A DEEP LEARNING BASED MODEL TO ASSIST BLIND PEOPLE IN THEIR NAVIGATION

Nitin Kumar*	Lovely Professional University, Phagwara, India	nitin.14652@lpu.co.in
Anuj Jain	Lovely Professional University, Phagwara, India	a1978jain@gmail.com

* Corresponding author

ABSTRACT

Aim/Purpose	This paper proposes a new approach to developing a deep learning-based prototyping wearable model which can assist blind and visually disabled people to recognize their environments and navigate through them. As a result, visually impaired people will be able to manage day-to-day activities and navigate through the world around them more easily.
Background	In recent decades, the development of navigational devices has posed challenges for researchers to design smart guidance systems for visually impaired and blind individuals in navigating through known or unknown environments. Efforts need to be made to analyze the existing research from a historical perspective. Early studies of electronic travel aids should be integrated with the use of assistive technology-based artificial vision models for visually impaired persons.
Methodology	This paper is an advancement of our previous research work, where we performed a sensor-based navigation system. In this research, the navigation of the visually disabled person is carried out with a vision-based 3D-designed wearable model and a vision-based smart stick. The wearable model used a neural network-based You Only Look Once (YOLO) algorithm to detect the course of the navigational path which is augmented by a GPS-based smart Stick. Over 100 images of each of the three classes, namely straight path, left path and right path, are being trained using supervised learning. The model accurately predicts a straight path with 79% mean average precision (mAP), the right path with 83% mAP, and the left path with 85% mAP. The average accuracy of the wearable model is 82.33% and that of the smart stick is 96.14% which combined gives an overall accuracy of 89.24%.
Contribution	This research contributes to the design of a low-cost navigational standalone system that will be handy to use and help people to navigate safely in real-time scenarios. The challenging self-built dataset of various paths is generated and

	transfer learning is performed on the YOLO-v5 model after augmentation and manual annotation. To analyze and evaluate the model, various metrics, such as model losses, recall value, precision, and maP, are used.
Findings	These were the main findings of the study:

- To detect objects, the deep learning model uses a higher version of YOLO, i.e., a YOLOv5 detector, that may help those with visual impairments to improve their quality of navigational mobilities in known or unknown environments.
- The developed standalone model has an option to be integrated into any other assistive applications like Electronic Travel Aids (ETAs)
- It is the single neural network technology that allows the model to achieve high levels of detection accuracy of around 0.823 mAP with a custom dataset as compared to 0.895 with the COCO dataset. Due to its lightning-speed of 45 FPS object detection technology, it has become popular.

Recommendations for Practitioners	Practitioners can help the model's efficiency by increasing the sample size and classes used in training the model.
Recommendations for Researchers	To detect objects in an image or live cam, there are various algorithms, e.g., R-CNN, Retina Net, Single Shot Detector (SSD), YOLO. Researchers can choose to use the YOLO version owing to its superior performance. Moreover, one of the YOLO versions, YOLOv5, outperforms its other versions such as YOLOv3 and YOLOv4 in terms of speed and accuracy.
Impact on Society	We discuss new low-cost technologies that enable visually impaired people to navigate effectively in indoor environments.
Future Research	The future of deep learning could incorporate recurrent neural networks on a larger set of data with special AI-based processors to avoid latency.
Keywords	visually impaired, handheld assistive technology, assistive technology, wearable devices, blind, navigation, object detection

INTRODUCTION

Approximately one thousand million people in the world experience vision impairment losses caused by eye diseases, such as cataracts, keratoconus (KC), or diabetic retinopathy. A resolution adopted in the 73rd assembly of the World Health Organization (WHO) in the year 2020, which aims to promote integrated, people-centered eye health care, and prevention of blindness and vision impairment, was passed in the year 2019 (Keel et al., 2021). Until now, people with visual impairment have relied on low-cost traditional methods like a white cane, which is an effective touch-based technique but doesn't usually detect the shape of obstacles. It is also believed to be a low-cost navigational method if guided dogs are used, but dogs are less accurate due to their self-specified mood while moving. It also uses sensor-based assistive devices like GPS as a good alternative, but GPS is self-restrained in avoiding obstacles while on a mission. Electronic travel aids (ETAs) are also being used to guide people with visual impairments. One example of this, proposed by Bhatlawande et al. (2014), is an Electronic Mobility Cane (EMC) which constructs a logical map based on the surroundings. Takatori et al. (2006) have designed an IC tag system that is used for indoor walking which makes moving through a defined space easier. The floor color lanes were identified using an RFID tag and a one-chip processor (Mocanu et al., 2018). The process for selecting and recognizing 11 different clothes is also provided by X. Chen et al. (2019). Deep Convolutional Neural Networks (DCNN) for object detection reduces the amount of data and costs associated with hardware by using advanced

technology (X. Chen et al., 2019). Convolutional Neural Network (CNN) based algorithms are much more accurate and can be used both for previously known environments as well as unknown environments.

An assistive framework, such as the DEEP-SEE framework proposed by Mocanu et al. (2018), could be employed for distinguishing known faces from unknown ones. Scientists have developed systems with RFID tags, Bluetooth, ultrasonic, Wi-Fi, and a camera for assisting visually impaired people to navigate their environment. A powerful AI navigation device for visually impaired people, such as NVIDIA Jetson TX2 embedded with DLSNF (Deep learning based Sensory Navigational Framework) provides an efficient device for navigation. In addition to detecting cracks in pavements, researchers have also discovered a way to detect currency. There are mobile apps that can help visually disabled people navigate a range of environments safely. With the assistance of acoustic waves, a person could also access information. The You Only Look Once (YOLO) algorithm can also be employed for checking the quality of objects, such as corn; the whole simulation was performed on NVIDIA TX2. By using tensor flow and the coco data set to train with 328K images, YOLO Object detection in the path of the user can be detected. Binaural sounds are generated via the HRTF (Head Related Transfer Function), which is used along with cameras to estimate and navigate efficiently in the environment.

Optical character recognition (OCR) and Text to speech (TTS) techniques were used to read text on the obstacles and smoothly navigate. To guide a user in map building, Simulation Localization and Mapping (SLAM) can be applied using sensors like RPLIDAR A2 and Kinect V1. The entire simulation can be run on Robot Operating System (ROS) with Turtlebot2. The validity of the YOLO algorithm has been verified with multiple data sets including COCO, VOC, and VisDrone.

Based on the above-mentioned research work, this research aims:

- To present a method to produce a module that feeds the user with vital data with greater accuracy, such as identifying individuals and detecting obstacles in the path.

- To facilitate practical implementation, we propose a deep learning algorithm, i.e., YOLO, which is compatible with user-friendly hardware.

The organization of this paper is as follows: the next section presents the literature review and related work. The third section presents the proposed methodology. The fourth section presents results and discussion and, finally, the fifth section concludes this work and provides a recommendation for future work.

LITERATURE REVIEW AND RELATED WORK

This section briefly discusses the research conducted on assistive technology for visually impaired people. In this paper, the majority of the discussion centers on these topics and the simulation presented. With the introduction of convolutional neural networks-based deep learning algorithms, some of the best solutions are now available, and they make a major impact, especially when integrated with IoT (Internet of Things). In this article, we discuss those deep learning algorithms and their merits, disadvantages, and limitations.

VARIOUS ALGORITHMS IN NAVIGATIONAL ASSISTANCE FOR VISUALLY IMPAIRED PERSONS

For object detection, the Scale Invariant Feature Transform (SIFT) algorithm is considered to be an image local feature description and one of the most popular algorithms that can extract key features from a frame. The solutions developed for detecting and tracking objects in videos are presented. Using the improved k-means clustering algorithm, key features are grouped into groups that detect moving objects (Sharif et al., 2019). Zheng et al. (2018) used log-polar transform to stabilize the

video. They then compare the results of their proposed solution with the SIFT-ME algorithm to determine their affinity for transformation. Karami et al. (2015) combined SURF and SIFT to detect objects and produced an object-tracking algorithm that is robust. To speed up the identification of objects, Speeded Up Robust Features (SURF) was developed. A study of the SURF algorithm was performed by P. Ding et al. (2018), Geng and Qiao (2017), Karami et al. (2015), Tareen and Saleem (2018), T. Wang et al. (2018), and Xingteng et al. (2015). As part of this research, SURF performance enhancements are being investigated as well as comparative analyses with algorithms such as SIFT, ORB, and BRIEF. By improving the SURF algorithm, illumination invariance and matching rates will be achieved.

One of the most widely used algorithms for identifying objects is optical character recognition (OCR). The main application of this technique has been to detect text in images. The OCR algorithm can be used to identify objects according to work by Adriano et al. (2019), Deshpande and Shriram (2016), Liem et al. (2018), and Liu et al. (2016) who discuss how it could be used to accommodate more applications. Among these studies are experiments on food label detection, expiry date detection, and retrieving ID card information using the OCR-based model. Zhu et al. (2018) provided a detailed description of their proposed technique for detecting elevator buttons. The purpose of this study is to assist service robots in navigating to their desired floors by detecting buttons. As a result of combining OCR and Faster R-CNN, the authors came up with OCR-RCNN, a single neural network. In this case, they have compiled a custom dataset of elevator panels for applying the OCR-RCNN algorithm.

Since its release, the You Only Look Once (YOLO) algorithm has been very popular. YOLO is a method of detecting objects that only need to be looked at once, as its name suggests. The YOLO detector is a single-shot detector, which means that instead of doing image classification and localization in two separate steps, it performs it in just one step. There is a minor degradation in accuracy (Khalfaou et al., 2022) resulting in speeding up the overall process of object recognition. In the past few years, the YOLO algorithm has undergone major upgrades – versions v2, v3, v4, and v5 – while version-1 is still considered the original algorithm. Several studies are using YOLO to detect objects and compare them to FasterRCNN (Benjdira et al., 2019; Cao et al., 2019; Y. Chen et al., 2018; Y. Kim & Jung, 2018; Park et al., 2018; Redmon & Farhadi, 2018). In these systems, a deep learning-based neural network is trained on a dataset to create a pre-trained model.

A vision-based system was proposed by H. Wang et al. (2017) to assist visually impaired persons during navigation. In this system, there is a camera, a haptic feedback device, and an embedded computer to assist blind people in locating objects. As a wearable system, this system provides mobility to blind users. In this experiment, blind people wearing the system are made to walk through a maze. Based on the results of the experiments, this system aids blind users in navigating a path without colliding with obstacles. However, navigation is slower using this system than when using a cane. In combination with a cane, this system can improve the speed of navigation for blind users. According to Liu et al. (2016), an assistive system that uses the OCR algorithm can be used to help blind persons read. It consists of a glove that is equipped with an embedded camera index that is worn by the blind to assist in navigation. The blind person uses the index finger of the glove hand over the first sentence from left to right. As the picture of the text is processed by the camera beneath the finger, audio will be output. According to the experiment results, the finger can read the text in the image through the camera; however, the overall process is lagged as it creates a high-resolution image by combining several images of the same frame. Due to the experiments being conducted through a webcam mounted on a table for reading text, it is also necessary to evaluate the project's feasibility. Maolanon and Sukvichai (2018) and Rajesh et al. (2017) propose a model to aid blind users. An embedded computer is used by their system to process the text in images, which is handled by Raspberry Pi. On the one hand, Rajesh et al. (2017) use the OCR algorithm, while on the other hand Maolanon and Sukvichai (2018) use the YOLOv1 algorithm. The Text to Speech (TTS) library helps blind people to hear the text read out to them. The results of the experiment indicate that it

performs well when it comes to reading text, but it has poor performance when it comes to detecting faces. As a result of its fast and accurate detection capabilities, YOLO can attain an accuracy of 83%. A comparative list between algorithms is shown in Table 1.

Table 1. Shows comparative analysis of various algorithms

Algorithm	Author	Advantages	Limitations
SIFT (Scale Invariant Feature Transform)	Deekshitha et al. (2018), Jabnoun et al. (2015), Karami et al. (2015), Sharif et al. (2019), Zheng et al. (2018)	High accuracy and invariance to rotation and scaling.	The algorithm is very computationally intensive, slow, performs poorly against blurring, occlusion, and changes in illumination, and is patented.
SURF (Speeded Up Robust Features)	P. Ding et al. (2018), Geng and Qiao (2017), Y. Kim & Jung (2018), Tareen and Saleem (2018), R. Wang et al. (2019)	Fast, robust, requires little computation power and is robust to transformations.	Inherently less accurate than SIFT, fewer key points are detected than other algorithms and patented.
OCR (Optical Character Recognition)	Adriano et al. (2019), Borisyuk et al. (2018), Deshpande and Shriram (2016), Hosozawa et al. (2018), Liem et al. (2018), Liu et al.(2016), Zhu et al. (2018)	Accurately and quickly recognize text in the images.	Not light-resilient, low image detection accuracy.
Regional Proposed networks	Ammirato and Berg (2019), X. Chen et al. (2019), Chao et al. (2018), Hu et al.(2018), Li et al. (2019), Najibi et al. (2019)	It is fast and capable of handling transformations, as well as dealing with illumination changes well.	The training process is quite lengthy, including two stages, making it slower than SSD.
YOLO (You Only Look Once)	Benjdira et al. (2019), Cao et al. (2019), K. Kim et al. (2018), Maolanon and Sukvichai (2018), Park et al. (2018), Redmon and Farhadi (2018), Widyastuti and Yang (2018)	The algorithm performs all of its predictions with the help of a single fully connected layer of neural network in one step, making it the fastest. It is robust in the face of transformations, and it also works well when there is occlusion or illumination change.	High-speed detection.

PROPOSED METHODOLOGY

The proposed work features a software design as well as a hardware design, the software design being an Artificial Intelligence-based module that enables obstacle-aware navigation for people with disabilities. It consists of a 1.2 GHz quad-core Cortex A53 processor, bone conduction headphones, and a 1.2 GHz 64-bit quad-core Cortex A53 processor. It is built to support the tracking and detection of moving objects. Datasets used to train the proposed model include real-time scenarios.

We associate the trained CNN model (Kumar & Jain, 2021) with the live camera feed once the model has been trained. The bounding box for the detected area is drawn if the object is detected. A path is

chosen from left, right, or straight based on the live frame once the number of regions with the most likelihood of detection has been determined.
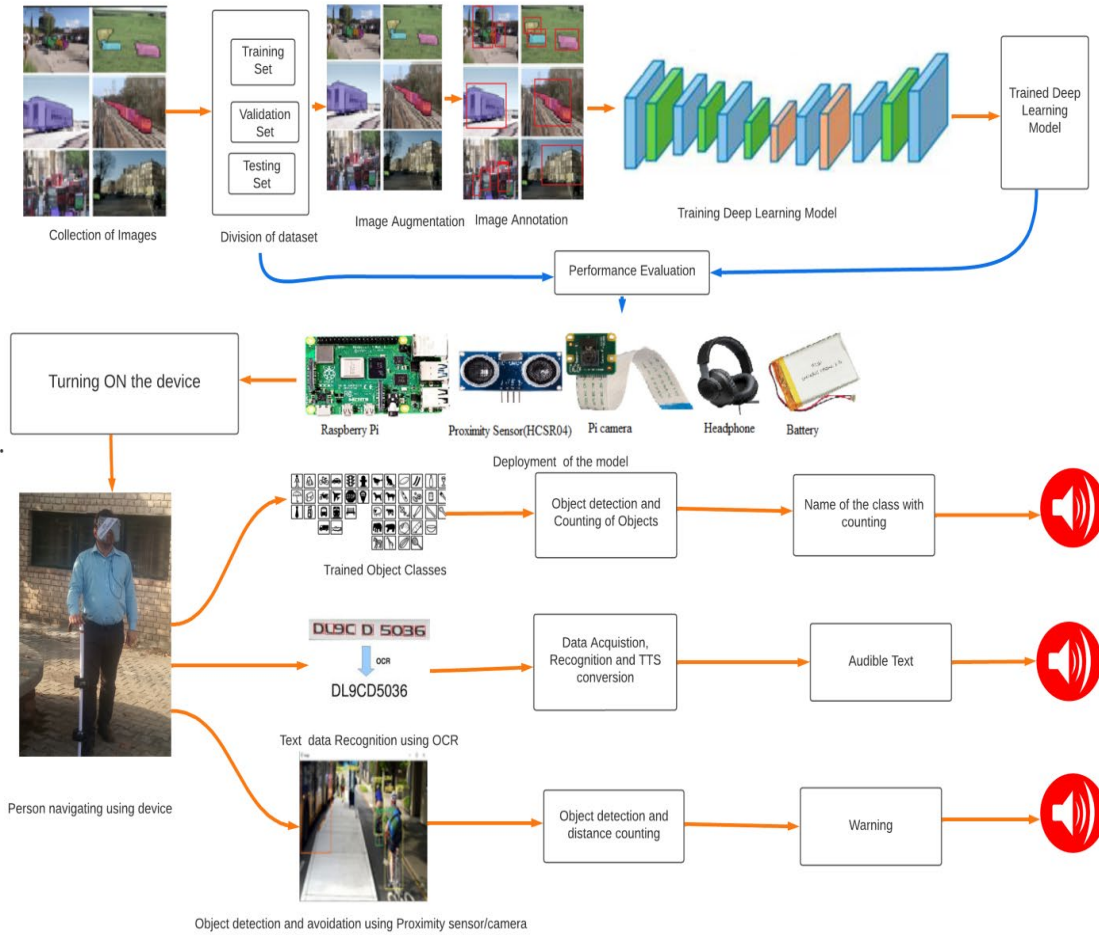

Figure 1. Block diagram of the system

Two major hardware models are present in the model, namely Wearable Mask and a Smart Cane. The user has to analyze the outputs of the wearable mask and smart cane for navigating through the path as shown in Figure 1.

HARDWARE DESIGN

The overall structure of a wearable mask and smart cane consists of components like a Raspberry Pi 3, an Ultrasonic Sensor, a GPS (Global Positioning System), a camera, a headphone, and a power supply.

In our system, Raspberry Pi 3 is the brain since it contains a 64-bit quad-core Cortex A53 processor running at 1.2 GHz with Wi-Fi and Bluetooth 4.1 and is almost 10 times faster than lower versions of Raspberry Pi, i.e., Raspberry Pi 1. It comes with 1 GB of RAM, an integrated camera interface, and a Video Core IV 3-D graphics chip. The YOLOv5 application is implemented through a Python script. This application reads frames from the video stream when it accesses our Open VINO environment. It performs near real-time object detection using a Raspberry Pi and YOLO.

The detection of objects will be done through an ultrasonic sensor (HC SR04). It detects obstacles in the way, mainly by using ultrasound. A static current of 2 mA is produced by the HC-SR04 module when it is powered by 5 V DC. It is capable of measuring distances between 2 and 450 cm. The output ranges from Low to High when detecting an object. Our model uses a Pi Camera of 5-megapixel RGB (Red Green Blue) model B camera capable of taking static images of 2592 x 1944 pixels. Additionally, 1080p30 and 640x480p60 video formats are supported. It is supported with an Audio Jack that supports 4 poles and 3.5 millimeters. A person who is visually disabled will be assisted by audio signals through headsets connected to this device. The overall system uses 18650 lithium batteries of 3.7 V with a capacity of 2200 mAh. The overall system is supported by GPS module APM 2.5 NEO-M8N which is a low-cost, low-powered position sensor. It offers a 9600 baud rate and is useful in applications such as Data Logger for position, velocity, and time. The operating voltage of the GPS is 3-5 volts. The whole model was 3-D printed for encapsulating entire hardware components. Acrylonitrile Butadiene Styrene (ABS) is used in the model. The model is created in Fusion 360 and printed in ABS using 3-D printing.

3D MODEL DESIGNING OF WEARABLE MASK AND A SMART CANE

Hardware models are designed in Fusion 360 as shown in Figure 2, a cloud-based software platform that supports 3-D modeling, CAD, CAM, CAE, and PCB for designing a variety of products. Our Wearable 3-D model has dimensions of 18 cm in length to 12 cm in breadth, and the depth of the model is 4 cm to use with ease.
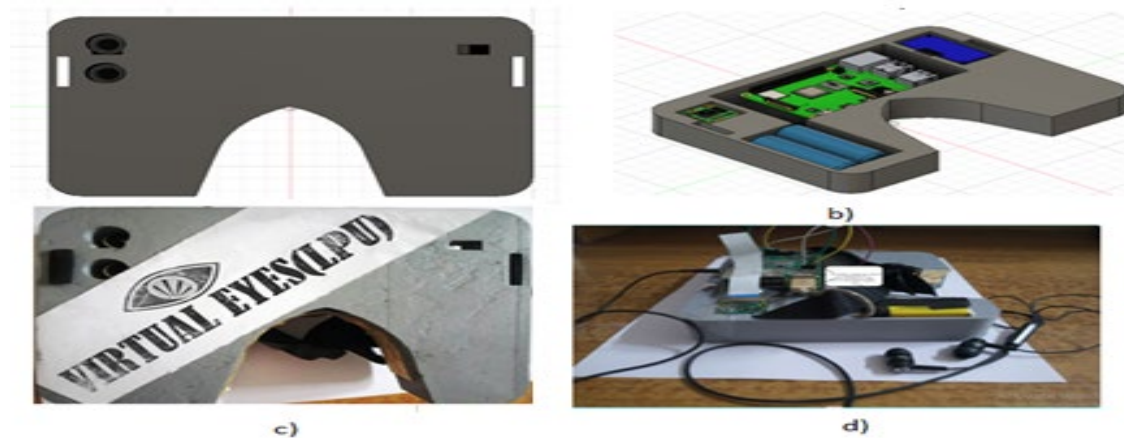
**Figure 2. 3-D model on Fusion 360: a) front view, b) rear view
c) printed front view, d) rear printed view**

The front view consists of two openings in the design for camera input and the ultrasonic sensor. The opening for the camera is 1 cm and that for the ultrasonic sensor consists of 1.5 cm each. The rear view consists of spacing for capsulation of 4 objects, namely ultrasonic sensor (45 mm X 20 mm X 1.5 mm), camera (25 mm X 23 mm X 9 mm), Raspberry pi (56.5 mm X 85.6 mm X 17 mm) and the batteries (50 mm X 29 mm). The model is worn on the head of the visually impaired person and rested on the nose of the user. The weight of the model is 240 grams which is easily worn all long day by the VI person.

For the Smart Cane, we have designed a 3-D model for our existing stick that has dimensions of 65 mm X 120 mm X 50 mm. It will incorporate a Raspberry pi (56.5 mm X 85.6 mm X 17 mm) and the batteries (50 mm X 29 mm). The model is designed for a fixed diameter (25 mm) stick which is depicted in Figure 3. It will enhance more visibility of the scene.
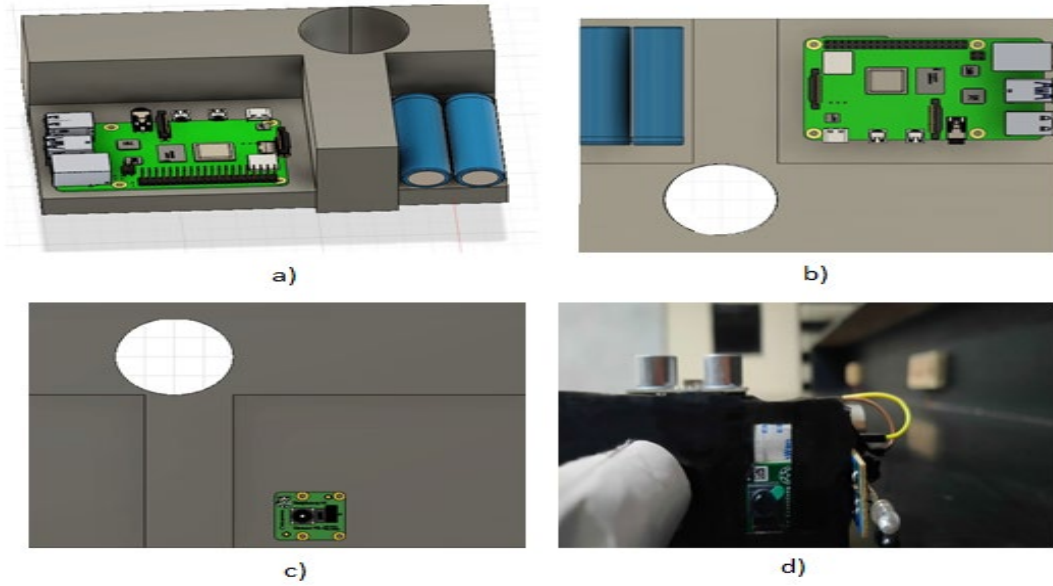
**Figure 3. 3-D models: a) side view, b) top view,
c) rear view, and d) original model designed on Fusion 360 software**

SOFTWARE DESIGN

The system includes modules for acquiring images, pre-processing, enhancing, and annotating them.

A) Image Acquisition: The images are acquired through a camera model Raspberry Pi 3, which can record up to 60 frames per second (FPS) at 640x480p.

B) Image Augmentation: Collected images were then enhanced using various methods such as flipping, brightness levels, noise levels, etc.

C) Image Annotation: Before training, images were annotated with the LabelImg tool, and then a bounding box was placed around the detected objects. During this process, images and bounding box positions were saved to an XML file.

D) Data Sets: In terms of data sets, there are many existing datasets for path recognition such as CIFAR (Calik & Demirci, 2018), IMAGENET (Deng et al., 2009), PASCAL (Everingham et al., 2009), COCO (Lin et al., 2014) and SUN (Xiao et al., 2010), but these are limited for a small number of classes.

E) Deep-learning model: There are many deep-learning models such as YOLO, SSD, and FASTER-CNN (L. Chen et. al., 2018) which have many advantages and disadvantages. With our model, we are utilizing a YOLOv5-based model for detecting a road for people with disabilities.

RESULTS AND DISCUSSION
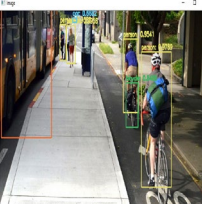
ANALYSIS OF RESULTS WITH DIFFERENT DATASETS

Analysis with COCO dataset

The datasets used in this research work were collected from Common Objects in Context (COCO) data sets. It is one of the largest scale object detection datasets which comprises 330K images with 80 object categories, some of which are listed as a person, bus, train, bicycle, or car. The other reliable datasets that can train our model are as follows: COCO, COCO128, VOC, Argoverse, VisDrone, GlobalWheat, xView, Objects365, and SKU-110K.

As depicted in Table 2, the model is tested for various known environments and the outcomes are among the 80 classes stated by the COCO model, i.e., traffic lights, cars, persons, trucks, cycles, etc. The amount of time taken to analyze this data varies from less than 1 second to 3 seconds, depending on the number of objects detected in the environment.

Table 2. Result analysis with predefined classes

Real-time environment data	Object detection in the environment	Time taken	Class type	No. of classes
		2.874795 seconds	Known	4 classes: 'car', 'traffic light', 'motorbike', 'person'
		1.765487 seconds	Known	4 classes: 'persons', 'bus', 'bicycle', 'car'
		1.765487 seconds	Known	4 classes: 'car', 'traffic light', 'truck'

Analysis with a manual dataset

For our model, we manually captured 100 images for each of the 3 classes: Straight-Path, Left-Path, and Right-Path. The images were captured in the daytime. Through these 300 images, the dataset is divided into 90% for training the model and 10% is kept for testing the model. We kept the optimal distance between the object and the wearable device mounted on the visually impaired to around 1-2 meters. The resolution of the digital camera is 3264 X 2448. The YOLOv5 model requires labeled data that comprises of class-label and position of all ground truth in images, which could be automated using annotation that reduces further errors. We annotated our images in YOLO format (.yml) and trained across 60 epochs. In addition to enhancement of the images, such as scaling, transformation flipping, and data augmentation, techniques were performed on the data as well. With our data we used to evaluate the IOU values at 0.5 (50%) and 0.95 (95%) through several metrics which included Precision, Recall, and mAP (mean average precision). The graphs in Figure 4 illustrate the metrics curves as they progressed through training. Based on the evaluation, the validity precision score of the YOLOv5 model was 0.8057, a recall score of 0.95, and the mAP score was 0.95 for @0.5I OU and 0.64 for @0.95 IOU with a detection average speed of 49.66 FPS.
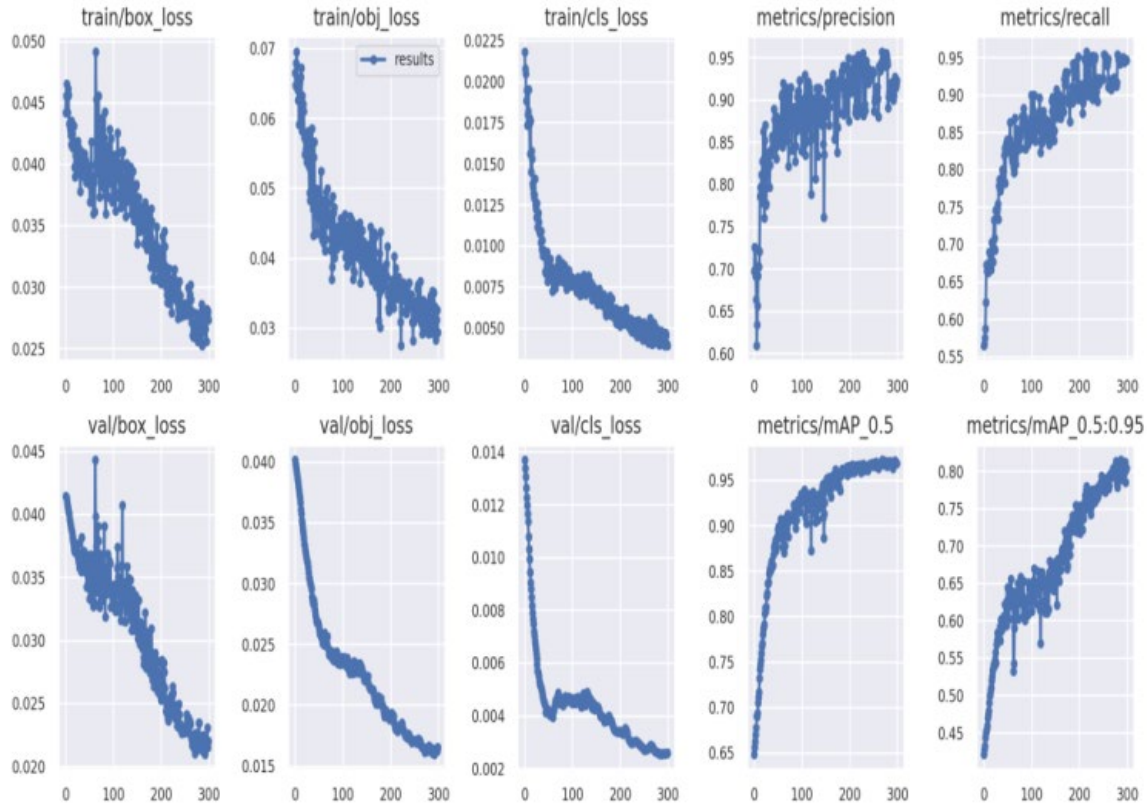
Figure 4: Graph of losses, recall, precision, and mAP with data training

Analyzing data: For analyzing the image, the image has to be converted into a grid of cells. A grid of S X S cells is used to split each frame of the image into cells responsible for prediction. Each grid cell predicts "B" bounding boxes and probabilities for C classes. There are five components in the predicted bounding box (x, y, w, h, confidence) as shown in Figure 5. The image center (x, y) is represented by the bounding box. All dimensions (x, y, w, h) are normalized to [0,1]. The YOLO prediction is based upon the formula S*S*(B*5+C) where S means grid cells, B means bounding boxes, and C means class probabilities, where we used S=13, B=2, and C=20 to give the equation (13*13(2*5+20). We trained the data with 1500 iterations.

Figure 5. Encoding an object

The confidence score in the image can be given by Equation 1:

$$CS = P_r(Obj) * IOU_{Groundtruth}^{Predicted} \tag{1}$$

Where CS = Confidence Score, $P_r(Obj)$ represents the probability of the object and the IOU Predicted Ground truth represents the IOU of predicted and ground truth bounding boxes. A confidence score (CS) of Zero means there is no object in the cell. A Confidence score tending towards value 1 is considered to be the best.

The cost function or loss function of YOLOv5 can be given by Equation 2 (Song et al., 2021):

$$\lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} \left[((\sqrt{w}_i - \sqrt{\widehat{w_i}})^2 + \right.$$
$$\left. (\sqrt{h}_i - \sqrt{\widehat{h_i}})^2 \right] + \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} [(C_i - \widehat{C_i})^2] + \lambda_{noobj} \tag{2}$$

In the above equation, 3λ constants represent more than one aspect of the loss function, whereas λ_{coord} represents the highest order.

YOLOv5 is a more advanced and upgraded version of YOLO (Redmon & Farhadi, 2017) and YOLOv5. A boundary box around an input frame is predicted by using logistic regression and Feature Pyramid Network (FPN) in YOLOv5. To detect objects, YOLOv5 uses 53 convolutional layers to extract features from Darknet-53.

ROUTE DETECTION MODULE

These days, special paths are specially designed for the visually impaired person. These paths are tactile and support easy movements. In our project, these tactile paths are converted into a smart path using augmented Quick Response (QR) code.

Tactile Surface Paving: Tactile paving is a walking surface indicator (Xiao et al., 2010) that can produce a warning when this is detected with long canes or by walking on it. Tile sizes are currently determined by ISO/FDIS 23599, which is designed as an assistive product for visually impaired and blind people. The tiles were different in their tactile characteristics, such as tiles with parallel blister lines or tiles with offset blister lines.

TRACKING MODULE

Tracking module consists of two components, namely Path Traversed Module and Mapper.

Path Traversed Module: We are looking for recording the path traversed by a visually impaired person. The VI user started from 36 blocks to 14 blocks in Lovely Professional University. This module records GPS data and stores it as a Comma Separated Value (CSV) file. There are two columns in the file – Latitude and Longitude – which are values from the GPS module. The GPS data is appended every 30 seconds to the excel sheet as shown in Table 3.

Table 3. Test model specifications and test conditions

Time(min)	Place	Latitude	Longitude
0	36 Block	31.258189	75.707936
5	Auditorium	31.254047	75.70484
10	13 Block	31.254643	75.705323
15	LPU Mall (14 blocks)	31.255074	75.705666
20	Baldevraj Mittal Hospital	31.256661	75.70631

Mapper: This evaluates the current data and the previously stored data. In the event that the user entered and moved through designated mapped steps then the output will be received by the user in the form of audio as "You are moving in the right direction" and if the user navigates off the path, then the message "It seems you have departed from the path" will be received through the microphone. As shown in Figure 6, the path comprises movement through the tactile path, stairs, a round bottleneck path, and QR-enabled pavements.

Figure 6. Shows the movement of the visually impaired person through various environments

We took various sets of GPS data (Redmon & Farhadi, 2018; Rodríguez et al., 2012) from various test environments for our proposed model. It includes data from 36 blocks to 14 blocks in Lovely Professional University (LPU), Punjab, India. The set of data includes data from the 9th floor of a 36-block building which has Blister-type tactile flooring and data from transient classrooms of that level to the movement in and around 14 blocks as shown in Figure 7.

Figure 7. Mapping of a visually impaired person

As shown in Table 4, the model accuracy is calculated by averaging recorded data from each ultrasonic sensor 3 times with the same objects. The data is taken while the user is moving on a rounded bottleneck path which is guided by 2 side walls. From the data, we find that as the distance between the object and sensor increases, the accuracy of the model decreases gradually. The average accuracy of the vision-based system is 96.14%.

Table 4. Data received from multiple ultrasonic sensors

Actual Distance (cm)	Measured Distance (cm)			Average	Accuracy (%)	Error (%)
	1	2	3			
25	24.1	24.3	24.8	24.4	97.5	2.47
50	49.25	48.75	48	48.7	97.4	2.63
75	72.25	74.13	72.02	72.8	97.1	2.89
100	97.75	94.3	98.35	96.8	96.8	3.25
125	123.0	121.3	117.5	120.6	96.5	3.47
150	145.0	148.8	139.7	144.5	96.3	3.68
175	172.9	169.5	162.2	168.2	96.1	3.87
200	191.8	188.3	193.5	191.2	95.6	4.37
225	222.8	219.75	200.65	214.4	95.3	4.67
250	237.8	234.3	238.3	236.8	94.7	5.28
300	288.9	284.3	275.5	282.9	94.3	5.67

As depicted in Table 4, the highest accuracy turns out to be 97.5 when the user is 25 cm from the obstacle and the level of accuracy level falls when the user moves away from the obstacle The accuracy is worst for the cases above 300 cm and less than 3 cm.

Testing/Evaluation: For testing, we chose 2 places and 2 candidates for evaluation of our device, as shown in Figure 8. We chose the environments from our university, i.e., (1) Garden Area, (2) Classroom Area, and (3) Area enclosed from 36 blocks to 14 blocks. Our first environment was the Garden Area comprising chairs, trees, plants, polls, signboards, etc. Our second environment was the Classroom area which comprised the tactile path where various Quick Response codes were tagged with black tapes that provided useful information to the VI user. The last environment comprised the path from 36 blocks to 14 blocks in Lovely Professional University, Phagwara, India.

As shown in Table 5, the comparison set of YOLOv5 and SSD (single shot detector) is analyzed in the situation where a blind and visually impaired person is made to initially navigate through a known environment and later made to test through an unknown environment. For the environment, it is believed that the path can have 3 major movements: a straight path, a left path, and a right path. We tested this environment first with the YOLOv5 model (Zhang et al., 2022) and then compared it with the SSD algorithm. YOLOv5 produces better results with an average accuracy of 82.34% whereas SSD gives an average accuracy of 80%. The results indicate better Frames Per Second (FPS) for YOLOv5 than SSD, i.e., approximately 45 FPS as compared to 34 FPS, as shown in Figure 9 and Figure 10. For the model, we collected 100 images for each class. For the straight path, we got 79% mean Average Precision (mAP), for the right path 83% mAP, and for the left path 85% mAP. The model detected a straight path took about 42 ms, the right path took 49 ms and the left path 47 ms.

The model also gave an average accuracy of 82.33%. With a smart cane, the collective accuracy increased to 89.24%.

Scene	User into scene	View from the device	Information detected	Device and Technique
Patent 1 in the garden area			3 Classes namely: 'Poll', 'Chair', 'Plant'	Wearable Device and YOLO
Patent 2 in the classroom area			Room no 102 'Seating right of the entrance'	Vision-Based Stick and OCR

Figure 8. Information detected with different devices

Table 5. mAP and time elapsed for different environments

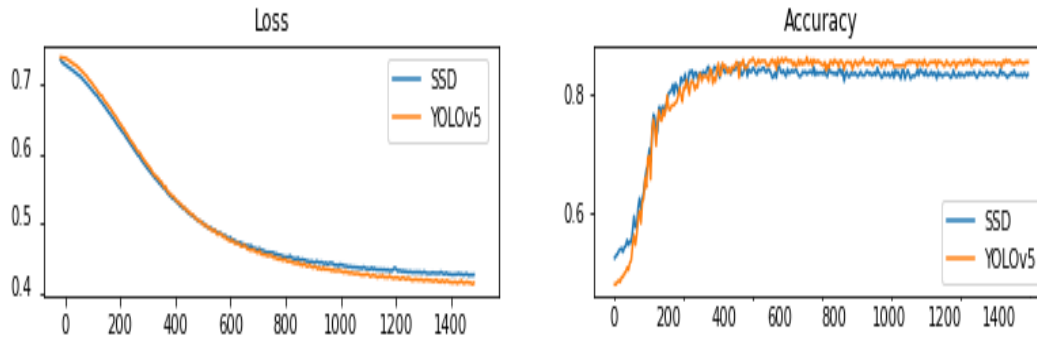| Class	Number of images	YOLOv5						SSD(Single Shot Detector)					
		True	False	None	Accuracy	Time Elapsed (ms)	FPS	True	False	None	Accuracy	Time Elapsed (ms)	FPS
Straight Road	100	79	12	9	0.79	42	50	76	11	8	0.76	40	37
Left Turn	100	85	9	6	0.85	47	45	83	10	7	0.83	43	32
Right Turn	100	83	10	7	0.83	49	42	81	10	9	0.81	48	35

Figure 9. Comparison between SSD and YOLOv5 algorithm for losses and accuracy
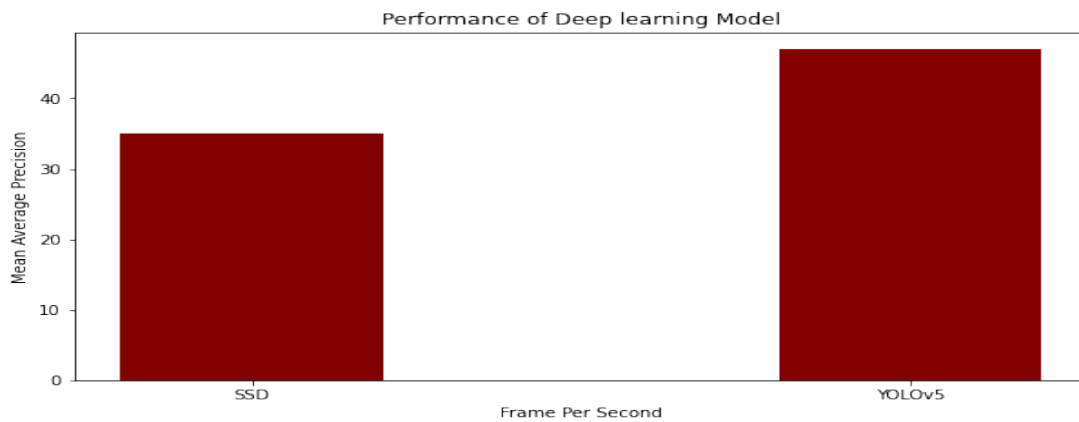

Figure 10. Comparison between SSD and YOLOv5 algorithm with mAP and FPS

CONCLUSION AND FUTURE WORK

This study aimed to develop an accurate and cost-effective solution that could be deployed to ease navigational accessibility for visually impaired people. Our model is unique due to its hybrid model that comprises a wearable device and a vision-based smart stick. The model is 3D designed in CAD through Fusion 360 and thereafter printed. The wearable device is trained by a machine-learning algorithm to detect major objects that fall in the path of the user and a vision-based stick uses GPS, ultrasonic, and a camera which adds accuracy to the existing model. The wearable model in a real-time navigational system was 3D modeled to achieve this objective and then a machine learning module was introduced to make the system more robust and adaptable to environmental changes. We propose a low-cost wearable device that could assist visually impaired individuals in finding their way through an environment by identifying their surroundings such as cars, bicycles, persons, trucks, and buses. Our model has been tested with a known dataset provided by COCO which has 80 classes, and also with our custom data set which is a modification to the existing design. The custom design comprises three additional classes – Straight-Path, Left-Path, and Right-Path – which help a user in navigating a particular area. The proposed model combines a single neural network with a full image to use an entirely different approach. Using this Network, we can determine the bounding boxes and probabilities of each region of the image. As opposed to systems like R-CNN, which require thousands of bounding boxes for a single image, the bounding boxes for the predictions are weighted according to the probabilities. YOLOv5-based models are very fast, over 1000x faster than R-CNN and 100x faster than Fast R-CNN as proposed by K. Ding et al. (2022), Karthi et al. (2021), and Wu et al. (2021). These models are very different from other models in these categories. Moreover,

YOLOv5 exhibits better accuracy and speed than other versions of YOLOv4 and YOLOv3 (Zhang et al., 2022).

The model was trained on over 300 images of various indoor and outdoor environments. We then tested the trained model with 30 different pictures of other environments. The model provides an anonymous path that contains segment classes of left, right, and straight paths which support the travel of people with visual impairments with accurate positional information and travel directions. Hence, we believe that our approach will be beneficial to visually impaired people. It is proposed that the model can solve the navigation problems of visually impaired people in indoor and outdoor environments and help them understand their environment with an overall accuracy of 89.24% as compared with previous models (Kallara et al., 2017). This model improves precision, as well as speed, by eliminating past deficiencies of non-overlapping bounding boxes.

REFERENCES

Adriano, J. E., Calma, K. A., Lopez, N. T., Parado, J. A., Rabago, L. W., & Cabardo, J. M. (2019). Digital conversion model for hand-filled forms using optical character recognition (OCR). *IOP Conference Series: Materials Science and Engineering, 482*, 012049. https://doi.org/10.1088/1757-899x/482/1/012049

Ammirato, P., & Berg, A. C. (2019). A Mask-RCNN baseline for probabilistic object detection. *arXiv:1908.03621*. https://doi.org/10.48550/arXiv.1908.03621

Benjdira, B., Khursheed, T., Koubaa, A., Ammar, A., & Ouni, K. (2019, February). Car detection using unmanned aerial vehicles: Comparison between faster R-CNN and YOLOv3. *Proceedings of the 2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS), Muscat, Oman,* 1-6. https://doi.org/10.1109/uvs.2019.8658300

Bhatlawande, S., Mahadevappa, M., Mukherjee, J., Biswas, M., Das, D., & Gupta, S. (2014). Design, development, and clinical evaluation of the electronic mobility cane for vision rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 22*(6), 1148-1159. https://doi.org/10.1109/tnsre.2014.2324974

Borisyuk, F., Gordo, A., & Sivakumar, V. (2018). Rosetta. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 71-79. https://doi.org/10.1145/3219819.3219861

Calik, R. C., & Demirci, M. F. (2018, May). In embedded systems image classification with convolutional neural network. *Proceedings of the 26th Signal Processing and Communications Applications Conference, Izmir, Turkey.* https://doi.org/10.1109/siu.2018.8404296

Cao, C., Zheng, J., Huang, Y., Liu, J., & Yang, C. (2019). Investigation of a promoted You Only Look Once algorithm and its application in traffic flow monitoring. *Applied Sciences, 9*(17), 3619. https://doi.org/10.3390/app9173619

Chao, Y., Liu, Y., Liu, X., Zeng, H., & Deng, J. (2018, March). Learning to detect human-object interactions. *Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision, Lake Tahoe, Nevada, USA,* 381-389. https://doi.org/10.1109/wacv.2018.00048

Chen, L., Hermans, A., Papandreou, G., Schroff, F., Wang, P., & Adam, H. (2018, June). MaskLab: Instance segmentation by refining object detection with semantic and direction features. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Lake Tahoe, Nevada, USA,* 4013-4022. https://doi.org/10.1109/cvpr.2018.00422

Chen, X., Xu, J., & Yu, Z. (2019). A 68-mw 2.2 tops/w low bit width and multiplierless DCNN object detection processor for visually impaired people. *IEEE Transactions on Circuits and Systems for Video Technology, 29*(11), 3444-3453. https://doi.org/10.1109/tcsvt.2018.2883087

Chen, Y., Li, W., Sakaridis, C., Dai, D., & Van Gool, L. (2018, June). Domain adaptive faster R-CNN for object detection in the wild. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Lake Tahoe, Nevada, USA,* 3339-3348. https://doi.org/10.1109/cvpr.2018.00352

Deekshitha, P., Reddy, B. T., Badadha, S., Dhruthi, K. K., & Pandiaraj, S. (2018). Object motion perception and tracking using sift with k-means clustering. *Journal of Network Communications and Emerging Technologies*, *8*(4), 242-246. https://www.jncet.org/Manuscripts/Volume-8/Issue-4/Vol-8-issue-4-M-55.pdf

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, Li. (2009, June). ImageNet: A large-scale hierarchical image database. *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, *Miami, Florida, USA*, 248-255. https://doi.org/10.1109/cvpr.2009.5206848

Deshpande, S., & Shriram, R. (2016, September). Real time text detection and recognition on hand held objects to assist blind people. *Proceedings of the 2016 International Conference on Automatic Control and Dynamic Optimization Techniques, Pune, India,* 1020-1024. https://doi.org/10.1109/icacdot.2016.7877741

Ding, K., Li, X., Guo, W., & Wu, L. (2022, January). Improved object detection algorithm for drone-captured dataset based on yolov5. *Proceedings of the 2022 2nd International Conference on Consumer Electronics and Computer Engineering, Guangzhou, China,* 895-899. https://doi.org/10.1109/iccece54139.2022.9712813

Ding, P., Wang, F., Gu, D., Zhou, H., Gao, Q., & Xiang, X. (2018, July). Research on optimization of SURF algorithm based on embedded CUDA platform. *Proceedings of the 2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems, Tianjin, China,* 1351-1355. https://doi.org/10.1109/cyber.2018.8688294

Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. (2009). The Pascal Visual Object Classes (VOC) challenge. *International Journal of Computer Vision*, *88*(2), 303-338. https://doi.org/10.1007/s11263-009-0275-4

Geng, Z. X., & Qiao, Y. Q. (2017, October). An improved illumination invariant SURF image feature descriptor. *Proceedings of the 2017 International Conference on Virtual Reality and Visualization, Zhengzhou, China,* 389-390. https://doi.org/10.1109/icvrv.2017.00090

Hosozawa, K., Wijaya, R. H., Linh, T. D., Seya, H., Arai, M., Maekawa, T., & Mizutani, K. (2018). Recognition of expiration dates written on food packages with open source OCR. *International Journal of Computer Theory and Engineering*, *10*(5), 170-174. https://doi.org/10.7763/ijcte.2018.v10.1220

Hu, H., Gu, J., Zhang, Z., Dai, J., & Wei, Y. (2018, June). Relation networks for object detection. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, *Salt Lake City, Utah, USA,* 3588-3597 https://doi.org/10.1109/cvpr.2018.00378

Jabnoun, H., Benzarti, F., & Amiri, H. (2015, December). Object detection and identification for blind people in video scene. *Proceedings of the 2015 15th International Conference on Intelligent Systems Design and Applications, Marrakech, Morocco,* 363-367. https://doi.org/10.1109/isda.2015.7489256

Karami, E., Prasad, S., & Shehata, M. (2015). Image matching using SIFT, SURF, BRIEF and ORB: Performance comparison for distorted images. *arXiv:1710.02726*. https://doi.org/10.48550/arXiv.1710.02726

Karthi, M., Muthulakshmi, V., Priscilla, R., Praveen, P., & Vanisri, K. (2021, September). Evolution of YOLO-V5 algorithm for object detection: Automated detection of library books and performance validation of dataset. *Proceedings of the 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems, Chennai, India,* 1-6. https://doi.org/10.1109/icses52305.2021.9633834

Keel, S., Müller, A., Block, S., Bourne, R., Burton, M. J., Chatterji, S., He, M., Lansingh, V. C., Mathenge, W., Mariotti, S., Muirhead, D., Rabiu, M. M., Ravilla, T. D., Resnikoff, S., Silva, J. C., Tapply, I., Vos, T., Wang, N., & Cieza, A. (2021). Keeping an eye on eye care: Monitoring progress towards effective coverage. *The Lancet Global Health*, *9*(10), e1460-e1464. https://doi.org/10.1016/s2214-109x(21)00212-6

Kim, K., Kim, P., Chung, Y., & Choi, D. (2018, November). Performance enhancement of YOLOv3 by adding prediction layers with spatial pyramid pooling for vehicle detection. *Proceedings of the 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance, Auckland, New Zealand,* 1-6. https://doi.org/10.1109/avss.2018.8639438

Kim, Y., & Jung, H. (2018). Reconfigurable hardware architecture for faster descriptor extraction in SURF. *Electronics Letters*, *54*(4), 210-212. https://doi.org/10.1049/el.2017.3133

Kumar, N., & Jain, A. (2021, December). Smart navigation detection using deep-learning for visually impaired person. *Proceedings of the 2021 IEEE 2nd International Conference on Electrical Power and Energy Systems, Bhopal, India,* 1-5. https://doi.org/10.1109/icepes52894.2021.9699479

Li, B., Munoz, J. P., Rong, X., Chen, Q., Xiao, J., Tian, Y., Arditi, A., & Yousuf, M. (2019). Vision-based mobile indoor assistive navigation aid for blind people. *IEEE Transactions on Mobile Computing, 18*(3), 702-714. https://doi.org/10.1109/tmc.2018.2842751

Liem, H. D., Minh, N. D., Trung, N. B., Duc, H. T., Hiep, P. H., Dung, D. V., & Vu, D. H. (2018, January). FVI: An end-to-end Vietnamese identification card detection and recognition in images. *Proceedings of the 2018 5th NAFOSTED Conference on Information and Computer Science, Ho Chi Minh City, Vietnam,* 338-340. https://doi.org/10.1109/nics.2018.8606831

Lin, T-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars, T. (Eds.) *Computer Vision – ECCV 2014*, 740-755. Springer. https://doi.org/10.1007/978-3-319-10602-1_48

Liu, Z., Luo, Y., Cordero, J., Zhao, N., & Shen, Y. (2016, June). Finger-eye: A wearable text reading assistive system for the blind and visually impaired. *Proceedings of the 2016 IEEE International Conference on Real-time Computing and Robotics, Angkor Wat, Cambodia,* 123-128. https://doi.org/10.1109/rcar.2016.7784012

Maolanon, P., & Sukvichai, K. (2018, December). Development of a wearable household objects finder and localizer device using CNNs on Raspberry Pi 3. Proceedings of the *2018 IEEE International WIE Conference on Electrical and Computer Engineering, Chonburi, Thailand,* 25-28. https://doi.org/10.1109/wiecon-ece.2018.8782924

Mocanu, B., Tapu, R., & Zaharia, T. (2018). Deep-see face: A mobile face recognition system dedicated to visually impaired people. *IEEE Access, 6,* 51975-51985. https://doi.org/10.1109/access.2018.2870334

Najibi, M., Singh, B., & Davis, L. S. (2019, June). FA-RPN: Floating region proposals for face detection. *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA,* 7715-7724. https://doi.org/10.1109/cvpr.2019.00791

Park, J., Yun, Y., Eun, S., Cha, S., So, S., & Jung, J. (2018). Deep neural networks based user interface detection for mobile applications using symbol marker. *Proceedings of the 2018 Conference on Research in Adaptive and Convergent Systems*, 66-67. https://doi.org/10.1145/3264746.3264808

Rajesh, M., Rajan, B. K., Roy, A., Thomas, K. A., Thomas, A., Tharakan, T. B., & Dinesh, C. (2017). Text recognition and face detection aid for visually impaired person using raspberry PI. *Proceedings of the 2017 International Conference on Circuit, Power and Computing Technologies,* 1-5. https://doi.org/10.1109/iccpct.2017.8074355

Redmon, J., & Farhadi, A. (2017, July). YOLO9000: Better, faster, stronger. *Proceedings of the* 2017 *IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA,* 6517-6525. https://doi.org/10.1109/cvpr.2017.690

Redmon, J., & Farhadi, A. (2018). YOLO-v3: An incremental improvement. *arXiv:1804.02767*. https://doi.org/10.48550/arXiv.1804.02767

Rodríguez, A., Yebes, J. J., Alcantarilla, P., Bergasa, L., Almazán, J., & Cela, A. (2012). Assisting the visually impaired: Obstacle detection and warning system by acoustic feedback. *Sensors, 12*(12), 17476-17496. https://doi.org/10.3390/s121217476

Sharif, M., Khan, S., Saba, T., Raza, M., & Rehman, A. (2019, April). Improved video stabilization using SIFT-log polar technique for unmanned aerial vehicles. *Proceedings of the 2019 International Conference on Computer and Information Sciences, Sakaka, Saudi Arabia,* 1-7.. https://doi.org/10.1109/iccisci.2019.8716427

Song, Q., Li, S., Bai, Q., Yang, J., Zhang, X., Li, Z., & Duan, Z. (2021). Object detection method for grasping robot based on improved YOLOv5. *Micromachines, 12*(11), 1273. https://doi.org/10.3390/mi12111273

Takatori, N., Nojima, K., Matsumoto, M., Yanashima, K., & Magatani, K. (2006, August-September). Development of voice navigation system for the visually impaired by using IC tags. *Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, *New York, NY, USA,* 5181-5184. https://doi.org/10.1109/iembs.2006.260437

Tareen, S. A., & Saleem, Z. (2018, March). A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK. *Proceedings of the 2018 International Conference on Computing, Mathematics and Engineering Technologies, Sukkur, Pakistan*, 1-10. https://doi.org/10.1109/icomet.2018.8346440

Wang, H., Katzschmann, R. K., Teng, S., Araki, B., Giarre, L., & Rus, D. (2017, May-June). Enabling independent navigation for visually impaired people through a wearable vision-based feedback system. *Proceedings of the 2017 IEEE International Conference on Robotics and Automation, Singapore*, 6533-6540. https://doi.org/10.1109/icra.2017.7989772

Wang, R., Shi, Y., & Cao, W. (2019). GA-SURF: A new speeded-up robust feature extraction algorithm for multispectral images based on geometric algebra. *Pattern Recognition Letters*, *127*, 11-17. https://doi.org/10.1016/j.patrec.2018.11.001

Wang, T., Huang, H., Lin, J., Hu, C., Zeng, K., & Sun, M. (2018). Omnidirectional CNN for visual place recognition and navigation. *2018 IEEE International Conference on Robotics and Automation (ICRA)*. https://doi.org/10.1109/icra.2018.8463173

Widyastuti, R., & Yang, C.-K. (2018). Cat's nose recognition using You Only Look Once (Yolo) and Scale-Invariant Feature Transform (SIFT). *Proceedings of the 2018 IEEE 7th Global Conference on Consumer Electronics, Nara, Japan*, 55-56. https://doi.org/10.1109/gcce.2018.8574870

Wu, Z., Zhang, D., Shao, Y., Zhang, X., Zhang, X., Feng, Y., & Cui, P. (2021, August). Using YOLOv5 for garbage classification. *Proceedings of the 2021 4th International Conference on Pattern Recognition and Artificial Intelligence, Yibin, China*, 35-38. https://doi.org/10.1109/prai53619.2021.9550790

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010, June). SUN database: Large-scale scene recognition from abbey to zoo. *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA*, 3485-3492. https://doi.org/10.1109/cvpr.2010.5539970

Xingteng, J., Xuan, W., & Zhe, D. (2015, October). Image matching method based on improved SURF algorithm. *2015 IEEE International Conference on Computer and Communications, Chengdu, China*, 142-145. https://doi.org/10.1109/compcomm.2015.7387556

Zhang, T., Zhang, B., Zhao, F., & Zhang, S. (2022, April). COVID-19 localization and recognition on chest radiographs based on Yolov5 and EfficientNet. *Proceedings of the 2022 7th International Conference on Intelligent Computing and Signal Processing, Xi'an, China*, 1827-1830. https://doi.org/10.1109/icsp54964.2022.9778327

Zheng, L., Yang, Y., & Tian, Q. (2018). SIFT meets CNN: A decade survey of instance retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(5), 1224-1244. https://doi.org/10.1109/tpami.2017.2709749

Zhu, D., Li, T., Ho, D., Zhou, T., & Meng, M. Q. (2018, October). A novel OCR-RCNN for elevator button recognition. *Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems, Madrid, Spain*, 3626-3631. https://doi.org/10.1109/iros.2018.8594071

AUTHORS

Nitin Kumar is currently working as an Assistant Professor in the School of Electronics and Electrical Engineering, Lovely Professional University, Phagwara, Punjab, India He has completed his MTech from IIIT-Allahabad and currently is pursuing his PhD in the field of Robotics and Control from Lovely Professional University. He has over 12 years of experience in Academics and Research and Development. His research interests are in the fields of Robotics, Machine Learning, and Artificial Intelligence. He has attended multiple international conferences and published numerous papers in international journals.

Dr. Anuj Jain is currently working as a Professor in the School of Electronics and Electrical Engineering, Lovely Professional University, Phagwara, Punjab, India He has completed his PhD in the field of Wireless Communication from Mewar University, Rajasthan, India. He has over 17 years of experience in Academics and Research and Development. His research interests are in the fields of Robotics, Machine Learning, and Wireless Networks.