International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.014 ISSN: 2581-4621

Smart Assistive System for the Visually Impaired

Mejalin Arno B[1], Vinothagan J[1], Arun Kumar R[1], Gowthami V[2]
[1]UG – Electronics and Communication Engineering, Anna university, MIT Campus,Chennai, Tamilnadu
[2]Teaching Fellow, Electronics and Communication Engineering, Anna university, MIT Campus, Chennai, Tamilnadu
Orchid ID: https://orcid.org/0009-0003-8983-1176

ABSTRACT
Recent advances in innovation have made ordinary people's lives richer, simpler, and easier. Many people suffer from vision impairment, which causes many difficulties in their daily work, according to World Health Organization (WHO) statistics. Our goal is to develop a framework that is unobtrusive, safe, portable, and versatile in order to assist blind people in their daily lives. It intends to develop an effective system for assisting the visually impaired with obstacle detection and scene classification. The Raspberry-Pi 3B+ and a camera module are used in the proposed method. It photographs the scene and then preprocesses the images with Viola-Jones and Yolov5 object detection algorithms. The aforementioned techniques are used to detect people and objects. We also used a sound system that speaks when an object or person is detected. The presented study implements simple computations and detects obstacles with high efficiency. Unlike other frameworks, this framework is lightweight and easy to transport
Keywords—Viola Jones Algorithm, Yolov5s Algorithm, Face and Object Detection.

1.Introduction

The five basic senses are essential to human life. The most important is obviously sight. Visual impairment is defined as the inability to see at a certain level with the eyes, even when using conventional means such as a lens or glasses. Blind people have no such vision. As a result, carrying out daily tasks becomes extremely difficult for them. This can result in difficulties that can only be temporarily overcome by a few helpers, and certain situations can be fatal not only to the individual but also to everyone around them. Many studies specifying the design of devices for visually impaired have resulted from scientific advancements. These devices were simple and long-lasting, but they lacked usability and precision. Visually impaired people require a capable device in order to move around independently. Today, we can see that many techniques and devices have been developed to assist these people in moving freely even in a dynamic environment. [1].

2.Motivation

As the modern world has become more reliant on computers and artificial intelligence, these technologies have become more dependable and efficient. Even detecting a particular Vehicle in large traffic is possible now [2]. However, these technologies still have many gaps. A review of prior literature reveals that RFID sensors, Logitech cameras, and embedded systems have all been used to design efficient systems in the past. Visual impairment diseases have been found to increase over time. The motivation for this project is thus to create a system that enables visually impaired people to classify scene and avoid obstacles through object recognition, allowing them to live normally in the world. If an obstacle in the user's path is detected, the system will track it and notify the user. Many obstacles can occur along the length of the route during this process. This information is critical for the user. For example, a stool in the user's path should be recognized. Failure to do so may endanger the user. This flaw should be identified early on so that it can be avoided later. In contrast, if something as small as a pebble is present, the system can dismiss it. Because even the most ordinary person overlooks minor annoyances. As a result, the system recognizes the scenes of the intended user. To avoid harm, all of this information is conveyed to the person through speech. Visually impaired people require a capable device in order to move around independently. We experimented with various hardware to develop a system that helps visually impaired people achieve optimal performance on our system. The hardware selected should be energy efficient in terms of battery life, night vision, and weight. Similarly, we researched various techniques but chose the one with the simplest and easiest-to-implement algorithms. A Raspberry- Pi 3B+ and a camera module are used in our proposed project.

3.Literature Survey

3.1 Wearable Assistive System Using Deep Learning

Yimin Lin, Kai Wang, Wanxin Yi, Shiguo Lian " Deep Learning based Wearable Assistive System for Visually Impaired People".

Through the use of a wearable assistive device that was developed utilizing deep learning techniques, it enables the visually impaired people be aware of the events that are happening around them. In this research, the author presented a deep learning-based assistance system to enhance visually impaired people's environment perception. A powerful processor used mostly for deep learning inferences, a wearable terminal containing an RGBD (Red Green Blue Depth) camera and an earphone, as well as a smartphone enabling touch-based interface make up the system. Utilizing RGBD data and the pre-existing semantic map, a data-driven learning strategy is suggested to predict secure and trustworthy walking instructions. Through clever touchscreen interactions, this map is also used to help visually impaired grasp the objects and arrangement in their 3D surroundings. This learning-based approach to obstacle avoidance yields good performance in addressing low lying obstructions in both indoor and outdoor datasets, according to the quantitative and qualitative experimental results.

3.2 Facial Features Detection

Modesto Castrillón, Oscar Déniz, Daniel Hernández, Javier Lorenzo, "A comparison of face and facial feature detectors based on the Viola– Jones general object detection framework".

Viola Jones Algorithm used to detect the facial features of the person in-front of the visually impaired people. Any system combining vision based human computer interaction must be capable of quick and accurate face and facial feature detection since the human face gives useful information during contact. This capability has been the focus of various methodologies, but only opensource implementations have seen widespread use by researchers. The Viola-Jones object identification framework, which has been utilized widely in the field of facial processing, is a suitable example. For the face detection scenario, the OpenCV community shares a number of public domain classifiers. These classifiers, however, have seldom been tested on the same datasets and have typically been trained under various circumstances and using various sets of data. In this study, we attempt to close that gap by reviewing the merits and downsides of each public classifier's performance in order to establish a benchmark for future methods. Researchers can select a specific classifier for their particular situation with the use of reliable comparisons.

3.3 Text to Speech Conversion

Trupti Shah, Sangeeta Parshionikar, "Efficient Portable Camera Based Text to Speech Converter for Blind Person." In this paper, Text Reader with Camera Module with Portability is a prototype built using Raspberry Pi 3b and Python to read text from handheld objects for blind people. This paper proposes a more effective method for text localization and extraction in order to detect text regions in images. Text size is an important consideration that should be selected carefully in order to make the procedure more general and insensitive to different font shapes and sizes. The proposed method includes his four steps of object detection, text localization, text extraction, and text-to speech conversion. Text localization algorithms are used to locate and extract text from regions of interest extracted from the cluttered background. The extracted text from the ROI is converted to speech. It

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.014 ISSN: 2581-4621

works better with Optical Character Recognition. A recurrent neural convolutional network has been proposed to train words individually. Experimentation and training are performed on the synth 90k word dataset. Finally, the proposed model was developed using OCR and CRNN.

4. Methodology

As presented in Fig 1, our design's architecture is comprised of Raspberry Pi, which serves as the heart of our guiding system.

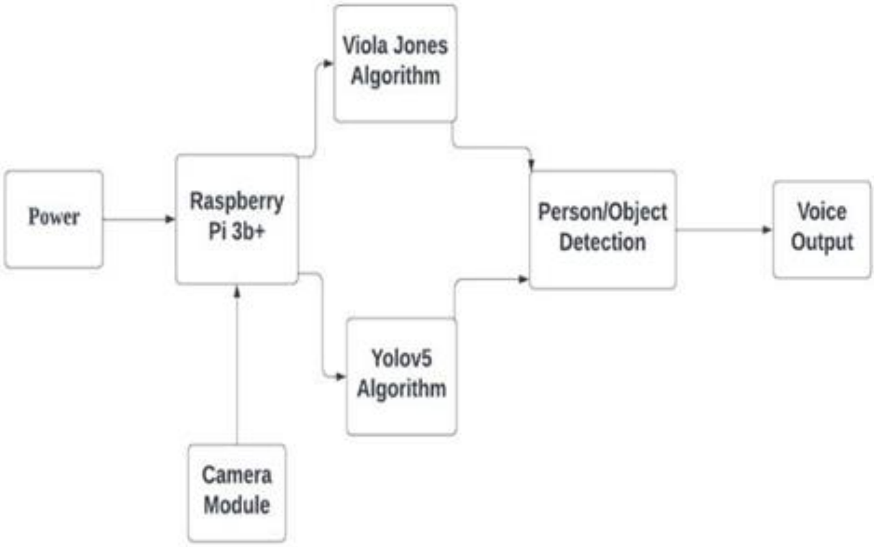


Fig 1. Architecture of the proposed methodology

It receives a 5V regulated power supply and includes a camera module. The Viola Jones and Yolov5 frameworks are then loaded into our microcontroller. It is displayed if an object or a person is detected. Furthermore, because our system must serve visually impaired people, we have included speech signal output. When a person or object is detected, the speech output can be heard through an earphone. Ninety-one different objects were detected by our system. The OpenCV community has made available a set of public domain classifiers for the scenario of face and object detection [3].

4.1 Viola Jones Algorithm

The Viola-Jones object recognition framework, developed by Paul Viola and Michael Jones in 2001, can be trained to recognize various object types. The success rate of this algorithm is about 78.4% [4]. A detection frame seeks features that comprise the sum of image pixels within a rectangular region. The functions of Viola and Jones are typically more complex, as they all rely on multiple rectangular regions. The sum of the pixels in the simple rectangle minus the pixel density of the dark rectangle equals the value of each feature. Primitives are rectangular features of this type. This scheme also adapts to vertical and horizontal characteristics, but with a much coarser response. A slightly modified framework employing the Viola-Jones algorithm produces significantly better results. With its implementation and integration in OpenCV, the described framework has achieved remarkable results. The cascading classifier made good use of combined features. All parts can be recognized by the framework without any critical errors. However, because processing at this level takes a long time, the time lag and processor load should be reduced. This is critical for sensor development because visually impaired people will use it. To improve our concept, the AdaBoost features have been classified. A set of horizontal features, such as edges and line features, is chosen for this purpose. It was then improved to include strong classifications. This is required because if many weak classifiers are used for face recognition, combine all of them into one strong classifier. This increased accuracy while decreasing wasted time. All parts of the image are unfused and thus distinguishable. It produces very high-quality output with positive attributes. This precise upper body (object) detection method is especially useful for face enhancement, removing the need for object matching to identify upper body regions in unconstrained still images and videos. Using the Viola Jones face detection

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.014 ISSN: 2581-4621

computation, the Cascade Objects Indicator detects human bodies and identifies face objects. This model recognizes the upper body region, also known as the head and shoulder region, as well as a rearrangement of the chest and face. Furthermore, the eyes are recognised because they are darker than the rest of the face. As a result, vertical haar features capture the eyes. Eye detection is thus possible from small changes in the integral image with low noise and haze. As a result, the model's accuracy is high, and the results are excellent. Each image sub-window has over 180,000 rectangle features, which is far more than the number of pixels. [5]

4.2 Haar Features

Every human face has distinct features. These regularities can be as eyes, lips, and nose, the values of which are determined by directed gradients of pixel intensities. Viola and Jones made use of two rectangular elements. There are three kinds
: 2, 3, and 4. matched using haar features. Human faces have several distinguishing characteristics, including eye areas that are darker than the upper cheeks and a bridge of the nose that is whiter than the eyes. The machine recognizes each feature by looking at a box with a light and dark side, as shown in the features below. As with the brow's edge, there are times when one side is lighter than the other. The center section may occasionally shine brighter than the surrounding boxes, giving the impression of a nose. Fig 2 below shows different kinds of haar features namely: edge features, line features and four-sided features.

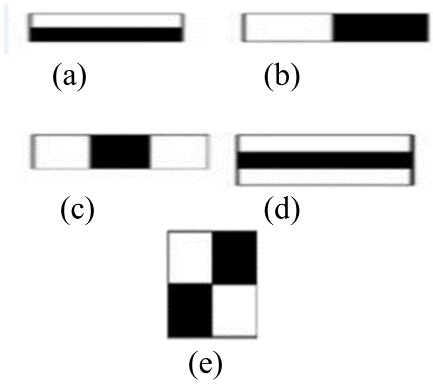


Fig 2. Different Types of Haar features

4.3 Integral Images

An integral image is a representation of image that analyzes rectangular features in real time and outperforms the more comprehensive version. This is because the rectangular region of each feature is always adjacent to at least one other rectangle. 20 Each 2 Rectangular Feature, 3 Rectangular Feature, and 4 Rectangular Feature can be deduced into 6 ArrayRefs, 8 ArrayRefs, and 9 ArrayRefs.In the previous section, we calculated the value of the feature. Since a large feature has a significantly higher number of pixels, these calculations can be extremely time-consuming. The integral image helps us complete these time-consuming computations quickly and determine whether a feature or a group of features meets the requirements.

4.4 Adaboost Training and Cascaded Classifiers

The object identification framework employs a variant of the AdaBoost learning method to select the most useful features and train a classifier that employs them. This method generates strong classifiers by linearly combining weighted simple weak classifiers.

As shown in Fig 3, a multistage classifier that finds patterns fast and reliably is a cascade classifier. The AdaBoost method generates a strong classifier at each stage. Stages in the cascade are constructed by training classifiers using Adaboost. The number of weak classifiers for a strong classifier grows over time. On the input, a sequential (gradual) evaluation is performed. If the classifier returns a negative test result for a specific step, the inputs are automatically removed. If the outcome is favorable, the input is forwarded to the next level. This multi-step method, according to Viola and

Jones, provides a simpler method for quickly rejecting most negative (non-facial) data while focusing on positive (facial) data. Allows for the creation of classifiers. In other words, it differentiates faces from all other objects in the image under consideration.

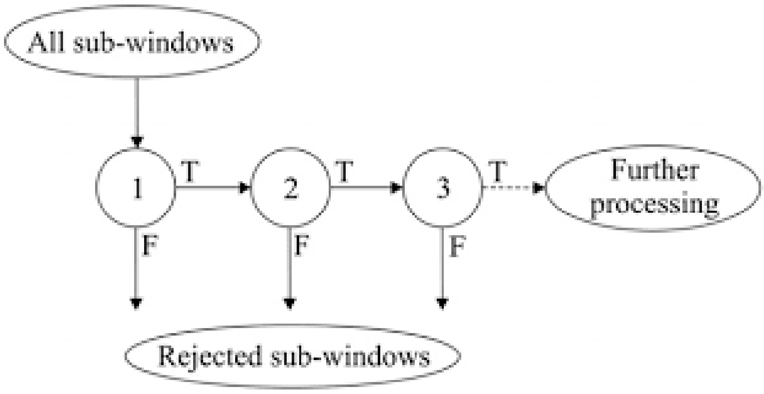


Fig 3.Cascade Classification

E.Yolov5 Algorithm

There are two types of object detection models: two-stage object detectors and single-stage object detectors. Single- stage object detector architectures (such as YOLO) are made up of three parts to make dense predictions: a backbone, a neck, and a head. The given image was divided into grids of S x S in the YOLO-based object detection, where S = the number of grid cells in each of the axes[6].There, each unit of the grid was accountable to detect the targets which were getting into it. A trained network extracts rich feature representations from images as the backbone[7]. This helps to reduce the spatial resolution of the image while increasing the feature (channel) resolution. The model neck is used to extract the feature pyramids. This helps the model generalise to objects of different sizes and scales. To complete the final stage operations, the model head is used. It generates the final output, which includes classes, objectness scores, and bounding boxes, by using anchor boxes on feature maps. The improved ASPP (Atrous Spatial Pyramid Pooling) module is used to replace the SPP (Spatial Pyramid Pooling) module of the original YOLOv5s algorithm to obtain the receptive fields of images of various sizes, allowing for better context capture [8]. This software can even further be extended to vehicle and pedestrian detection[9].

4.5 Hardware

The Raspberry Pi is compatible with a wide range of programming languages. The Raspberry Pi Foundation recommends Python as a programming language for beginners. Any programming language that compiles for ARMv6 will, in theory, run on the Raspberry Pi. As a result, users are not restricted to using only Python. The Raspberry Pi comes preloaded with several programming languages, including C, C++, Java, Scratch, and Ruby.

Raspbian Wheezy is a free operating system based on the Debian package management system. It was created by a small team of Raspberry Pi enthusiasts. Raspbian is designed specifically for the Raspberry Pi hardware and includes over 35,000 packages and precompiled software. Raspbian is still actively developed, with the goal of improving the stability and performance of Debian packages. Raspbian is officially recommended for beginners, and it includes the LXDE graphical desktop environment. Raspbian Wheezy is one of the quickest ways to get your Rasp Pi running.

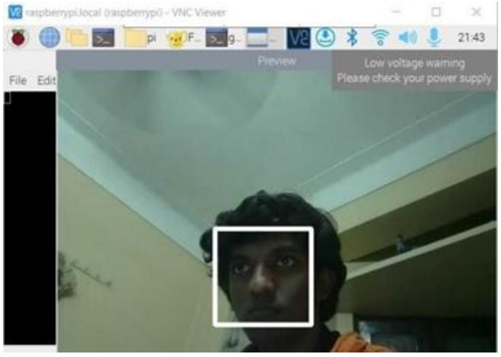
The Pi NoIR camera module is one of these additional components. NoIR is an abbreviation for "No Infrared." This means the Pi NoIR camera module lacks an infrared filter. As a result, images taken in daylight appear strange, but you can see them in the dark by using an infrared flash. The camera module is small, roughly the size of a quarter coin. The maximum resolution of the camera is 5 megapixels (2592 x 1944 pixels). It has a fixed focus lens and the Omnivision 5647 sensor. The camera module is also capable of recording videos in full HD resolution. The camera module works with all Raspberry Pi models and can be accessed through the MMAL and V4L APIs. There are

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.014 ISSN: 2581-4621

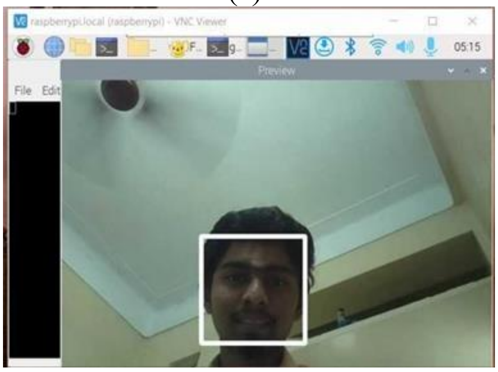
currently several third-party libraries designed for camera modules. The Picamera Python library is one of these. Our hardware implementation can be further improved to various forms using wearable sensors [10]

5 Results

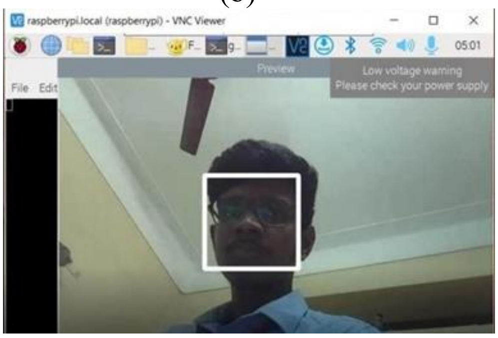
The results of our output are shown below in two phases: Face detection and the object detection.



(a)

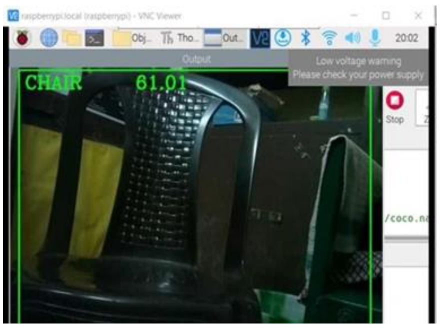


(b)



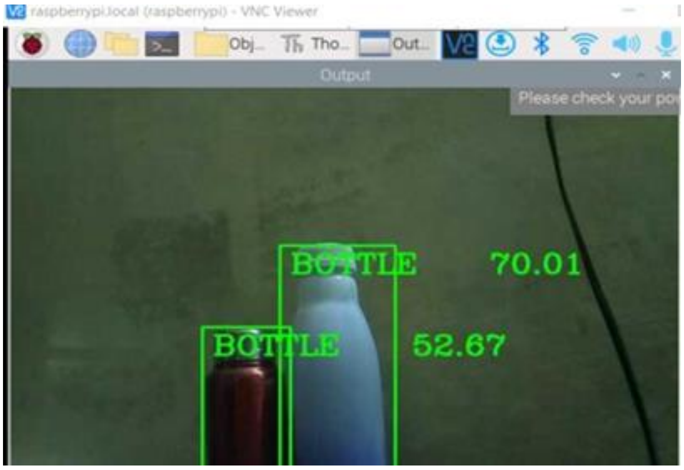
(c)

Fig 4. Face Detection of Team Members

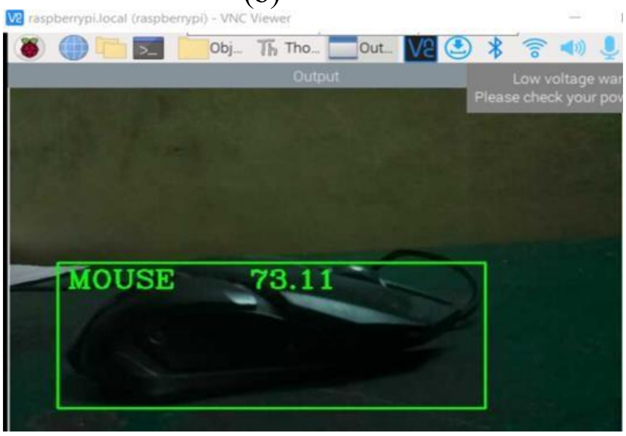


(a)

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.014 ISSN: 2581-4621



(b)



(c)

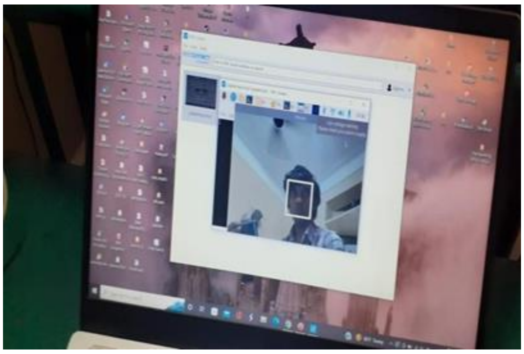
Fig 6. Object Detection

Table 1. Accuracy of Detected Objects

S.No	Object	Accuracy
1	Chair	61.01%
2	Bottle 1	52.67%
3	Bottle 2	70.01%
4	Mouse	73.11%

(a)

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i04.014 ISSN: 2581-4621



(b)
Fig 7. Hardware Setup

CONCLUSION

The Raspberry Pi and Camera module were used to implement the face and object detection. Face and object detection are handled by Viola Jones and Yolov5, respectively. The image processing technique of Viola-Jones algorithm works efficiently in our assistive system. The object detecting algorithm is also very accurate and it displays its accuracy. This system is capable of detecting a wide range of objects, including animals and various obstacles. The assistive system also includes the output of the speech signal. The speech signal is implemented with the help of a wireless headphone, via Bluetooth connection. So this system can assist any visually impaired person with ease. Finally, by using the Raspberry Pi display, the system can be made more portable. Power could also be supplied to the Pi via a portable battery. The system's accuracy can also be improved by using a better camera module.

References

1. H. Sharma, M. Tripathi, A. Kumar, and M. S. Gaur, ‘‘Embedded assistive stick for visually impaired persons,’’ in Proc. 9th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT), Jul. 2018 .
2. Liu Xiaomeng, Feng Jun, Chen Peng, “Vehicle Detection in Traffic Monitoring Scenes Based on Improved YOLOV5s”, IEEE, Jul. 2022.
3.M. Castrillón, O. Déniz, D. Hernández, and J. Lorenzo, ‘‘A comparison of face and facial feature detectors based on the Viola– Jones general object detection framework,’’ Mach. Vis. Appl., vol. 22, no. 3, pp. 481–494, 2010.
4.M. N. Chaudhari, M. Deshmukh, G. Ramrakhiani, and R. Parvatikar, ‘‘Face detection using Viola Jones algorithm and neural networks,’’ in Proc. 4th Int. Conf. Comput. Commun. Control Autom. (ICCUBEA),Aug. 2018
5.P. Viola and M. Jones, ‘‘Rapid object detection using a boosted cascade of simple features,’’ in Proc. Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2001.
6.R. Joshi, S. Yadav, M. Dutta, and C. Travieso-Gonzalez, ‘‘Efficient multi-object detection and smart navigation using artificial intelligence for visually impaired people,’’ Entropy, vol. 22, p. 941, Oct. 2020.
7. Usman Masud , Tareq Saeed , Hunida M. Malaikah , Fezen Ul Islam , AND Ghulam Abbasi , “Smart Assistive System for Visually Impaired People Obstruction Avoidance Through Object Detection and Classification”, IEEE Access, 17 January 2022.
8.Yaroslav Paramud, Dmytro Kushnir, “Model for real-time object searching and recognizing on mobile platform”, IEEE Access, 2020.
9.Ying Wang, Rui Li, Qi Qi, Xiaojie Su, “An improved one-stage detector for vehicle and pedestrian detection on campus AGV”, IEEE Access, 2021.
10.Z. Bauer, A. Dominguez, E. Cruz, F. Gomez-Donoso, S. Orts- Escolano, and M. Cazorla, ‘‘Enhancing perception for the visually impaired with deep learning techniques and low-cost wearable sensors,’’ Pattern Recognit. Lett., vol. 137, pp. 27–36, Dec. 202

11. Girshick, Ross. "Fast r-cnn." In Proceedings of the IEEE international conference on computer vision, pp. 1440-1448. 2015.
12. Shoeb Ahmed Syed; Steve Ales; Rajesh Kumar Behera; Kamalakanta Muduli. "Challenges, Opportunities and Analysis of the Machining Characteristics in hybrid Aluminium Composites (Al6061-SiC-Al2O3) Produced by Stir Casting Method". International Research Journal on Advanced Science Hub, 4, 08, 2022, 205-216. doi: 10.47392/irjash.2022.051
13. Ashima Saxena; Preeti Chawla. "A Study on the Role of Demographic Variables on Online Payment in Delhi NCR". International Research Journal on Advanced Science Hub, 4, 08, 2022, 217-221. doi: 10.47392/irjash.2022.052
14. M. S. N. K. Nijamudeen, G. Muthuarasu, G. Gokulkumar, A. Nagarjunan, and T. Pravin, “Investigation on mechanical properties of aluminium with copper and silicon carbide using powder metallurgy technique,” Advances in Natural and Applied Sciences, vol. 11, no. 4, pp. 277–280, 2017.
15. R. Devi Priya, R. Sivaraj, Ajith Abraham, T. Pravin, P. Sivasankar and N. Anitha. "MultiObjective Particle Swarm Optimization Based Preprocessing of Multi-Class Extremely Imbalanced Datasets". International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems Vol. 30, No. 05, pp. 735-755 (2022). Doi: 10.1142/S0218488522500209
16. Pravin T, M. Subramanian, R. Ranjith, Clarifying the phenomenon of Ultrasonic Assisted Electric discharge machining, “Journal of the Indian Chemical Society”, Volume 99, Issue 10, 2022, 100705, ISSN 0019-4522, Doi: 10.1016/j.jics.2022.100705
17. T. Pravin, C. Somu, R. Rajavel, M. Subramanian, P. Prince Reynold, Integrated Taguchi cum grey relational experimental analysis technique (GREAT) for optimization and material characterization of FSP surface composites on AA6061 aluminium alloys, Materials Today: Proceedings, Volume 33, Part 8, 2020, Pages 5156-5161, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2020.02.863.
18. V.S. Rajashekhar; T. Pravin; K. Thiruppathi , “Control of a snake robot with 3R joint mechanism”, International Journal of Mechanisms and Robotic Systems (IJMRS), Vol. 4, No. 3, 2018. Doi: 10.1504/IJMRS.2018.10017186
19. J. Redmon and A. Farhadi, "YOLO9000: Better faster stronger", 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2261-2269, Jul. 2017.
20. Ando, Bruno, Salvatore Baglio, Vincenzo Marletta, and Angelo Valastro. "A haptic solution to assist visually impaired in mobility tasks." IEEE Transactions on Human-Machine Systems 45,2015
21. Soni, Laxmi Narayan, Ashutosh Datar, and Shilpa Datar. "Implementation of Viola-Jones Algorithm based approach for human face detection." Int. J. Curr. Eng. Technol 7, 2017
Swaminathan, Avinash, Mridul Chaba, Deepak Kumar Sharma, and Yogesh Chaba. "Gender classification using facial embeddings: A novel approach." Procedia Computer Science, 2020