually Impaired. A device is needed to help them to overcome their problems [8]. This project mainly deals with two problems Currency recognition and Traffic light detection [9]. These features are selected based on the survey done by WHO, which states that around 22% of all road deaths are Pedestrians. So, when it comes to Visually Impaired, every step is complicated. When a person wants to travel abroad, currency is the major issue. For, Visually Impaired it is very difficult to identify the currency so, they get cheated easily. In this project, we propose a system that helps them to make them feel free and navigate independently. As per the WHO statement [6] it is found that nearly 285 million people suffer from visual problems in which 246 people are suffering from mild vision problem and the remaining are suffering from blindness. In the world nearly 12.6 % of people from Eastern Mediterranean Region suffer from blindness.

II. LITERATURE SURVEY

Literature survey deals with existing methodologies using different algorithms and techniques. Older methodologies can be used or taken as a reference to configure new systems. Machine learning algorithms and Image processing techniques are taken from old systems and a new model is configured [10].

The author [1] describes the real-time traffic light recognition system used for different vehicles. In the proposed systemthey have used image processing techniques with some object recognition methods [11]. They have done it in grayscale. Different traffic lights from various countries are recognized in this system. They have used cascade classifiers for the learning process [12]. They also have made a prototype vehicle and tested in real conditions using the videos of different countries including France, China, and the U.S.A. They have good accuracy with minimal errors. In this system tracking algorithms can be added to increase the accuracy and decrease the risk factor.

The author [1] describes an efficient way to detect traffic light and countdown time also. An application named Driver Assistance System (DAS) is used to provide guidance for the driver. This application is used to identify the traffic signal. There are two parts in this proposed system. The first part deals with the detection of the colour whether it is Green or Red by colour segmentation [13]. The number of pixels of green and red is calculated, and the result is obtained by the threshold. The second part is to identify the countdown time [14]. The author [3] describes Currency recognition using image processing. They have used more than 200 different countries

around the world. They have extracted both visible and hidden features in the Currency. This helped for an efficient classification. This proposed system can reduce the human power and supervision. They have used different images of blurred and damaged noted [15]. They have used MATLAB to detect the currency using digital image processing, and neural network. They have five different currencies and the proposed system was working well for four of them.

The author [2] describes a system for automated Currency recognition. In this, the proposed system will detect both the country and the denomination. They have used pixel by pixel comparison. They have extracted the denomination value using different features or characteristics. This includes size, colour and the text. They have taken 20 different currencies and identified with good accuracy. The average time takes was 5 seconds which is very high and needs to be improved. This system also cannot find the mutilated notes. This system provides a method to detect and read text in natural images. Electronic travel aids (ETAs), electronic orientation aids (EOAs), and place locator devices are the three main categories in this framework (PLDs). This framework performs research on portable/wearable obstacle detection/avoidance systems (a subcategory of ETAs) in order to educate the research community and users about these systems' capabilities and advances in assistive technology [16]. The survey considers a number of variables and performance factors of the structures that categorizes them, presenting both qualitative and quantitative results. Finally, it assigns a rating that serves only as a guideline rather than a criticism of these systems. A novel texture-based approach for discovering texts in images is presented in current framework. The textural properties of texts are analysed using support vector Machine (SVM). The intensities of the raw pixels that make up the textural pattern are fed directly to the SVM, which works well even in high-dimensional spaces, rather than using an external texture function extraction module.

III. PROPOSED SYSTEM

In this paper, Machine learning and deep learning algorithms are used to configure a ML model. Machine learning is a logical investigation of calculations and quantitative models that Personal Computer (PC) systems use to carry out a specific task without using unambiguous instructions and instead relying on examples and induction. In this model training data is used to train the model. More accuracy can be obtained using Machine learning techniques. OpenCV framework is also used in the ML model.

Considering Traffic light detection, there is no existing system for pedestrians to cross the road. We have traffic light detection for vehicles only which also does not a good accuracy. Coming to Currency recognition, Microsoft has developed a system to find the currency for IOS devices and that system only tells which currency it is. There are many Smart Glasses for people to improve their life but, there are very few for Visually Impaired [18]. A lot of work has been done in order to recognize currencies automatically [19].

Scanner-based systems Trupti and Bawane [4] proposed a process that scans the entire piece of paper for identification. In addition to colour and surface, they found size to be one of the three parameters of paper currency. The image histogram and the variety of colours in a paper currency are taken into account. The technique described in this paper can be used to identify paper money of various countries. It also reflects an ensemble neural network-based currency recognition scheme (ENN). Negative correlation learning is used to train the individual neural networks in an ENN [17]. Negative relationship learning is used to train individuals in an ensemble on various sections or portions of input patterns. They used ENN to determine the currencies were new, old, or noisy. For categorising various types of currency, the Ensemble network is used. It is less likely to misclassify than a single network or an ensemble network with independent training. They didn't run any tests.

LookTel Money Reader [6] It is an iOS-based mobile application that helps users recognize and count their currency bills. It assists fully blind people by reminding them about the importance of a note discovered using TTS. It also supports blind people by displaying the number of known note in broad font in the screen's middle. It has a wide stock of twenty-one currencies, including the U.S. dollar. The US Dollar, Euro, United Arab Emirates Dirham, Japanese Yen, and Kuwaiti Dinar are among the twenty-one currencies available. It also operates in a number of languages, including English, Spanish, French, and Japanese. When capturing an image, it has few constraints on the device's location and operates in real-time.

ViaOptaDaily [6] It is an Android app that provides blind users access to an interactive world with six main services. The first is to aid blinds in deciding the temperature. It is capable of delivering weather reports for the existing and subsequently three days, and also notify the blind. It has a magnifying glass that helps you to take a picture with text using the device's camera. It is then used to edit the image after it has been captured. It facilitates zooming in. It can also read the text in the image and present it to the user in an audible format. The third feature is a simple timer that counts down and alerts the user when it is enabled.

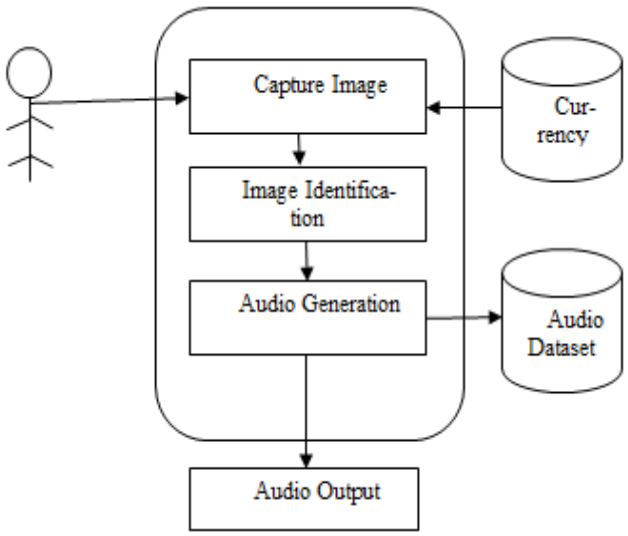


Fig: 1 General Architecture

In the proposed system images of traffic light and currency are taken as data, and it is preprocessed. We have configured a ML model and this is trained using the training data. Testing data is used to test the model whether the accuracy is good or not. The ML model consists of CNN, Object detection model and image processing techniques. For Currency recognition, CNN is used and the output is the value of the currency they have. Traffic light detection is done in four steps. First the traffic light (Object) is detected using a deep learning algorithm named You Only Look Once (YOLO) algorithm. Next the colors (Red, Green) in the image are detected using OpenCV. The detected object i.e. traffic is alone taken and rest is left behind. In the final step using CNN the output is given, that is whether the light is red or green color, if it is green the number of seconds, they have to cross the road. Finally, the output is given to visually impaired by voice.

Data Owner In this module after collecting the data, data is pre-processed and ML model is trained and tested with the images. Whenever the user gives the input data to the model, the output is given.

End User In this module, whenever the user i.e. visually impaired gives the input to the ML Model. The input is processed and output is given. For Currency recognition, the output is the value of the currency and for Traffic light detection, the output is color and countdown time. The output is given by voice. The full view of the process is as shown in Fig 2

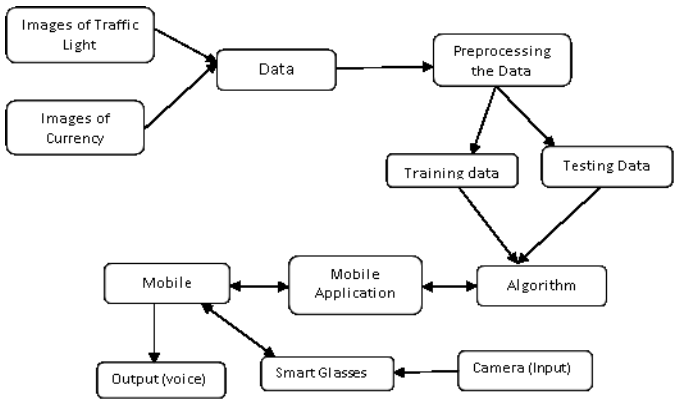


Fig 2: Overall Flow Process

IV. IMPLEMENTATION

Implementation is the phase in which the plan is converted or transformed into a working framework. The main and important stage is helping the client or user by giving them an efficient framework. Many arrangements have to be done for the execution, which includes examining the framework, looking after the requirements, planning the techniques and assessment of strategies.

Input and Output for Traffic light detection

The task includes preparing a Machine learning model for Traffic light detection. The datasets that we have used for traffic light detection consists of different locations, from different angle and distances. For testing the model, we have used images of currency that are taken in different angles and notes which are mutilated also. The images of traffic light taken from long distance and different angle are used for testing. The images with different countdown times are given for testing to improve the accuracy. For training the model, we have used the original currency notes both front and back for better accuracy. We have also used traffic lights with different countdown times and in different locations with different angles as in Fig 3A, 3B.

Fig 3A: Training Images of Traffic Light

Testing is very important in understanding the mistakes done while configuring a model. The mistakes found can be rectified and accuracy can be improved. This helps in testing the segments, parts, functions, and a complete item. The main goal is to configure a software framework which is up to the necessities and what client desires.

Fig 3B: Testing Images of Traffic Light

Real Time Scenario:

As the Visually impaired needs minimum of 15 seconds to cross the road, if the traffic signal is less than 15 seconds the output is number of seconds and Stop or else the output will be number seconds. If the traffic light is red, then the output is Red and Stop. When made into a device, first the traffic signal will be detected if the numbers of seconds are greater than 15 then the output is given. Otherwise the output is given Stop and after every 5 seconds an image is captured and tested in background unless and until the seconds are greater than 15.The entire flow is shown in Fig 4.

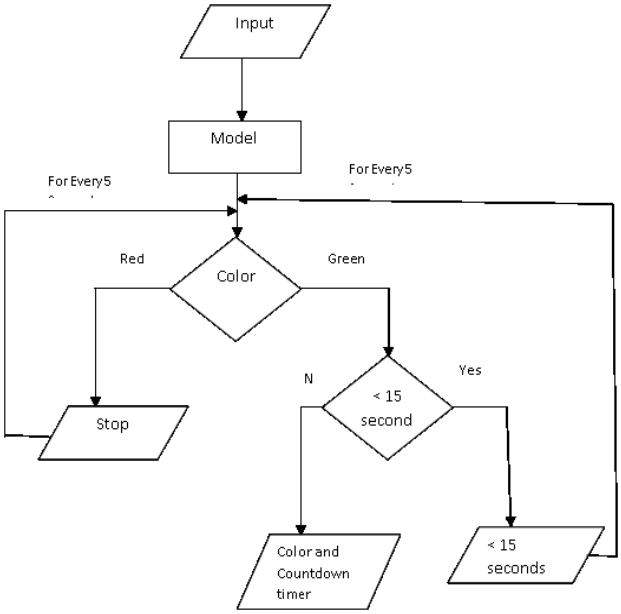


Fig 4: Real Time Implementation

Input and Output for Currency Recognition

Currency recognition, we have used many images that consist of original images and images taken by us from different angles. Some Currency notes have also been folded and mutilated. The image acquisition stage is the first phase in any vision system. After the image has been obtained, it can be processed using a number of methods to accomplish a variety of tasks. In image processing, image acquisition is the first step in the workflow since processing is impossible without an image.

There are several methods for obtaining an image, by using a camera or a scanner. All of the features should be maintained in the acquired image.

Preprocessing:

The main aim is to increase the visual quality of images and effect of datasets. The operations that are usually wanted prior to the main data investigation and mining of information are referred to as image pre-processing. The correction of distortion, decay, and noise introduced during the imaging process is what image pre-processing, also known as image restoration. Image pre-processing can remarkably increase the accuracy of an optical inspection. Image Adjusting is completed with the assistance of image interpolation. Interpolation is the technique largely used for tasks such as zooming, rotating, shrinking etc. Removal of noise is an important part, if not done can affect dissection and pattern matching. When carrying out smoothing operation on a pixel, the pixel's neighbor is used to perform any conversion. After that, a new pixel value is generated.

Background Removal

Images are taken in a number of conditions, in relation to lighting condition and context, as shown in architecture, while currency in the picture can be impaired. Image segmentation is critical for reducing the amount of data to process and removing undesirable features (background Region) that could influence decision-making. We begin with a fixed rectangular region of interest (ROI) that is forty pixels smaller than the image on all four sides. We anticipate that a significant portion of the currency will be present within this area. After you've obtained this area, you'll need to expand it to a segmentation of the entire picture. We're using the Grab cut algorithm to remove the unnecessary context in this case.

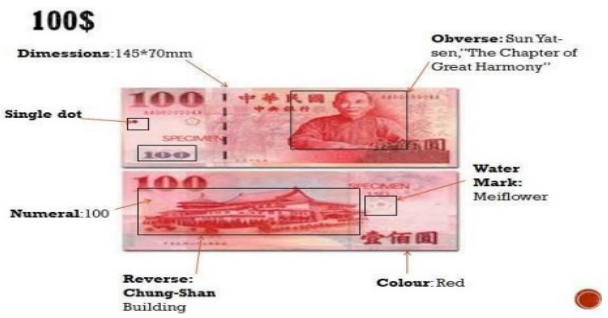
Feature Extraction

Feature extraction is a form of dimension reduction that is unique. When an algorithm's input is too large to process and is no longer needed, the data is transformed into a reduced set of features. Feature extraction is a method of translating input data into collection of features. It is assumed that if the extracted features are carefully chosen, the feature set will be complete. If the features extracted are carefully chosen, the features set should be able to extract related information from the input data in order to achieve necessary task using this reduced representation rather than full-size input.

Classification

In order to confirm image similarity, key points in the test image are checked if they are in spatial consistency with the retrieved images. Popular method of geometric verification (GV) is used by fitting fundamental matrix to find out the number of key points of the test image that are spatially consistent with those of the retrieved images. The recognized text codes are recorded in script files. Then text to speech converter is employed to load these files and display the audio output of text information. Blind users can adjust speech rate, volume and language.

The ability to identify currency (both coins and bills) without human input is unfavorable for a number of applications. Probably the most important one is assisting visually impaired people. The image processing-based currency recognition technique involves simple steps, including image collection, processing, and currency recognition. Image retrieval is typically achieved with a camera or a scanner. The images are then processed using various image processing methods, and various features are extracted from the images, which is the main principle behind currency classification. Three denominations have been taken here. i.e.100NTD, 500NTD, 1000NTD.The features that are taken in these denominations are shown in Fig 5.



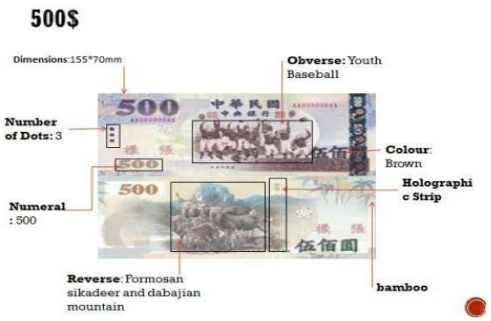


Fig 5: Features of Currency

V. CHALLENGES FACED

We have faced many challenges while doing this project. While doing the traffic signal detection using YOLO algorithm, we have faced problem with the time taken. So, we have taken tiny datasets to increase the speed. While doing the color recognition, as each color, position of image from different places effect the intensity of the color. We had to take many images as data to confirm the final range of color. One more challenge was, other lights in the image which are green, red or orange are also detected which effects the image processing. So, we have done background removal to overcome this challenge. The main challenge was to differentiate between 100 NTD and 1000 NTD currency by the system. System was unable to predict the correct output for mutilated and folded currencies of NTD 100 and NTD 1000. We used Template matching to match the good features/good matches to identify the currency but the accuracy was less and it took more time to process which will create problem in real time. To differentiate between NTD 100 and NTD 1000, we used many features of these currencies. Then we used ORB algorithm i.e., Oriented FAST and rotated BRIEF. Here we managed to do a feature matching and recognize the currency note.

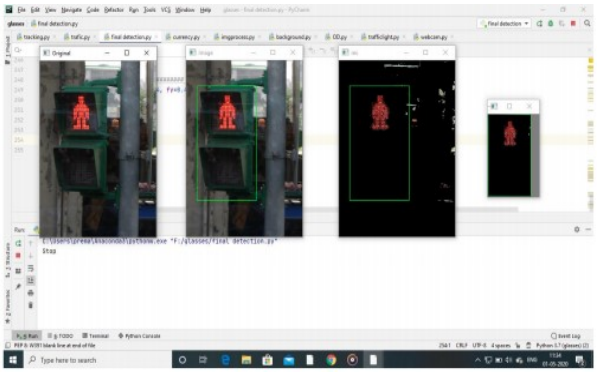


Fig 6: Output of Traffic Light

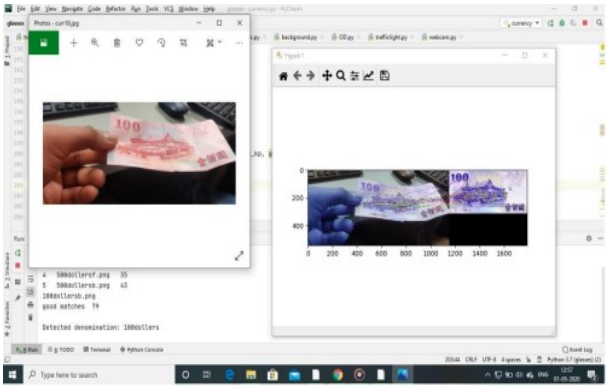


Fig 7: Output of Currency

Efficiency and Comparison of Proposed with Existing

Efficiency of the Proposed System the Efficiency of the proposed scheme is very high as different training images are used from different angles and distances. The proposed system will detect the value of the currency and identifies the traffic light color and countdown time. This system can also detect mutilated notes. This helps the visually impaired to improve their standard of living and make them feel socially convenient.

Considering Traffic light detection, there is no existing system for pedestrians to cross the road. The traffic light detection for vehicles only which also does not a good accuracy. This proposed system helps the visually impaired to know whether the light is red or green color, if it is green the number of seconds, they have to cross the road. Coming to Currency recognition, Microsoft has developed a system to find the currency for IOS devices and that system only tells which currency it is. This proposed system helps the visually challenged to know the value of the Currency. Even mutilated notes can be detected from this system. There are many Smart Glasses for people to improve their life but, there are very few for Visually Impaired.

Advantages of the Proposed System

- Make the Visually Impaired socially convenient.
- Improves their standard of life
- Helps the Visually Impaired to navigate independently.
- Visually Impaired can feel physically convenient.

VI. PERFORMANCE EVALUATION

For performance evaluation, we have tested our system on various data-set sizes. The testing process was initially done by Manual data that was collect by us using a normal digital camera with a minimum of 14 images in which 7 where currencies and 7 were traffic signals. In which we received a processing time of 31.02 seconds and matching time was 15.25 sec. Then consecutively we increased the number of images and we found that there was a maximum of 56 images we got 180.25 sec for processing time and 70.32 sec for Matching

time.The analysis is shown in TABLE 1 and the chart representation is shown in Fig8

TABLE 1: Performance Evaluation

Dataset Size	Dataset Processing Time (in seconds)	Matching Time (in seconds)
14	31.02	15.25
28	73.96	38.24
42	142.21	52.78
56	180.256	70.325

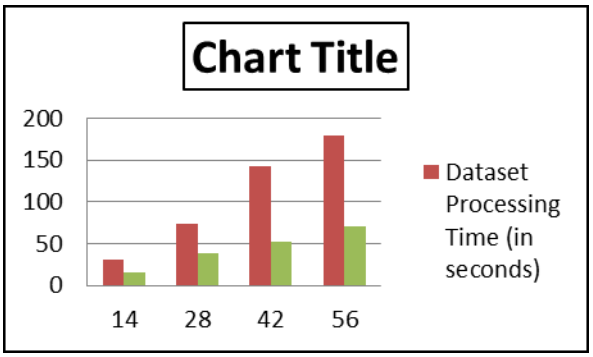


Fig8: Performance Evaluation

VII. CONCLUSION

We have developed software for Smart Glasses which helps Visually Impaired to know the value of the currency with them and the traffic light, when they want to cross the road. Smart devices area unit thought as for future, however, this project tries to bring that within the gift and got success to some extent. Smart Glass might be thought of as higher than the product within the market relating to cost-effectiveness and practicality. The results show however it can be regenerated to a product from the epitome with minor modification in style. The device will address an outsized population of visually impaired as well as the vision loss because of accidents, or any diseases. This helps them to feel free and socially convenient.

VIII. FUTURE ENHANCEMENT

In future, we can develop a device from this proposed system and we can add many more features for the Smart Glasses to identify Bus or Train numbers, Restroom detection, and Color detection ...etc. and also improve the existing features. For Currency recognition, we can recognize the Currency of different countries which helps Visually Impaired when they travel abroad. This can help the Visually Impaired for improving their life and standard of living. The accuracy should also be increased so that errors will not happen. A navigation system can also be included in Smart Glasses which makes them more comfortable.

REFERENCES

[1] Dr. D.K. Shedge et.al . "Road Sign Detection Using Cascade Classifier" International Journal of Scientific & Engineering Research, Volume 8, Issue 4, April-2017

[2] F. Daraee,, and S. Mozaffari. "Eroded money notes recognition using wavelet transform." In 6th Iranian Machine Vision and Image Processing (MVIP), 2010, pp. 1-5. IEEE, 2010.

[3] R. M. Haralick,, and L. G. Shapiro, Computer and Robot Vision, Volume I, Addison Wesley, 1992, pp. 28-48.

[4] A. Ms Trupti Pathrabe, and B. Dr NG Bawane. "Paper currency recognition system using characteristics extraction and Negatively Correlated NN Ensemble." International Journal of Latest Trends in Computing (E-ISSN: 2045-5364) vol. 1, no. 2 (2010), pp: 121-124.

[5] J. H. Park, and C. S. Jeong (2009). Real-time Signal Light Detection, International Journal of Signal Processing and Pattern Recognition, 2 (2), 1-10.

[6] Noura A. Semary , Sondos M. Fadl , Magda S. Essa and Ahmed F. Gad :Currency Recognition System for Visually Impaired: Egyptian Banknote as a Study Case IEEE 2015 International Conference on Information and Communication technology and Accessibility, ICTA 2015

[7] P. Prabu and S. Duraisamy, "Impact of Pair Programming for Effective Software Development Process." Research India Publications, International Journal of Applied Engineering Research, vol. 10, pp.18969-18986, 2015.

[8] P. Prabu, P and S. Duraisamy, "Reducing Defects in Software Development Process using Pair Programming.", International Journal of Computer Science and Communication Networks, vol. 5, pp. 128-135,2015.

[9] B. Kirubandand and P. Prabu , "Hybrid Server With Zigbee Technology Using Job Scheduling And Queuing Petrinet.",International Journal of Research and Technology,vol. 3, pp. 123-130, 2016.

[10] Jini Jojo Stephen and P. Prabu, "Detecting the magnitude of depression in Twitter users using sentiment analysis.", International Journal of Electrical and Computer Engineering,vol .9, pp. 3247-3255, 2019.

[11] Megana Ghosh and P. Prabu, "Empirical analysis of ensemble methods for the classification of robocalls in telecommunications.", International Journal of Electrical and Computer Engineering,vol .9, pp. 3108-314, 2019.

[12] Jay Kotecha and P. Prabu, "An Investigation on android background services for controlling the unauthorized accesses using android LOG system.", International Journal of Engineering & Technology(UAE), vol .7, pp. 301-305, 2018.

[13] Ankur Rameshbhai Khunt and P. Prabu, "An Empirical Analysis of Android Permission System Based on User Activities.", Journal of Computer Science, vol .14, pp. 324-333, 2018.

[14] K. Venkatachalam., et al. "Cross-Layer Hidden Markov Analysis for Intrusion Detection.",CMC-Computers, Materials & Continua, vol .70, pp. 3685-3700, 2021.

[15] K. Venkatachalam, P. Prabu, Abdulwahab Almutairi and Mohamed Abouhawwash," Secure biometric authentication with de-duplication on distributed cloud storage.", PeerJ Computer Science,vol .7, pp. e569, 2021.

[16] Mohd Anul Haq, Abdul Khadar Jilani and P. Prabu, "Deep Learning Based Modeling of Groundwater Storage Change.", CMC-Computers, Materials & Continua, vol .70, pp. 4599-4617, 2021.

[17] K.S.Bhuvansehwari, et a1. "Improved Dragonfly Optimizer for Instrusion Detection Using Deep Clustering CNN-PSO Classifier.", CMC-Computers, Materials & Continua, vol.70, pp. 5949-5965, 2021.

[18] Soumi De, P. Prabu, Joy Paulose, "Effective ML Techniques to Predict Customer Churn", IEEE Xplore 2021 Third International Conference on Inventive Research in Computing Applications, pp. 895-902, 2021.

[19] K. Amrutha, P. Prabu, "ML Based Sign Language Recognition System", 2021 International Conference on Innovative Trends in Information Technology (ICITIIT), pp. 1-6, 2021.