Object Localization for Impaired Vision using Artificial Intelligence

Karan Banker, Prince Shrivastava, M. Vaidhehi Rakesh

Abstract: This paper will suggest object localization module, an artificial intelligence powered driven system modeled to help visually challenged individuals or a person suffering from dementia. This model will help them in their daily routine tasks in locating misplaced objects. To accomplish this task we are going to implement artificial intelligence based techniques such as speech recognition, generating speech, object detection and image processing. This will help the system to understand the user's request and will respond accordingly. This system consists of four major units: a) a speech unit, b) a image processing unit, c) an object detection unit, and d) a logic unit. The speech unit interacts with user and will listen to user's verbal query and verbally answers to the user. The image processing unit processes the image with the help of deep learning modules. The object detection unit detects all the target and non-target objects in a scene. Then, the logic unit sends the object location and description to the speech unit. This will help us to create most usable system to help visually impaired people.
Keywords: Localization Impaired, Artificial Intelligence.

I. INTRODUCTION

As per surveys and statistics there are about 250 million individuals who are visually challenged and 35 million are completely blind and rest have distorted vision. Blindness is the complete or partial loss of vision. It is the loss of useful sight. It can be temporary or permanent. There are various ongoing research to develop assistants for visually impaired people. For them even performing menial daily routine tasks like navigation, identifying hurdles in the way, identifying and pin pointing the object's location requires advance knowledge of the surrounding. Utilizing the recent breakthroughs in the field of artificial intelligence, especially in the speech recognition area and object detection, AI can not only support them but also increase their independence in their day to day life. In order to handle this crucial problem encountered by individuals with visual disabilities, we introduce object localization system to counter this problem with the help of artificial intelligence-based techniques- speech recognition, object detection. We are going to guide the user to their misplaced item without physical contact with the device. Adding to that, this system also helps them to reduce their job of memorizing the location of each item.

Karan Banker*, Student, Bachelor's in technology with specialization in Computer Science Engineering from SRM Institute of Science and Technology

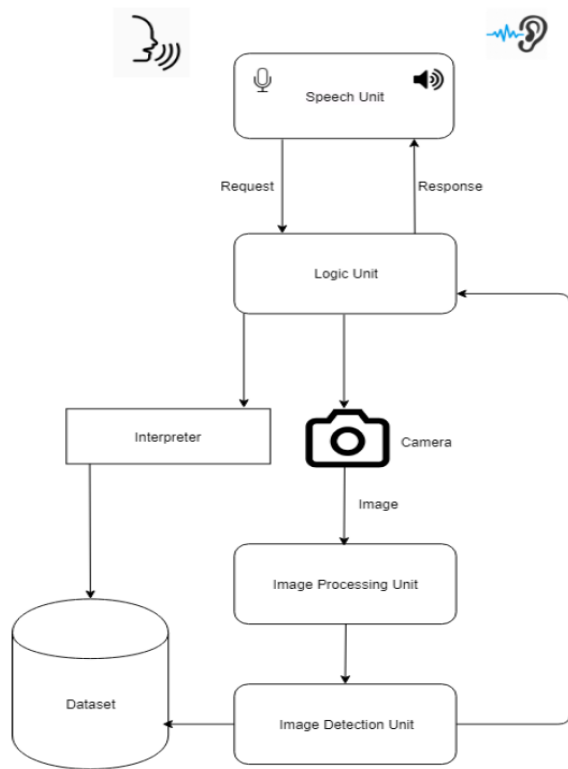**Prince Shrivastava,** Student, Bachelor's in technology with specialization in Computer Science Engineering from SRM Institute of Science and Technology

Mrs. M. Vaidhehi Rakesh, professor in SRM Institute of Science and Technology.

There has been various ongoing research in order to reduce the displeasure and uncertain situations when traveling to some unknown areas, such as system that navigates and detects obstacles. Some research suggests that combining computer vision and speech models helps to educate children in order to identify various objects without guidance. Our method can be served as a useful tool in applications such as baggage inspection, where the identification of threat objects is a matter of utmost importance.

Object detection systems are directed by the success of region proposal method and convolutional neural networks. These methods depends upon low-cost features and economical schemes. In order to detect objects, we need to provide the target image of the scene. That scene can be complex and the detected targets can be texture-less. In this case, feature points are not enough to express the results. So, feature-based detection approach cannot be applied. For these issues, we use the template matching approach, as it is more efficient with much less computational complexity and it doesn't require much time on training. Therefore, this method is suitable for the situations where the background scene is complex and texture-less.

We are going to use speech recognition technique in order to accept the user's request and to respond them verbally. Speech recognition involves the process of conversion of an acoustic waveform into text which is similar to the information provided by the user. Many researchers all-around the globe are trying to develop a powerful and noise tolerant speech recognition system. Pattern matching as well as feature extraction techniques plays a vital role in speech recognition system in order to recognize the speech of various persons. This helps to make a better and efficient system for recognizing speech. For this, we are going to use google speech API i.e. gTTS.

II. EASE OF USE

This will be the most usable system for visually impaired people to aid them in their daily routine tasks. As there will be no physical contact of user with the system units needed. The interaction between the user and system will be communicated through speech module. The speech module is responsible for both handling user request and providing audio response.

System Overview Diagram

III. SYSTEM OVERVIEW

A. Speech Recognition Unit

The speech unit is accountable for understanding verbal request from the user and gives the verbal response. This response has the information about the direction of misplaced item.

We will implement speech module provided by Python's library gtts (Google text-to-speech). As gtts provides humanlike response as the output. We will use speech API in order to translate the input speech by a user into the text that can be interpreted by other units.

Due to widespread adoption of these speech assistance on devices makes this a very reliable system. The main mode of communication for visually impaired people is their voice. As per the surveys, some visual impaired individuals already use voice assistance gadgets because of the simplicity of these device. For example, you can switch on the room lights using these voice assistant devices or you can set up the alarms and many other tasks.

B. Object Detection Unit

The main objective of this object detection unit is to recognize all the objects that are in particular scene. In order to accomplish this task we should have a proper dataset which contains all those objects information that can be found in the user's surrounding. Then, we need to train our model with all the details of the user's environment. For example, we need to feed the model with the details of some objects that can be act as a landmark for that particular scene like refrigerator.

Previously, object detection consists of two steps i.e., dividing the image into multiple sub parts and then passing the sub parts into an image classifier which predicts whether the image contains an object or not. If yes, then classify it as the detected object in the original image.
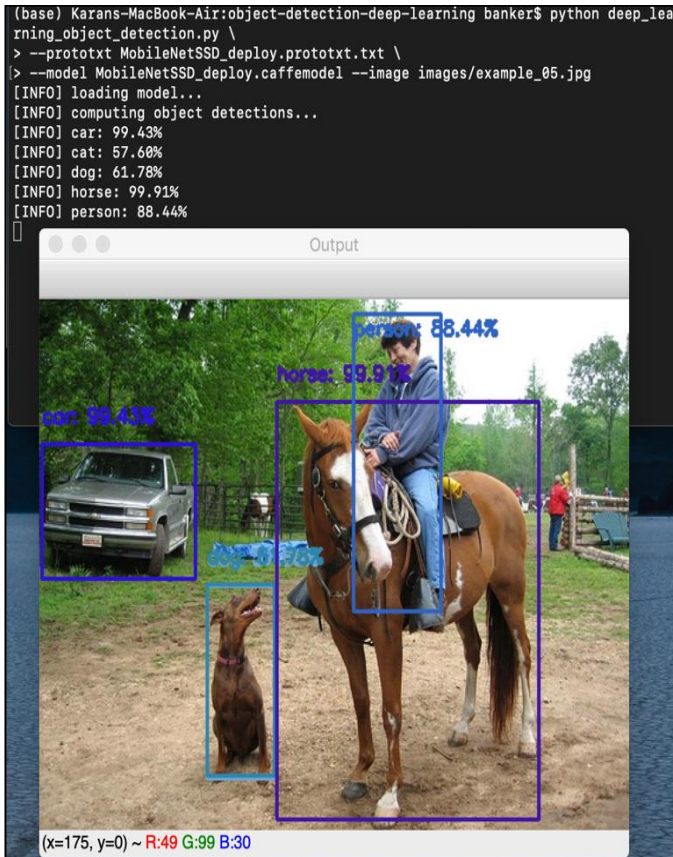
We will be using MS-COCO (Microsoft Common Object in Context) dataset. This is very rich and large in terms of object detection, segmentation and labeling a dataset. This dataset contains about 330K images with having 1.5 million instances of object over 80 object categories. This dataset is used because of its high-quality data and it is widely used for objection than any other dataset.

Object detection is predicting the location of the object along with the class. As we aware of CNN which has three layers first is input layer, second is hidden layer and third one is output layer. The input layer receives the input of the different portions of the image and then it passes them to the hidden layer which consists of different functions and filters, such as color filter, size filter, shape filter and so on. This is where the main magic happens, different CNN (convolutional neural network) based techniques like R-CNN, Fast R-CNN, Faster R-CNN, Grid based CNN and so on are implemented. Finally, after going through many functions and filters it produces the output with the highest probability. So, for this instead of using the CNN based region propositional network that utilizes the selective search approach which works fine but the problem doesn't get solved with this, we urge to decrease the response time of the system so that if a user gives a query to the system then it can get a response as fast as possible. So, we present an idea of object detection system which depends on position-sensitive grid feature maps which are quick to process and highly responsive. In this, we build position-sensitive convolutional layers, called Grid Convolutional Layers (GCLs), which activates the object's particular locations in the feature maps in grid form. With the proper training, the RoI (reason of interest) grid pooling layer guides the last set of convolutional layers to learn specialized grid feature maps.

Dimensions	Selective Search	Grid Based
3269*2497	4.98 seconds	0.55 seconds
2560*1440	28.99 seconds	1.11 seconds
1280*720	62.6 seconds	1.98 seconds

Execution Time Comparison

As we can infer from the above table that the response time of Grid Based approach is way faster than the Selective Search approach. Because the selective search approach emphasized on a hierarchical grouping-based segmentation algorithm.

```
(base) Karans-MacBook-Air:object-detection-deep-learning banker$ python deep_lea
rning_object_detection.py \
> --prototxt MobileNetSSD_deploy.prototxt.txt \
> --model MobileNetSSD_deploy.caffemodel --image images/example_05.jpg
[INFO] loading model...
[INFO] computing object detections...
[INFO] car: 99.43%
[INFO] cat: 57.60%
[INFO] dog: 61.78%
[INFO] horse: 99.91%
[INFO] person: 88.44%
```

C. Logic Unit

The objective of this logic unit is to handle the interaction between the speech unit and the object recognition unit. This helps to interpret the user's wish and it will develop the logic description based on the user's query. There are various function performed by this unit – it is used to initiate the camera to capture the image so that the object detection unit can identify all the objects of that scene. It is also used to determine the target object based on the user's query passed by the speech unit. Then finally, it helps to develop a logic description. When this description gets generated it forwards it to speech unit to process the output to the user.

This logic unit has a response builder that is used to find the closest landmark object in respect to the target object based on the user's query. Then, it calculates the relative location of that object and generate the response for the user.

IV. RELATED WORK

In [1] Viraj Shinde, Tushar Bacchav, Jitendra Pawar, Mangesh Sanap have stated the usage of recognizing hand gestures using a web camera various algorithms and gestures recognition techniques. Some of algorithms used are Point Pattern Matching algorithm and Image processing and object recognition. In this it provides High Portability as it reduces External Interface like mouse and keyboard. It only requires web-camera to capture I/P image and no physical contact with device is needed.

In [2] Ren, Shaoqing has implemented Real-Time Object Detection using Region Proposal Networks Algorithm. PASCAL VOC dataset has been used. To improve the object detection RPNs (Region Proposal Networks) with Fast R-CNNs is used as region proposal step is nearly cost-free and RPN improves overall object detection accuracy.

In [3] Saksamudre, Suman K., P. P. Shrishrimal, and R. R. Deshmukh throws light on different approaches for speech recognition. Some of the strategies used are Feature Extraction method and Pattern classification. They have classified speech recognition system over various categories based on utterances, speaker model and vocabulary.

In [4] Jiao Jichao presented an algorithm low texture object in a complex environment using POD (Fast template matching algorithm). Oxford dataset and lab based data set has been used. Template matching method to performed to extract target edges. It requires less computational complexity and it does not need to spend much time on training.

In [5] Wu, Yuanhao, Ze Wu and Chenglong Fu have proposed Continuous Arm Gesture Recognition. Strategies used are Gesture modelling and recognition and Logistic regression. The IMU dataset with 54024 data frames and 1628 gesture instances is used. In this both static gestures and dynamic gestures are modelled. This provides Gesture models to have good comprehensibility.

In [6] Riffo, Vladimir, and Domingo Mery have worked on Automated Detection of Threat Objects. Target object characterization and detection strategy has been used. The dataset contains X-ray images of 200 bags. The relevance lies in the detection of threat objects using X-ray images and recognizing objects based on implicit shape models. The Effectiveness and robustness of this method is for the detection of different threat objects.

In [7] Ton, Carolyn has implemented Spatial Sensing for the Visually Impaired people. They used the strategy called Light detection and ranging (LIDAR) assist spatial sensing system through sensors. Its relevance lies in the Obstacle identification, enhanced spatial perception of surrounding areas. This provides people who are visually challenged of their surrounding spatial information through stereo sound at varying pitch.

In [8] Poularakis, Stergios, and Ioannis Katsavounidis throws some light on Gesture Recognition System for Continuous Streams of Digits and Letters. They used the algorithm of maximum cosine similarity and fast nearest neighbour (NN) techniques. 2-D Graffiti [11], Kinect [13] and 6DMG [12] dataset has been used. Its relevance lies in the Trajectory classification of digits and letters. It is a promising approach for implementations on low-power embedded systems.

In [9] Kurzekar, Pratik K worked on Continuous Speech Recognition System. They have also focused on languages like Korean, Dutch, Tamil, Telugu, Marathi and Arabian other than English. Some of the datasets used are DARPA, TIMIT, VIOS and so on. Its relevance lies to develop continuous speech system based on large vocabulary and to increase the accuracy of such system. As it very advantageous as this system developed in different languages around the world.

In [10] Islam Md Milon has thrown some light over Development of Walking Assistants for Visually Impaired People. It is sensor, computer vision and smartphone based walking assistant. The ADE20K dataset has been taken in use.

It generates an audio signal or vibration with the presence of obstacles for indoor and outdoor surroundings. Its advantage is that this can be incorporated in any framework that can assist visually impaired people.

In [11] Balasuriya B. K. has worked to create a Learning Platform which can assist Visually Impaired Children. It has two important module namely Object Recognition Module and Speech Recognition Module. MS COCO having 400K image caption dataset has been used. In this Deep Region Based Convolutional Networks, Recurrent Neural Networks and Speech models are used to provide an interactive learning experience. This proves very beneficial to visually challenged people Learning to identify objects without the supervision of any third-party.

In [12] Wang, Linda worked on Assistant for Object Localization for Impaired Vision. They proposed AI-based speech recognition, speech generation, and object detection. MS COCO, ILSVRC dataset has been taken in use. Its relevance is in the OLIV which is voice-driven and thus omits the use of hands and is unobtrusive. A more portable hands-free solution and feasibility of OLIV as a useful aid for individuals with impaired vision.

In [13] Dekel, Tali have worked on Robust Template Matching. They have worked on Best-Buddies Similarity Algorithm and improved it to some extent. It is robust against complex geometric deformations and high levels of outliers. But one has to take into consideration that Method may fail when the template is very small compared to target image.

In [14] Ye, Cang, and Xiangfei Qian have put their efforts in 3-D Object Recognition, Data acquisition and plane extraction. The proposed method is robust to the object's visual appearance change. It is ideal for detecting structural objects. It has high scalability and parallelism.

In [15] Auti, Aishwarya have worked on Sign-Talk with Hand Gesture Recognition. Some of the techniques and strategies used are Data Procurement, Feature Extraction, Gesture Classification and Speech Output. This will help to bridge the gap between the impaired people and normal people. This method use to recognize static sign language gestures however, the system can also be expanded to include dynamic gestures.

V.   PROPOSED SYSTEM

A. Speech Recognition Technique

We need speech because it is the most common means of communication around the world and majority of the population around the world relies on speech to communicate with each another. This overcomes the barrier of literacy so that both literate and illiterate people can make use of it.

As far as the speech unit is concerned it will serve two purposes one is that it will convert the speech (query by the user to system in the form of speech) to text and responding to those queries in the speech (converting text to speech) back to the user.

One of the major challenges faced by the speech recognition systems is that there are different types of languages spoken around the world and the usage of same language may differ from a child to grown up person. The way a child speaks a word may differ from the way it is spoken by an older person. This is taken care by the google speech API. The google speech API supports around 120 different languages around the world.

We will be implanting speech module with the help google speech API. The speech recognition system translates the spoken utterances to text. Using the recognizer class the speech will be converted to text and will be forwarded to the system for further process. We will be using microphone to take speech input, it will be served by the pyAudio package in python.

For converting back the text to use we will be using gTTS which is google Text To Speech which is a very simple and easy to use tool that can convert the text into audio file of mp3 format. The gTTS is very useful because it supports various languages like Hindi, English, Tamil, German and many more. This will help to build a more generic system that can be used by most of the people around the globe.

The google speech API plays an important role in this system. It currently can identify around four languages itself based on the utterances. It does real-time translation, as soon as the user speaks the unit translates it back to speech and the times come to produce the output to user in the form of speech to text it does this job well too.

We are using google speech API because no other APIs are available such as Microsoft Cognitive Services, IBM Watson, Final Thoughts and so on. It is because as google is almost the backbone of today's era and being its product it is no wonder that this API is popular as well. One of the reasons it catches attention is due to its accuracy and ability to adapt different machine learning models as per the required system.

B. Datasets

There are number of datasets available for object recognition like MS COCO (Microsoft Common Objects in Context), ImageNet, YOLO, CIFAR-100, core50, etc. The best performing algorithms usually consider COCO detection dataset. The recent one is YOLO dataset and it actually outperformed all.

MS COCO (Microsoft Common Objects in Context) is a rich dataset of more than 300K images and which will be quite beneficial to use. Its wide range of images gives opportunities to different models which are developed for object detection. This dataset contains a wide variety of images like train, cars, bottles and so on.

It has some various features which makes it a very suitable dataset like it has more than 300K of images out of which more than 200K images are labelled. It has around 1.5 million instances of objects, 80 different object categories, about 90 stuff categories, 5 captions per image and 250K people with key points.

C. Convolutional Neural Network (CNN)

For any image basically there are three channels, one will be red, another will be green and finally blue channel (which is popularly known as RGB). For each of these channel, they have their own respective pixel values. So, when I say the image size is B*A*3, that means that there are B rows, A columns and 3 channels. Eg.28*28*3 (28 rows, 28 columns, 3 channels). So, this is how a computer sees an image. This is for coloured images.

For black and white images, there are only two channels.

We do not use Fully Connected Networks for Image Recognition. Consider an image which has 28*28*3 pixels and when we feed this image to a fully connected network then the total number of weights required in the first hidden layer will be 2352. But in real life, the images are not that small. So, the images that we have, has generally 200*200*3 pixels and feed it to the fully connected network. So, at that time the number of weights required at the first hidden layer itself will be 120,000. So, we need to deal with such huge amount of parameters and obviously it requires more number of neurons so that can eventually lead to overfitting. So, that's why we can't use Fully Connected Network for Image classification.

Convolutional Neural Network are special type of feed-forward artificial neural network which are inspired from visual cortex. Visual cortex is nothing but a small regions of cells in our brain that are sensitive to specific regions of visual field. In a convolutional neural network, a neuron in the layer will only be connected to a small region of the layer before it. Hence, we need to handle less amount of weights and in turn we need less number of neurons as well. That's why we need convolutional neural network. Convolutional Neural Network has various layers-convolutional layer, ReLU layer, pooling layer, and fully connected layer.

For example- to classify an alphabet, that alphabet can be represented in multiple forms. Those forms are nothing but the deformed images of that alphabet. But we need to classify all those deformed images to its original alphabet image.

We know that computer understands an image using numbers at each pixels. So, just consider a matrix and draw an alphabet in that. The blocks which are empty, mark it as -1 and assign 1 to the remaining blocks (FIG). So, now with the help of CNN we take small patches of our image. These patches are known as features or filters. We will try to match all the feature of deformed image with the same position of that particular feature in original image one by one. Basically, we need to find out the similarity between the two images (original and deformed image). Just choose a feature (small pieces of bigger image) and put it on the input image, if it matches then the image is classified correctly.

a) Convolution Layer

This is the first layer. There are some steps for this particular layer. First, we need to line up the feature and the image and then multiply each image pixel by the corresponding feature pixel. After that we need to add them up and divide by the total number of pixels in the feature. Then, create a map and put the value of the filter at that particular place. Similarly, we are going to move this filter throughout the image and follow the same steps. Finally, we will get an output matrix for a particular feature.

(b) ReLU Layer

This layer is nothing but an activation function. That means, it will only activate a node if the input is above a certain quantity while the input is below zero the output is also zero. When the input rises above certain threshold it has a linear relationship with the dependent variable.

$$f(x) = \begin{cases} 0 \ if \ x < 0 \\ x \ if \ x \geq 0 \end{cases}$$

This function says that f(x)=0 if x<0 and it is equal to x when x>=0. So, this is how ReLU function works. We are using ReLU function here because we want to remove all the negative value from our output matrix that we got through convolution layer.
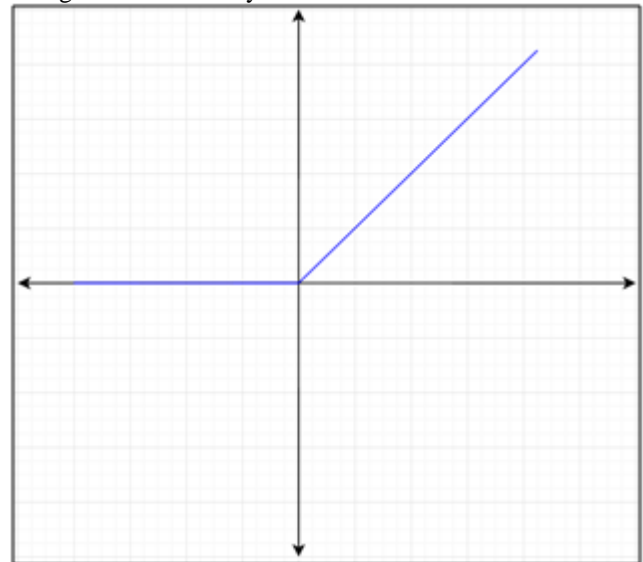
ReLU Function Graph

b) Pooling Layer

In this layer, we take window size of 2 or 3 and we move it across the entire matrix that we have got after passing through ReLU layer. While moving the window, choose only maximum value in order to shrink the image i.e. reducing the size of the image. This will the give the output matrix with less number of rows and columns.

Now, stack up all the layers. We can add, multiple layers of convolution, ReLU and pooling to shrink the image.

c) Fully Connected Layer

This is the final layer where the actual classification happens. Here, we are going to take the shrink images and put it into a single list. Basically, after stacking up all the layers convert the matrix into a single list or vector.

There are certain values in the list which is high. Just add those and note the position of those values in that list. Now, take an input image, pass it through all the layers. Once it has passed through all the layers we will get a list similar to the deformed image list. To classify it, we will compare this list with our deformed image lists. For comparing, take the input image list and add all the values of those corresponding positions which we noted earlier. Divide this value with the sum we got for a particular image. Do the same process for all the symbols in the dataset. Now, check for the highest value that we got and then classify the input image to that particular image class.

We are also going to use TensorFlow. TensorFlow is a computational framework used to build machine learning models. It provides a variety of different toolkits for construction of these models.

It is known for its high performance in working with deep learning models.
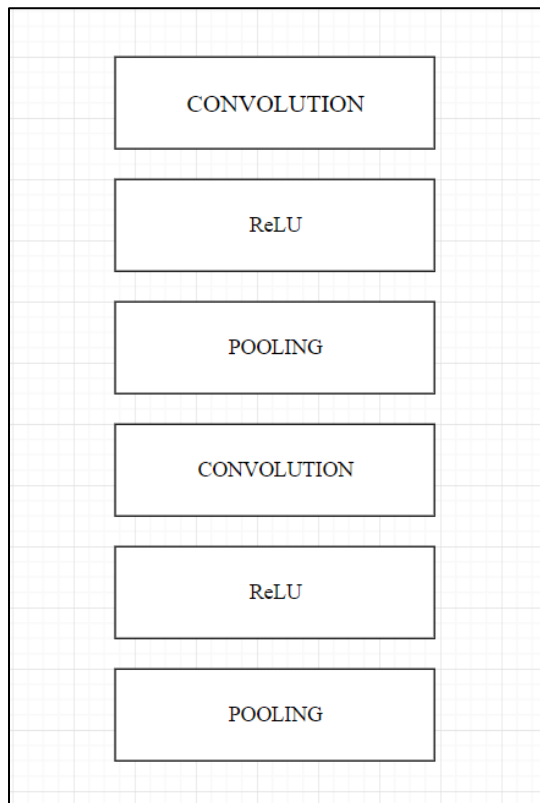
Stacking up all the layers

VI.   FUTURE SCOPE

First and foremost thing will be to remove language barrier to allow individuals speaking languages other than English can utilize this module and will get the response in the same language they can understand. For this we are going implement Google speech APIs.

Currently this system is applicable for indoor object recognition, so in future we will expand our module to work on outdoor environment as well. This can be achieved by distributing our module to different commercial handheld devices like watches, portable speakers and etc.

As per surveys and various articles, we came to know that various visually impaired individuals hurt themselves with some objects like knives, blades, etc. in an unknown environment. So, in order to tackle the above situation we are going to implement this model in order to recognize various objects that can be   proved dangerous to the user like knifes, blades, guns and we will label them as "threat objects". Whenever our module detects any threat objects in the user environment, it will alert the user about that object location and type. To achieve the above idea, we will classify items than can impose any threat as "threat objects" in our dataset.
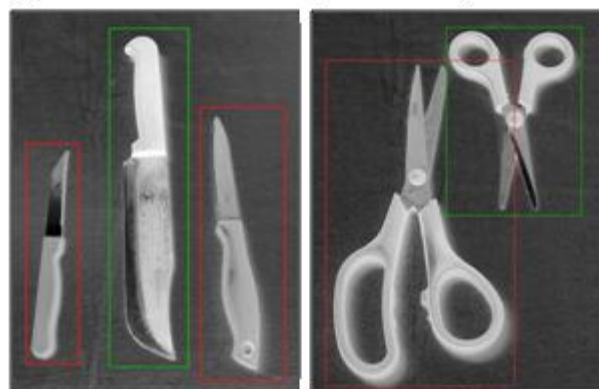
Identification of Threat Objects

This model can also be proved to be beneficial for helping the visually challenged individuals in navigation with the help of AI. So, if a user goes to some unfamiliar environment, the system can remember the pattern of that surrounding and will notify the user when the user visits the place again.

VII.   RESULTS

As a proof of concept, this system is built with the object detection module that is trained on a particular dataset. In this example, the user asks "where is my mouse?".

User Query	"where is my mouse"
Class	Mouse
Confidence	99.6%
Answer	"mouse is on the left of laptop"
	"mouse is on the right of laptop"

Based on the object detection module, the location of the mouse relative to the nearby objects is verbally conveyed to the user using speech module i.e., "mouse is on the right of laptop" and "mouse is on the right of cell phone".

VIII. COCNLUSION

The proposed work will help to eliminate all the problem faced by visually impaired people in their daily life. It only requires a camera and voice assistant device to help them in locating the displaced object. This system will lead to a new era of computer human interaction with the help of computer vision and artificial intelligence. In this model there will be no physical contact of user with any device is needed. Anyone can utilize this system for navigation, detection by simply using voice command.

As we know that dementia is a condition in which a person's memory, problem solving, language and other thinking skills gets disturbed which can affects a person's ability to perform their day to day tasks. Like, a person may not remember where they placed their items. So, this system can be utilized not only by visually impaired people but also those who are suffering from dementia. This system can inform them about the location of those objects.

We will implement this module with the help of high-level languages like C++ and Python. But we will try to expand our work in Python language as it is platform independent language.

REFERENCES

1. Wang, Linda, et al. "OLIV: An Artificial Intelligence-Powered Assistant for Object Localization for Impaired Vision." *Journal of Computational Vision and Imaging Systems* 4.1 (2018): 3-3.
2. Shinde, Viraj, et al. "Hand Gesture Recognition Using Web Camera." *International Journal of Advanced Engineering & Innovative Technology* 1.1 (2014).
3. Auti, Aishwarya, et al. "Sign-Talk: Hand gesture recognition system." *International Journal of Computer Applications* 160.9 (2017).
4. Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems*. 2015.
5. Saksamudre, Suman K., P. P. Shrishrimal, and R. R. Deshmukh. "A review on different approaches for speech recognition system." *International Journal of Computer Applications* 115.22 (2015).
6. Kurzekar, Pratik K., et al. "Continuous Speech Recognition System A Review." *Asian Journal of Computer Science and Information Technology* (2014): 62-66.
7. Jiao, Jichao, et al. "A fast template matching algorithm based on principal orientation difference." *International Journal of Advanced Robotic Systems* 15.3 (2018):
8. Dekel, Tali, et al. "Best-buddies similarity for robust template matching." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015.
9. Wu, Yuanhao, Ze Wu, and Chenglong Fu. "Continuous Arm Gesture Recognition Based on Natural Features and Logistic Regression." *IEEE Sensors Journal* 18.19 (2018)
10. Riffo, Vladimir, and Domingo Mery. "Automated detection of threat objects using adapted implicit shape model." *IEEE Transactions on Systems, Man, and Cybernetics: Systems*46.4 (2015): 472-482.
11. Ye, Cang, and Xiangfei Qian. "3-D object recognition of a robotic navigation aid for the visually impaired." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26.2 (2017): 441-450.
12. Islam, Md Milon, et al. "Developing walking assistants for visually impaired people: A review." *IEEE Sensors Journal*19.8 (2019): 2814-2828.
13. Ton, Carolyn, et al. "LIDAR Assist spatial sensing for the visually impaired and performance analysis." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26.9 (2018): 1727-1734.
14. Poularakis, Stergios, and Ioannis Katsavounidis. "Low-complexity hand gesture recognition system for continuous streams of digits and letters." *IEEE transactions on cybernetics* 46.9 (2015): 2094-2108.
15. Balasuriya, B. K., et al. "Learning platform for visually impaired children through artificial intelligence and computer vision." *2017 11th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*. IEEE, 2017.

AUTHORS PROFILE

Karan Banker, is a fourth-year student currently pursuing his Bachelor's in technology with specialization in Computer Science Engineering from SRM Institute of Science and Technology. He has completed courses like Artificial Intelligence, Deep Learning et cetera from Coursera. An avid learner, he looks forward to work in reputed company in the IT industry.

Prince Shrivastava, is a fourth-year student currently pursuing his Bachelor's in technology with specialization in Computer Science Engineering from SRM Institute of Science and Technology. He has completed courses like Machine Learning, Web Development using PHP et cetera from Coursera. An avid learner, he looks forward to work in reputed company in the IT industry.

Mrs. M. Vaidhehi Rakesh, has completed her master of engineering in Computer Science Engineering. Currently she is working as an assistant professor in SRM Institute of Science and Technology. She is having 7 years of teaching experience. Her research work includes machine learning, deep learning and Artificial Intelligence.