BLIND PEOPLE NAVIGATION SYSTEM USING OBJECT RECOGNITION

A MINI PROJECT REPORT

Submitted by

H.BHARANI (811519104016)

D.BHAVADARANI (811519104017)

A.DHIVYA (811519104024)

R.KARTHIKA (811519104052)

in partial fulfilment of the award of the degree of

BACHELOR OF ENGINEERING
in
COMPUTER SCIENCE AND ENGINEERING

K.RAMAKRISHNAN COLLEGE OF ENGINEERING
(AUTONOMOUS),
SAMAYAPURAM, TRICHY – 621 112.

JUNE 2022

K.RAMAKRISHNAN COLLEGE OF ENGINEERING
(AUTONOMOUS)
BONAFIDE CERTIFICATE

Certified that this project report **"BLIND PEOPLE NAVIGATION SYSTEM USING OBJECT RECOGNITION"** is the bonafide work of **H.BHARANI ,D.BHAVADARANI ,A.DHIVYA** and **R.KARTHIKA** who carried out the project under my supervision.

SIGNATURE
Dr. T.M. NITHYA, M.E., Ph.D.,
Associate Professor,
HEAD OF THE DEPARTMENT
Department of Computer Science
And Engineering,
K.Ramakrishnan College of
Engineering (Autonomous),
Samayapuram,
Trichy-621 112.

SIGNATURE
Mr. R. RAMESH,M.E.,M.B.A.,
Assistant Professor,
SUPERVISOR
Department of Computer Science
And Engineering,
K. Ramakrishnan College of
Engineering (Autonomous),
Samayapuram,
Trichy-621 112.

Submitted for the Project Viva-Voce Examination held on _____

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

We thank the almighty GOD, without whom it would not have been possible for us to complete our project.

We wish to address our profound gratitude to **Dr.K.RAMAKRISHNAN**, Chairman, K.Ramakrishnan College of Engineering(Autonomous), who encouraged and gave us all help throughout the course.

We express our hearty gratitude and thanks to our honorable and grateful executive director **Dr.S.KUPPUSAMY, B.Sc., MBA., Ph.D.,** K. Ramakrishnan College of Engineering(Autonomous).

We are glad to thank our principal **Dr.D.SRINIVASAN, M.E., Ph.D., FIE., MIIW., MISTE., MISAE., C.Engg.,** for giving us permission to carry out this project.

We wish to convey our sincere thanks to **Dr. T.M. NITHYA, M.E., Ph.D.,**Head of the Department,Computer Science and Engineering for giving us constants encouragement and advice throughout the course.

We are graceful to **Mr.R. Ramesh,M.E.,M.B.A.,** Assistant Professor in the Department, Computer Science and Engineering, K. Ramakrishnan College of Engineering(Autonomous), for his guidance and valuable suggestions during the course of study.

Finally we sincerely acknowledged in no less term for all our staff members , our parents and friends for their co-operation and help at various stages of this project work.

DECLARATION

I hereby declare that the project work entitled **"BLIND PEOPLE NAVIGATION SYSTEM USING OBJECT RECOGNITION"** is submitted in partial fulfillment of the requirement for the reward of the degree in B.E., is a record of our own carried out by me during the academic year 2021-2022 under the supervision and guidance of **Mr.R.RAMESH**,Assistant Professor,**Department of Computer Science and Engineering**,**K.Ramakrishnan College of Engineering (Autonomous)**.The extent and source of information are derived from the existing literature and have been indicated through the dissertation at the appropriate places. The matter embodied in this work is original and has not formed the part of any other work submitted for award if any degree or diploma, either in this or any other University.

H. BHARANI
(811519104016)

I certify that the declaration made above by the candidate is true.

Mr. R. RAMESH
Assistant Professor/CSE

DECLARATION

I hereby declare that the project work entitled **"BLIND PEOPLE NAVIGATION SYSTEM USING OBJECT RECOGNITION"** is submitted in partial fulfillment of the requirement for the reward of the degree in B.E., is a record of our own carried out by me during the academic year 2021-2022 under the supervision and guidance of **Mr.R.RAMESH**,Assistant Professor,**Department of Computer Science and Engineering**,**K.Ramakrishnan College of Engineering (Autonomous)**. The extent and source of information are derived from the existing literature and have been indicated through the dissertation at the appropriate places. The matter embodied in this work is original and has not formed the part of any other work submitted for award if any degree or diploma, either in this or any other University.

D. BHAVADARANI
(811519104017)

I certify that the declaration made above by the candidate is true.

Mr. R. RAMESH
Assistant Professor/CSE

DECLARATION

I hereby declare that the project work entitled **"BLIND PEOPLE NAVIGATION SYSTEM USING OBJECT RECOGNITION"** is submitted in partial fulfillment of the requirement for the reward of the degree in B.E., is a record of our own carried out by me during the academic year 2021-2022 under the supervision and guidance of **Mr.R.RAMESH**,Assistant Professor,**Department of Computer Science and Engineering**,**K.Ramakrishnan College of Engineering (Autonomous)**. The extent and source of information are derived from the existing literature and have been indicated through the dissertation at the appropriate places. The matter embodied in this work is original and has not formed the part of any other work submitted for award if any degree or diploma, either in this or any other University.

A.DHIVYA
(811519104024)

I certify that the declaration made above by the candidate is true.

Mr. R. RAMESH
Assistant Professor/CSE

DECLARATION

I hereby declare that the project work entitled **"BLIND PEOPLE NAVIGATION SYSTEM USING OBJECT RECOGNITION"** is submitted in partial fulfillment of the requirement for the reward of the degree in B.E., is a record of our own carried out by me during the academic year 2021-2022 under the supervision and guidance of **Mr.R.Ramesh**,Assistant Professor,**Department of Computer Science and Engineering**,**K.Ramakrishnan College of Engineering (Autonomous)**. The extent and source of information are derived from the existing literature and have been indicated through the dissertation at the appropriate places. The matter embodied in this work is original and has not formed the part of any other work submitted for award if any degree or diploma, either in this or any other University.

R. KARTHIKA
(811519104052)

I certify that the declaration made above by the candidate is true.

Mr. R. RAMESH
Assistant Professor/CSE

TABLE OF CONTENT

ABSTRACT

Visual impairment and blindness caused by various diseases has been hugely reduced, but there are many people who are at risk of age-related visual impairment. Visual information is the basis for most navigational tasks, so visually impaired people are at disadvantage because necessary information about the surrounding environment is not available. With the recent advances in inclusive technology it is possible to extend the support given to people with visual impairment during their mobility. In this context the proposed a system, named smart way of blind people navigation system, whose objective is to give blind users the ability to move around in unfamiliar environment, whether indoor or outdoor, through a user friendly interface. This paper is focused mainly in the development of the computer vision module of the Smart Vision system. The system adds the ANN (Artificial Neural Network) based frame detection system which detects the points and converts the system as voice synthesis for easy audibility of blind people. The ANN system will extract the descriptor features Of the system where each object gets analyzed with the co-ordinate value which is extracted from the frame sequence. The extracted features get matched with the machine learning database system and the recognition of the object is done. The ANN system produces the classified objects recognition with 98.7% accuracy system. The labeled data are get analyzed with HMM based voice synthesis system. The prototype device is tested in a situation simulating a blind people being exposed to a new environment. The smart system helps the blind to easily navigate all over the environment.

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

Millions of people live in this world with incapacities of understanding the environment due to visual impairment. Although they can develop alternative approaches to deal with daily routines, they suffer from certain navigation difficulties as well as social awkwardness. For example, it is very difficult for them to find a particular room in an unfamiliar environment. And blind and visually impaired people find it difficult to know whether a person is talking to them or someone else during a conversation.

1.1 ARTIFICIAL INTELLIGENCE

Machine learning is an application of artificial intelligence (Al) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

Artificial Intelligence is an approach to make a computer, a robot, or a product to think how smart human think. Al is a study of how human brain think, learn, decide and work, when it tries to solve problems. And finally this study outputs intelligent software systems.The aim of Al is to improve computer functions which are related to human knowledge, for example, reasoning, learning, and problem-solving.The intelligence is intangible. It is composed of

- Reasoning
- Learning
- Problem Solving
- Perception
- Linguistic Intelligence

The objectives of Al research are reasoning, knowledge representation, planning, learning, natural language processing, realization, and ability to move and manipulate objects. There are long-term goals in the general intelligence sector. Approaches include statistical methods, computational intelligence, and traditional coding Al. During the Al research related to search and mathematical optimization, artificial

neural networks and methods based on statistics, probability, and economics, we use many tools. Computer science attracts Al in the field of science, mathematics, psychology, linguistics, philosophy and so on.

1.2 OBJECT DETECTION AND RECOGNITION

Object detection is associated with Computer Vision and describes a system that can identify the presence and location of a desired object or body within an image. Do note that there can be singular or multiple occurrences of the object to be detected. The output of an object detection process is an image with bounding boxes around the objects of interest and an indication as to the class instance of a single object

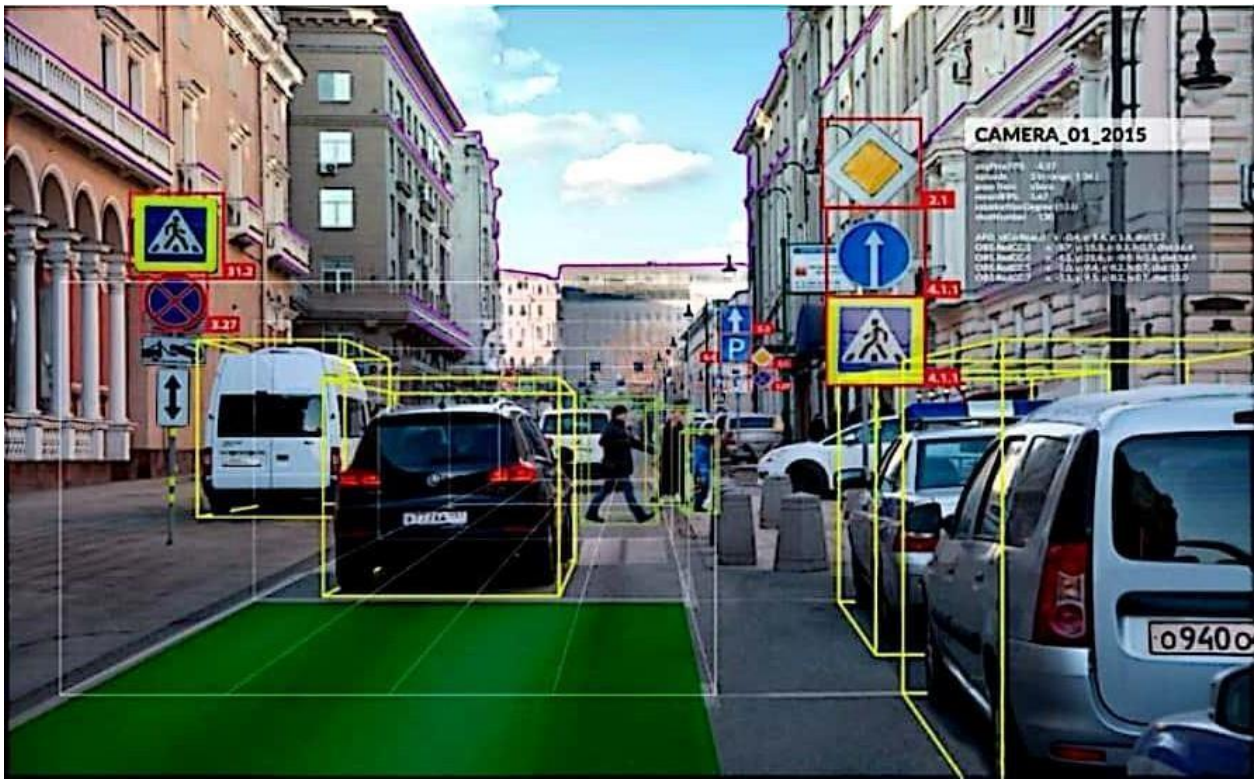


Fig 1.2 Object detection system using Artificial Intelligence

Other applications of Object Detection are:

- Face Detection: A term given to the task of implementing systems that can automatically recognize and localize human faces in images and videos. Face detection is present in applications associated with facial recognition, photography, and motion capture.

- Pose Estimation: The process of deducing the location of the main joints of a body from provided digital assets such as images, videos, or a sequence of images. Forms of pose estimation are

present in applications such as Action recognition, Human interactions, creation of assets for virtual reality and 3D graphics games, robotics and more

- Object Recognition: The process ofidentifying the class a target object is associated with. Objectrecognition and detection are techniques with similar end results and implementation approaches. Although the recognition process comes before the detection steps in various systems and algorithms.
- Tracking: A method of identifying, detecting, and following an object of interest within a sequence of images over some time. Applications of tracking within systems are found in many surveillance cameras and traffic monitoring devices.

1.2 COMPUTER VISION TECHNOLOGY

Computer vision technologies, especially the deep convolutional neural network, have been rapidly developed in recent years. It is promising to use the state-of-art computer vision techniques to help people with vision loss. This paper aims at exploring the possibility of using the hearing sense to understand visual objects. The sense of sight and hearing sense share a striking similarity: both visual object and audio sound can be spatially localized. It is not often realized by many people that we are capable at identifying the spatial location of a sound source just by hearing it with two ears. The aim of the work is to guide the blind people through the output of processor or controller by voice to navigate them. The methodology of this work includes Object Extraction, Feature Extraction, and Object Comparison.

1.3 OPTICAL CHARACTER RECOGNITION

There exist multiple tools to use computer vision technologies to assist blind people. The mobile app "Tap Tap See" uses computer vision and crowd sourcing to describe a picture captured by blind users In about 10 seconds. The Blind sight offers a mobile app Text Detective featuring optical character recognition (OCR) technology to detect and read text from pictures captured from the camera. However, these products were not focusing on enabling general visual sense for blind people and did not use the spatial sound techniques to further enhance the user experience.

1.4 VISUALLY IMPAIRED PEOPLE

Blindness is a state of lacking the visual perception due to neurological or physiological factors. The partial blindness represents the lack of integration in the growth of the optic visual or nerve centre of the eye, and total blindness is the full absence of the visual light perception. In this work, cheap, a simple friendly user, smart blind guidance system is designed and implemented to improve the mobility of both blind and visually impaired people in a specific area. The proposed work includes a wearable equipment consists of light weight blind stick and sensor based obstacle detection circuit is developed to help the blind person to navigate alone safely and to avoid any obstacles that may be encountered, whether fixed or mobile, to prevent any possible accident. The main component of this system is the infrared sensor which is used to scan a predetermined area around blind by emitting-reflecting waves.

Blind people face several problems in their life, one of these problems that is the most important one is detection the obstacles when they are walking. In this research, we suggested a system with two cameras

placed on blind person's glasses that their duty is taking images from different sides. By comparing these two images, we will be able to find the obstacles. In this method, first we investigate the probability of existence an object by use of special points that then we will call them "Equivalent points", then we utilize binary method, standardize and normalized cross-correlation for verifying this probability. This system was tested under three different conditions and the estimated error is acceptable range.

1.5 OBJECTIVES OF THE SYSTEM

The main objective of this project is to develop an application for blind people to detect the objects in various directions, detecting objects to make free to walk. Detecting objects using image processing can be used in multiple industrial as well as social application. This project is proposing to use object detection for blind people and give them audio/ vocal information about it. We are detecting an object using the mobile camera and giving voice instructions about the direction of an object. User must have to train the system first about the object information .We are then doing feature extraction to search for objects in the camera view. We are taking help of angle where object is placed to give direction about the object.

A dataset of objects gathered from daily scenes is created to apply the required recognition. Objects detection is used to find objects in the real world from an image of the world such as bicycles, chairs, doors, or tables that are common in the scenes of a blind. Based on their locations, and the camera is used to detect any objects. The proposed method for the blind aims at expanding possibilities to people with vision loss to achieve their full potential. The main object of the project is to design and implement a real time object recognition using blind glass.

CHAPTER 2
LITERATURE SURVEY2.1 TITLE: DEEPLY SUPERVISED SALIENT OBJECT DETECTION WITH SHORT CONNECTIONS

AUTHOR: Qibin Hou ; Ming-Ming Cheng ; Xiaowei Hu ; Ali Borji ; Zhuowen Tu ; Philip H. S. Torr

YEAR: 2018

DESCRIPTION:

Recent progress on salient object detection is substantial, benefiting mostly from the explosive development of Convolutional Neural Networks (CNNs). Semantic segmentation and salient object detection algorithms developed lately have been mostly based on Fully Convolutional Neural Networks (FCNs). There is still a large room for improvement over the generic FCN models that do not explicitly deal with the scale-space problem. The Holistically-Nested Edge Detector (HED) provides a skip-layer structure with deep supervision for edge and boundary detection, but the performance gain of HED on saliency detection is not obvious. In this paper, we propose a new salient object detection method by

introducing short connections to the skip-layer structures within the HED architecture. Our framework takes full advantage of multi-level and multi-scale features extracted from FCNs, providing more advanced representations at each layer, a property that is critically needed to perform segment detection. Our method produces state-of-the-art results on 5 widely tested salient object detection benchmarks, with advantages in terms of efficiency (0.08 seconds per image), effectiveness, and simplicity over the existing algorithms Beyond that, we conduct an exhaustive analysis of the role of training data on performance. We provide a training set for future research and fair comparisons.

2.2 TITLE: SALIENT OBJECT DETECTION VIA RANDOMFOREST

AUTHOR: Shuze Du ; Shifeng Chen

YEAR: 2013

DESCRIPTION:

Salient object detection plays an important role in image pre-processing. Existing approaches often neglect the contours of salient objects, thus resulting in inaccurate detection for large objects. Besides, they mainly focus on detecting only a single object. In this paper, we detect the salient object from the view of the object contour. We propose to exploit the random forest to measure patch rarities and compute similarities among patches. A global rarity map is calculated based on the patch's rareness over the whole image. The approximate contour of the salient object is extracted based on this rarity map by using an active contour model. Next, a local saliency map is obtained by the similarities of patches inside the contour and those outside. Finally, the local map is refined through image segmentation. Our method can detect not only a single object but also multiple objects. Experimental evaluation on the ASD-IOOO and SED2 datasets shows that our method outperforms the state-of-the-art methods.

2.3 TITLE: 3D OBJECT PROPOSALS USING STEREO

IMAGERY FOR ACCURATE OBJECT CLASS DETECTION

AUTHOR: Xiaozhi Chen ; KaustavKundu ; Yukun Zhu ; Huimin Ma ; Sanja Fidler ; Raquel Urtasun

YEAR: 2018DESCRIPTION:

The goal of this paper is to perform 3D object detection in the context of autonomous driving. Our method aims at generating a set of high-quality 3D object proposals by exploiting stereo imagery. We formulate the problem as minimizing an energy function that encodes object size priors, placement of objects on the ground plane as well as several depth informed features that reason about free space, point cloud densities and distance to the ground. We then exploit a CNN on top of these proposals to perform object detection. In particular, we employ a convolutional neural net (CNN) that exploits context and depth information to jointly regress to 3D bounding box coordinates and object pose. Our experiments show significant performance gains over existing RGB and RGB-D object proposal methods on the challenging KITTI benchmark. When combined with the CNN, our approach outperforms all existing results in object detection and orientation estimation tasks for all three KIITI object classes. Furthermore, we experiment also with the setting where LIDAR information is available, and show that using both LIDAR and stereo leads to the best result.

2.4 TITLE: WEAKLY SUPERVISED OBJECT DETECTION VIAOBJECT-SPECIFIC PIXEL GRADIENT

AUTHOR: Yunhang Shen ; Rongrong Ji ; Changhu wang ; Xi Li ; Xuelong

YEAR: 2018

DESCRIPTION:

Most existing object detection algorithms are trained based upon a set of fully annotated object regions or bounding boxes, which are typically labor-intensive. On the contrary, nowadays there is a significant amount of image-level annotations cheaply available on the Internet. It is hence a natural thought to explore such "weak" supervision to benefit the tramtng of object detectors. [n this paper, we propose a novel scheme to perform weakly supervised object localization, termed object-specific pixel gradient (OPG). The OPG is trained by using image-level annotations alone, which performs in an iterative manner to localize potential objects in a given image robustly and efficiently. In particular, we first extract an OPG map to reveal the contributions of individual pixels to a given object category, upon which an iterative mining scheme is further introduced to extract instances or components of this object. Moreover, a novel average and max pooling layer is introduced to improve the localization accuracy. In the task of weakly supervised object localization, the OPG achieves a state-of-the-art 44.5% top-5 error on ILSVRC 2013, which outperforms competing methods, including Oquab et al. and region-based convolutional neural networks on the Pascal VOC 2012, with gains of 2.6% and 2.3%, respectively. In the task of object detection, OPG achieves a comparable performance of 27.0% mean average precision on Pascal VOC 2007. In all experiments, the OPG only adopts the off-the-shelf pretrained CNN model, without using any object proposals. Therefore, it also significantly improves the detection speed, i.e., achieving three times faster compared with the state-of-the-art method.

2.5 TITLE: VISUAL AND SEMANTIC KNOWLEDGE TRANSFER FOR LARGE SCALE SEMI-SUPERVISED OBJECT DETECTION

AUTHOR: Yuxing Tang ; Josiah wang ; Xiaofang wang ; BoyangGao ; Emmanuel Dellandréa ; Robert Gaizauskas ; Liming Chen

YEAR: 2017

DESCRIPTION:

Deep CNN-based object detection systems have achieved remarkable success on several largescale object detection benchmarks. However, training such detectors requires a large number of labeled bounding boxes, which are more difficult to obtain than image-level annotations. Previous work addresses this issue by transforming image-level classifiers into object detectors. This is done by modeling the differences between the two on categories with both image-level and bounding box annotations, and transferring this information to convert classifiers to detectors for categories without bounding box annotations. We improve this previous work by incorporating knowledge about object similarities from visual and semantic domains during the transfer process. The intuition behind our proposed method is that visually and semantically similar categories should exhibit more common transferable properties than dissimilar categories, e.g. a better detector would result by transforming the differences between a dog classifier and dog detector onto the cat class, than would by transforming from the violin class. Experimental results on the challenging ILSVRC2013 detection dataset demonstrate that each of our proposed object similarity based knowledge transfer methods outperforms the baseline methods.

CHAPTER 3
SYSTEM STUDY

3.1 EXISTING SYSTEM

The work they present in this system is based on the use of new technologies to improve visually impair people mobility. Our research is on obstacle detection in order to reduce navigation difficulties for visually impaired people. Moving through an unknown environment becomes a real challenge when we can't rely on our own eyes. Since dynamic obstacles usually produce noise while moving, blind people develop the ability of hearing to localize them. However they are reduced to their sense of touch when the matter is to determine where an inanimate object exactly is. The common way for navigating of visionless person is using a walking stick cane or walking cane. The walking cane is a simple and mechanical device dedicated to detect static obstacles on the ground, uneven surfaces, and holes via simple tactileforce feedback. This device is light, portable, but range limited and it is not usable for the protection from obstacles near to head area. Another option that provides the best travel aid for the blind is the guide dogs. Based on the symbiosis between the disabled owner and his dog, the training and the relationship to the animal are the keys to success for this method. In existing system cross model object detection technique has been used. Here the visual trained features will identify the tactile points the proposed approach achieves an accuracy of 94.7%, which is comparable with the accuracy of the monomodal case.

3.1.1 Disadvantages

- Blind people suffers a lot while they are out of their home
- They always needs a guidance for them to cross the road or go out
- Due to this many visually impaired people doesn't go out and spend time without others help
- The accuracy is just 94 % achieved
- Out of 40 objects 15 objects are only classified and identified
- Labeling of the objects is not made
- Just identifying the object is not just useful for the people

3.2 PROPOSED SYSTEM

Normally, a blind person uses cane as a guide of him to protect him from obstacles. Most of area of surrounding is covered by the cane, especially the area near to his legs like stairs etc. But certain areas such as near to his head, especially when he is entering or leaving the door which is

short in height. This system is specially designed to protect the area near to his head. The product is designed to provide full navigation to user into the environment. It guides the user about obstacles as well as also provides information about appropriate or obstacle free path. We are using buzzer and Vibrator, two output modes to user. The developed system helps the visually impaired to navigate independently using real-time object detection and identification technology. The system makes use of the image processing technique to detect the object and speech synthesis to produce the voice output. Fast R-ANN (Region-based Artificial Neural Networks) algorithm has been implemented to detect the object with high accuracy. The detected image information is provided as a voice output using a speech synthesizer to the visually challenged persons to assist them in their mobility. The detection of images on moving objects has been a significant research area in computer vision which has been highly worked upon, and integrated with residential, commercial and industrial environments.

3.2.1 Advantages

- Blind people can use this system for the road crossing areas and as a identification system

  They don't need assistance always when they come out

- Object labeling with voice synthesis helps the user to identify the objects before them

  The accuracy is achieved with 97% where the object is completely detected.

3-3 MODULE DESCRIPTION

- Video streaming
- Key frame extraction
- Co- variance extraction
- Artificial Neural Network
- Object detection and recognition
- Voice synthesis

3.3.1 VIDEO STREAMING

Object recognition is a classical problem in computer vision: the task of determining if the image data contains a specific object and it is noted that general object recognition approaches exploit features extraction. Features that have received the most attention in the recent years are the local features. The main idea is to focus on the areas containing the most discriminative information. The detection of an object cannot be achieved with high accuracy in normal cases, as

most of the application provides the detection of objects captured as an image rather than detection in a live streaming video.

The proposed system captures the image through the continuous video stream rather than taking the picture of each object every time. The object detection and recognition have become easier when it is detecting from the image but in the case of the video stream, the processing speed should be high to detect the entire object in a frame. The functional process is to capture the images through the camera and process it through image processing algorithms

3.3.2 KEY FRAME EXTRACTION

The frame extraction play the vital role in many video processing applications like content based video retrieval, shot detection, segmentation, CC cameras , etc. The frame conversion can be get with the seconds of the video which are get with the video. Each frame will be analyzed to know the object in the scene. The frame conversion is the process of extracting the images from the video where the sequences of images can be delivered as frames with the given video. In this paper, a new technique for key frame extraction is presented. The scheme uses an aggregation mechanism to combine the visual features extracted from the correlation of RGB color channels, color histogram, and moments of inertia to extract key frames from the video. An adaptive formula is then used to combine the results of the current iteration with those from the previous. The use of the adaptive formula generates a smooth output function and also reduces redundancy. The results are compared to some of the other techniques based on objective criteria. The experimental results show that the proposed technique generates summaries that are closer to the summaries created by humans.

Fig 3.1 Summarized key frame extraction from video streaming

1) Uniform Sampling: Uniform sampling is one of the most common methods for keyframe extraction. The idea is to select every kth frame from the video where the value of k is dictated by the length of the video. A usual choice of length for a summarized video is 5% to 15% of the original Video, which means every 20th frame in case of 5% or every 7th frame in case of 15% length of the summarized video is chosen. For our experiment, we have chosen to use every 7th frame to summarize the video. This is a very simple concept which does not maintain semanticrelevance. Uniform sampling is often considered as a generalized baseline for video summarization.

2) Image histogram: Image histograms represent the total distribution of an image. It gives us the number of pixels for a specific brightness values rated from O to 256. Image histograms contain important information about images and they can be utilized to extract keyframes. We extract the histogram from all frames. Based on the difference between histograms of two frames, we decide whether the frames have significant dissimilarities among them. We infer that, a significant inter-frame image histogram dissimilarity indicates a rapid change of scene in the video which might contain interesting components. For our experiments, if histograms of two consecutive frames are 50% or more dissimilar, we extract that frame as a keyframe.

3) Scale Invariant Feature Transform: Scale Invariant Feature Transform (COVARIANCE) , has been one of the most prominent local features used in computer vision is applications en for each channel resulting in a 16Xl 6X16 tensor. Due to computational reasons, a simplified version of this histogram was computed, where each channel was treated separately, resulting in feature vectors for each frame belonging to R 48 . The nest step suggested for clustering is slightly different. But, the simplified color histograms give comparable performance to the true color histograms. Therefore, we modified VSUMM a bit by using the features extracted from VGG16 at the 2nd fully connected layer, and clustered it using kmeans.

5) ResNet16 on ImageNet: While reading about approach of VSUMM, we decided to test a different approach. We chose ResNet16 trained on image net, with different range of filters, and chopped of last loss layer, so as to obtain the embeddings of each image (512 dimension). We extracted frames out of the videos, and forward pass them through ResNet16, and after obtaining the embeddings for each frame in video, we clustered them using 2 algorithms: K-means and Gaussian Mixture Models. The number of cluster has been take as 15% of the video frame numbers.

We later chose the frames closest to the center of clusters as the keyframes.

3.3.3 COVARIANCE EXTRACTION

The main purpose of using features instead of raw pixel values as the input to a learning algorithm is to reduce/increase the in-class/out-of class variability compared to the raw input data, and thus making classification easier.

Types of features that can be extracted from image depend on the type of image, the level of granularity desired, and the context of the application. Once the features have been extracted, theirs representation depends on the technique used. The features extraction process should be precise, so that the same features are extracted on two images showing the same object.

Global image descriptor: features overall image are usually based on color indices and the most famous global color descriptor is the color histogram.

Local image descriptors: Local features are ones that have received the most attention in recent years. The main idea is to focus on the areas containing the most discriminated information.

Semi-local image descriptor: most shape descriptors fall into this category. This descriptor is based on extracting accurate contours of shapes in the image or in the region of interest. In this case, image segmentation is generally useful as a preprocessing step.

3.3.4 ARTIFICIAL NEURAL NETWORK FOR CLASSIFICATION

The artificial neural network classification system is used to classify the co-variance descriptor points. Feature matching using invariant features has obtained significant importance due to its application in various recognition problems. Such techniques have enabled us to match images irrespective of various geometric and photometric transformations between images.

For the object detection and labeling the machine learning method is used. The Artificial Neural Network is used for training and testing purpose. The region based segmentation is used and then they are segmented to identify the object. The object detection and recognition have become easier when it is detecting from the image but in the case of the video stream. We first take a pretrained Artificial Neural Network. Then, this model is retrained. We train the last layer of the network based on the number of classes that need to be detected. The third step is to get the Region of Interest for each image. We then reshape all these regions so that they can match the ANN input size. After getting the regions, we train ANN to classify objects and background. For each class, we train one binary ANN. Finally, we train a linear regression model to generate tighter bounding boxes for each identified object in the image.

Step (1) Obtain the set of key-points of objects

a. Select a large set of images of daily objects.

b. Extract the CO-VARIANCE feature points of all the images within the set and obtain the CO-VARIANCE descriptor for each feature point extracted from each image.

Step (2) Obtain the keypoints descriptor for the first video frame.

a. Extract CO-VARIANCE feature points of the given image.

b. Acquire CO-VARIANCE descriptor for each feature point.

c. Match the frame key-points with those of the objects and identify detected objects.

Step (3) For the next frame:

a. If it contains the same objects they will not be detected
b. New objects will be detected and identified.
c. Another method will be used in this step to identify similar and dissimilar frames for further treatments.

Step (4) For each object detected in video a video file is launched to notify the blind about the identity of objects.

3.3.5 OBJECT DETECTION AND RECOGNITION

The object detection method will be done in the testing process. With the region of interest the object in the scene are detected. The detected objects will be further labeled with the name of the object. Thus the object recognition is made here.

Object detection is the process of finding instances of real-world objects such as faces, bicycles, and buildings in images or videos. Object detection algorithms typically use extracted features and learning algorithms to recognize instances of an object category. It is commonly used in applications such as image retrieval, security, surveillance, and automated vehicle parking systems. Detecting a reference object (left) in a cluttered scene (right) using feature extraction and matching. RANSAC is used to estimate the location of the object in the test image. Local features and their descriptors are the building blocks ofmany computer vision algorithms. Their applications include image registration, object detection and classification, tracking, and motion estimation.

These algorithms use local features to better handle scale changes, rotation, and occlusion.

Fig 3.2 shows the object features extraction at each key frame

The features we find are described in a way which makes them invariant to size changes, rotation and position. These are quite powerful features and are used in a variety of tasks. We use a standard Java implementation of SIFT. The SIFT descriptor is a 128 dimensional description of a patch of pixels around the key point.

3.3.6 VOICE SYNTHESIS

The voice synthesis is added for the voice based output system. Here the classified objects are known with labeling where the labels of the object will be produced as voice output. This helps the blind people to know the object using the voice process

Speech recognition basically uses voice waveform analysis techniques. The speech pattern is generated by word or sentence, which is then used to recognize speech by comparing the learned pattern and the newly received pattern. This comparison process is carTied out using a formula related to cross-correlation. However, voice signals contain a mixture of noise components, and signals with a different frequency are also received as an input signal. In this study, a statistical model called the hidden Markov model (HMM) is used to accurately find the desired voice. HMM assumes that the system being modeled is a Markov process with unknown parameters, and it determines which parameters are hidden from previously observed parameters. It is used to estimate the parameters of the model in the learning stage, and then, to find the estimated parameters in the new incoming voice. Google announced that its voice recognition technology is becoming more sophisticated every year by learning parameters using large amounts of data, and said the error rate was only 4.9% in 2017. Based on these results, the

research in this study was conducted by applyingGoogle's voice recognition service to the voice recognition function.

ARTIFICIAL NEURAL NETWORK (ALGORITHM)

Inputs: V, the input video stream; to, the size
threshold; to, the temporal object appearance
threshold; td, the temporal detection threshold;

Outputs: K,the list of key frame labels; I. for each
frame f in V/* moving region extraction*/
2.extract the set of moving regions Rf
3.for each element r in Rf;
4. if size(r)<ts.
5. remove r from Rf
6. else
7. increment temporal- appearance(r) inverted tracking */
8. if temporal-appearance(r)>to
9. set the cell(r) to I in MAMf,/5 keyframe detection */
10. if is Keyframe(MAMf)/*Keyframe detection* /
11. compute Keyframe label If
12. increment temporal-count(if)
13. if temporal-count(if)>td
14. push if onto K
15. MAM
16. clear Mf
17 -end for

CHAPTER 4

SYSTEM REQUIREMENTS

4.1 HARDWARE SPECIFICATION

- Processor 2.4 GHz processor
- Ram:2.00GB or more
- Hard Disk IOOGB
- Monitor CRT Monitor 15inch
- Keyboard Multimedia Keyboard
- Mouse Optical mouse

4.2 SOFTWARE SPECIFICATION

- Operating System Windows 7 or Higher
- Front-End Python

CHAPTER 5
SYSTEM DESIGN

5.1 SYSTEM ARCHITECTURE

Numerous advantages are derived from the electronic health record (EHR).Though achieving such advantages depends on its architecture, at present no unique understanding of the architecture dimensions and specifications is available. Therefore, the aim of the present study is a systematic review of architecture perception of the electronic health record.

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system. System architecture can comprise system components, the externally visible properties of those components, the relationships (e.g. the behavior) between them. It can provide a plan from which products can be procured, and systems developed, that will work together to implement the overall system. There have been efforts to formalize languages to describe system architecture; collectively these are called architecture description languages (ADLs).

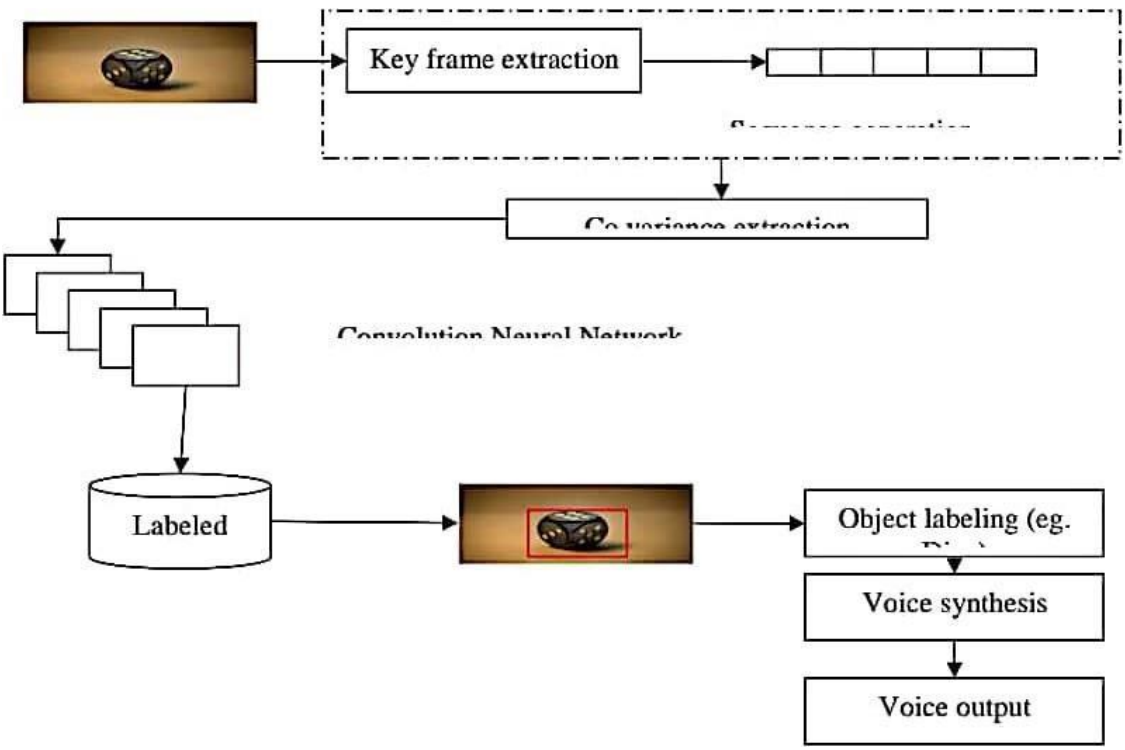
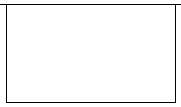
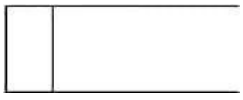
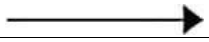


Fig 5.1 System architecture of blind people object detection

5.2 DATA FLOW DIAGRAM

A two-dimensional diagram explains how data is processed and transferred in a system. The graphical depiction identifies each source of data and how it interacts with other data sources to reach a common output. Individuals seeking to draft a data flow diagram must identify external inputs and outputs, determine how the inputs and outputs relate to each other, and explain with graphics how these connections relate and what they result in. This type of diagram helps business development and design teams visualize how data is processed and identify or improve certain aspects.

Data flow Symbols:

Symbol	Description
	An entity. A source of data or a destination for data.
	A process or task that is performed by the system.
	A data store, a place where data is held between processes.
	A data now.

LEVEL O

DFD Level O is also called a Context Diagram, It's a basic overview of the whole system or process being analyzed or modeled. It's designed to be an at-a-glance view, showing the system as a single high-level process, with its relationship to external entities. It should be easily

understood by a wide audience, including stakeholders, business analysts, data analysts and developers.

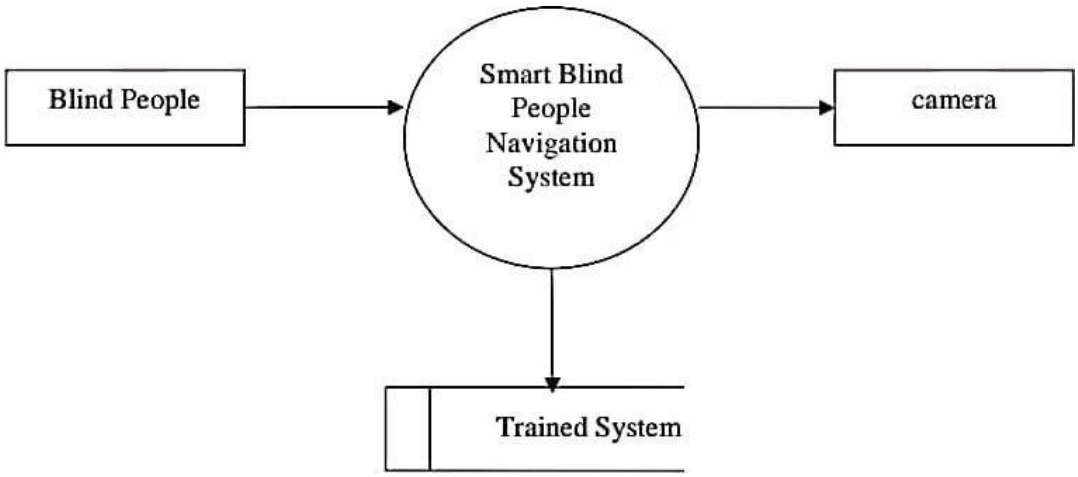


Fig 5.2.1 Level O DFD Diagram

LEVEL 1

DFD Level 1 provides a more detailed breakout of pieces of the Context Level Diagram. You will highlight the main functions carried out by the system, as you break down the high-level process of the Context Diagram into its sub — processes.

5.3. UML DIAGRAMS

5.3.1 USE CASE DIAGRAM

An interaction between a user and a system is described by use case diagram. Use case diagrams describe what a system does from the standpoint of an external observer. The emphasis is on what a system does rather than how. For initial development we can use this use case. In this use case diagram we can see seven use cases and two actors.

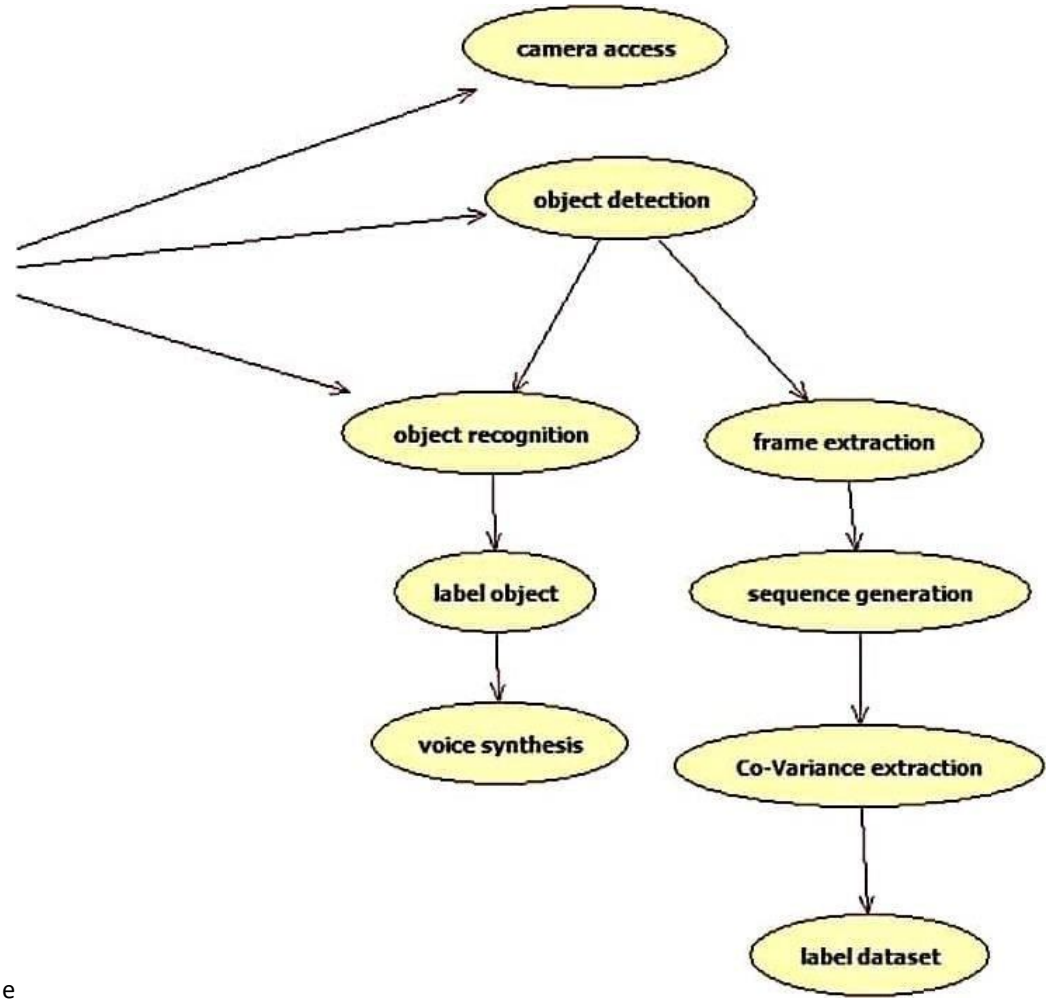


Fig 5.3.1 Use Case Diagram

5.3.2 Activity Diagram

We use Activity Diagrams to illustrate the flow of control in a system and refer to the steps involved in the execution of a use case. We model sequential and concurrent activities using activity diagrams. So, we basically depict workflows visually using an activity diagram An activity diagram focuses on condition of flow and the sequence in which it happens. We describe or depict what causes a particular event using an activity diagram. UML models basically three types of diagrams, namely, structure diagrams, interaction diagrams, and behavior diagrams. An activity diagram is a behavioral diagram i.e. it depicts the behavior of a system.

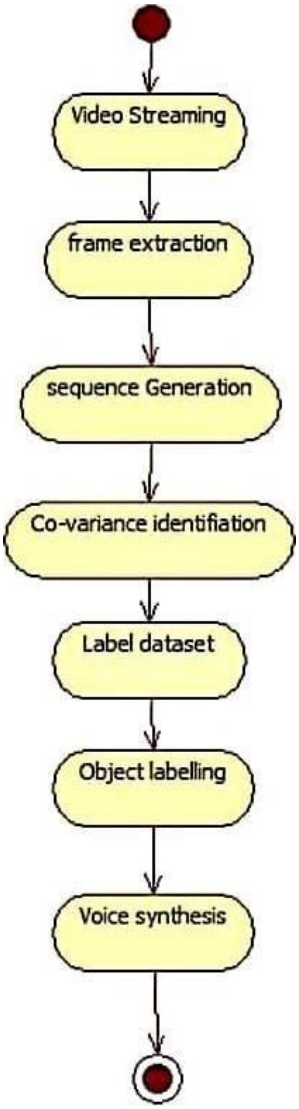


Fig 5.3.2 Activity Diagram

5.3.3 Sequence Diagram

Sequence diagrams are sometimes called event diagrams or event scenarios. A sequence diagram shows, as parallel vertical lines (lifelines), different processes or objects that live simultaneously, and, as horizontal arrows, the messages exchanged between them, in the order in which they occur. A sequence diagram is an interaction diagram that shows how objects operate with one another and in what order. It is a construct of a message sequence chart- A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called event diagrams or event scenarios. Objects as well as classes can be targets on a sequence diagram, which means that messages can be sent to them. A target is displayed as a

rectangle with some text in it. Below the target, its lifeline extends for as long as the target exists. The lifeline is displayed as a vertical dashed line.

CHAPTER 6
SYSTEM IMPLEMENTATION

6.1 PYTHON

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

6.1.1 HISTORY OF PYTHON

Python was developed by Guido van Rossum in the late eighties and early nineties at the National Research Institute for Mathematics and Computer Science in the Netherlands.

Python is derived from many other languages, including ABC, Modula-3, C, C-H-, Algol68, SmallTalk, and Unix shell and other scripting languages.

Python is copyrighted. Like Perl, Python source code is now available under the GNU General Public License (GPL).

Python is now maintained by a core development team at the institute, although Guido van Rossum still holds a vital role in directing its progress.

Often, programmers fall in love with Python because of the increased productivity it provides. Since there is no compilation step, the edit-test-debug cycle is incredibly fast. Debugging Python programs is easy: a bug or bad input will never cause a segmentation fault. Instead, when the interpreter discovers an error, it raises an exception. When the program doesn't catch the exception, the interpreter prints a stack trace. A source level debugger allows inspection of local and global variables, evaluation of arbitrary expressions, setting breakpoints, stepping through the code a line at a time, and so on. The debugger is written in Python itself, testifying to Python's introspective power. On the other hand, often the quickest way to debug a program is to add a few print statements to the source: the fast edit-test-debug cycle makes this simple approach very effective.

6.1.2 FEATURES IN PYTHON

There are many features in Python, some of which are discussed below_ 1
Easy to code:

Python is high level programming language. Python is very easy to learn language as compared to other language like c, c#, java script, javaetc.ltis very easy to code in python language and anybody can learn python basic in few hours or days.lt is also developer-friendly language.

2. Free and Open Source:
Python language is freely available at official website and you can download it from the given download link below click on the Download Python keyword.

Download Python
Since, it is open-source; this means that source code is also available to the public. So you can download it as, use it as well as share it.
3.0bject-Oriented Language:
One of the key features of python is Object-Oriented programming. Python supports object oriented language and concepts of classes, objects encapsulation etc.

4.GUI Programming Support:
Graphical Users interfaces can be made using a module such as PyQt5, PyQt4, wxPython or Tk in python.

PyQt5 is the most popular option for creating graphical apps with Python
5.High-Level Language:

Python is a high-level language. When we write programs in python, we do not need to remember the system architecture, nor do we need to manage the memory.

6. Extensible feature:

Python is a Extensible language. we can write our some python code into c or C-H- language and also we can compile that code in language.

7. Python is Portable language.

Python language is also a portable language. for example, if we have python code for windows and if we want to run this code on other platform such as Linux, Unix and Mac then we do not need to change it, we can run this code on any platform.

8. Python is integrated language:

Python is also an integrated language because we can easily integrated python with other language like c, C-H- etc.

9. Interpreted Language:

Python is an Interpreted Language. Because python code is executed line by line at a time. like other language c, c-H-, java etc there is no need to compile python code this makes it easier to debug our code-The source code of python is converted into an immediate form called bytecode.

10. Large Standard Library

Python has a large standard library which provides rich set of module and functions so you do not have to write your own code for every single thing. There are many libraries present in python for such as regular expressions, unit-testing, web browsers etc.

11. Dynamically Typed Language:

Python is dynamically-typed language. That means the type (for example- int, double, long etc) for a variable is decided at run time not in advance. Because of this feature we don't need to specify the type of variable.

Apart from the above-mentioned features, Python has a big list of good features, few are listed below_

- It supports functional and structured programming methods as well as OOP.
- It can be used as a scripting language or can be compiled to byte-code for building large applications.
- It provides very high-level dynamic data types and supports dynamic type checking. It supports automatic garbage collection. It can be easily integrated with C, C++, COM, ActiveX, CORBA, and Java.

CHAPTER 7
SYSTEM TESTING

7.1 INTRODUCTION

Software testing is a method of assessing functionality of a software program there are many different types of software testing but the two main categories are dynamic testing and static testing. Dynamic testing is an assessment that is considered while the program is executed; static testing. on the other hand is an examination of the program scored and associated documentation. Dynamic and static methods are often used together. Testing is a set activity that can be planned and conductor systematically. Testing begins at the module level and what towards the integration of entire computer based system. Nothing is complete without testing, as it is why to success of the system.

7.1.1 Testing objectives

There are several rules that can service testing objectives, they are

1. Testing is a process of executing a program with the intent of finding an error.
2. A good test case is one that has high probability of finding an undiscovered error.
3. A successful test is one that uncovers an undiscovered error.

If testing is conducted successfully according to the objectives as stated above, it would uncover errors in the software also testing demonstrates that software functions appear to the working according to the specification that performance requirements appear to have been met.

There are three ways to test a program

1. For correctness

2. For implementation efficiency
3. For computational complexity

Tests for correctness are supposed to verify today program does exactly what it was designed to do this is much more difficult than it mean at first appear especially for large programs.

Test used implementation efficiency attempted to find ways to make correct program faster or use less storage. It is a code refining process which we examine the implementation phase of algorithm development. Test for computational complexity amount to an experimental analysis of the complexity of an algorithm are an in experimental comparison of two or more algorithms which solve the same problem.

The data is entered in all forms separately and whenever an error occurred, it is corrected immediately. A quality team deputed by the management verified all the necessary documents and tested the software while entering the data at all levels. The development process involves various types of testing each type addresses a specific testing requirement . The most common type of testing involved in the development process are: 1.Unit test

2. Integration test
3. System test
4. Functional test

Unit testing

The first test in the development process is the unit test. The source code is normally divided into modules, which in turn are divided into smaller units called units. These units have specific behavior that test done on these units of code is called unit test. It depends upon the language on which the project is developed. Unit test ensure that each unique part of the project performs accurately to the documented specifications and contains clearly defined inputs and expected results.

Integration testing

In integration testing models are combined and tested as a group. Module are typically code modules, individual applications, source and destination applications on a network , etc. Integration follows unit testing and proceeds system testing after the product code complete. Betas are often widely distributed to the public at large in hopes that will buy the final product when it is released.

Acceptance testing

This testing is done to verify the readiness of the system for the implementation. Acceptance testing behaves when the system is complete its purpose is to provide the end user

with the confidence that the system is ready for use. It involves planning and execution of functional test performance tests and stress Tests in order to demonstrate that the implementation satisfies its requirements.

White box testing

Testing based on an analysis of internal workings and structure of a piece of software. This testing can be done in the percentage value of load and energy that is there should know what exactly is done in the internal program it includes techniques such as branch testing and path testing.

It also called is structural testing and glass box testing.

Black box testing

Black box tests conducted by without knowledge of the internal working of the system being tested. Tests are usually functional this test can be done by the user who was no knowledge of how the software is working.

Testing strategies and approach

A

Test ID	STOI
Unit tested	To test whether the complete object reco •tion is made
Purpose	To check that the output is executed rfectl
Pre re uirement	Pre trained dataset and requirements
Test Data	Co-ordinate values of ob•ect
Test status	0b •ect detected and labeled successfull
Test result	Pass

s t r a t e g y f

or software testing must accommodate low level test that are necessary to verify that all small source code segments has been correctly implemented as well as high level test that validate major system functions against customer requirements.

7.2 TEST CASE

A test case is a set of conditions or variables under which a tester will determine whether a system under test satisfies requirements or works correctly. The process of developing test cases can also help find problems in the requirements or design of an application.

Fig 7.1 represents the system testing for recognition of object and labeling

7.3 TEST PLAN

A test plan is a document describing software testing scope and activities. It is the basis for formally testing any software / product in a project.

- Test plan: A document describing the scope, approach, resources and schedule of intended test activities. It identifies amongst others test items, the features to be tested, the testing tasks, who will do each task, degree of tester independence, the test environment, the test design techniques and entry and exit criteria to be used, and the rationale for their choice, and any risks requiring contingency planning. It is a record of the test planning process.

- Master test plan: A test plan that typically addresses multiple test levels. ● Phase test plan: A test plan that typically addresses one test phase.Test Plan is influenced by the following factors:

- Organizational test policy and test sffategy 0 Software development life cycle model

0 Scope of testing

- Availability of resources.

At the beginning of the project, the test plan can be a draft or with very little details. But, as the project progresses and more information become available, the test plan needs to be fleshed out. Test planning is a continuous activity and is performed throughout the product's life cycle.

7.3.1 Test Plan Identifier:

- Provide a unique identifier for the document. (Adhere to the Configuration Management System if you have one.)

o Introduction:

- Provide an overview of the test plan.
- Specify the goals / objectives.

0 Specify any constraints.

References:

- List the related documents, with links to them if available, including the following:

0 Project Plan

- Configuration Management Plan Test Items:

- List the test items (software / products) and their versions. Features to be Tested:

- List the features of the software / product to be tested.
- Provide references to the Requirements and/or Design specifications of the features to be tested.

Features Not to Be Tested:

0 List the features of the software / product which will not be tested.
0 Specify the reasons these features won't be tested.

Approach:

0 Mention the overall approach to testing. 0 Specify the following:

- Testing levels [if it's a Master Test Plan]
- Testing types
- Testing methods

Item Pass / Fail Criteria:

0 Specify the criteria that will be used to determine whether each test item has passed or failed testing.

Suspension Criteria and Resumption Requirements:

- 0 Specify criteria to be used to suspend the testing activity.
- Specify what is required before testing can resume.

Test Deliverables:

- List test deliverables, and links to them if available, including the following:
- Test Plan (this document itself)
- Test Cases
- Test Scripts

Test Environment:

0 Specify the properties of test environment: hardware, software, network, etc.

- List any testing or related tools.

Estimate:

o Provide a summary of test estimates (cost or effort) and/or provide a link to the detailed estimation.

Schedule:

o Provide a summary of the schedule, specifying key test milestones, and/or provide a link to the detailed schedule.

Staffing and Training Needs:

0 Specify staffing needs by role and required skills.

0 Identify training that is necessary to provide those skills, if not already acquired.

Responsibilities:

- List the responsibilities of each team / role / individual.

List the risks that have been identified.

0 Specify the mitigation plan and the contingency plan for each risk.

Assumptions and Dependencies:

o List the assumptions that have been made during the preparation of this plan.

0 List the dependencies.

Approvals:

0 Specify the names and roles of all persons who must approve the plan.

Provide space for signatures and dates. (If the document is to be printed.)

7.3.2 Test Plan Guidelines

Make the plan concise. Avoid redundancy and super fluousness. If you think you do not need a section that has been mentioned in the template above, go ahead and delete that section in your test plan.

Be specific. For example, when you specify an operating system as a property of a test environment, mention the OS Edition / Version as well, not just the OS Name.

0 Make use of lists and tables wherever possible. Avoid lengthy paragraphs.

Have the test plan reviewed a number of times prior to base lining it or sending it for approval. The quality of your test plan speaks volumes about the quality of the testing you or your team is going to perform.

CHAPTER 8
CONCLUSION AND FUTURE ENHANCEMENT

8.1 CONCLUSION

In this paper, we design an object detection system using a deep learning object recognition technique and voice recognition technology. This system's voice synthesis provides convenience features for the visually impaired. As one of the areas where deep learning technology can be applied, our study was conducted by focusing on how to effectively aid the blind. As a result, voice recognition and voice guidance technologies were added to the system, and its performance was tested. This study can be used widely to provide the blind with privacy and convenience in everyday life. Also, it is expected to be applied to industrial areas where

diminished visibility occurs, such as coal mines and sea beds, to greatly help production and industrial development in extreme environments.

8.2 FUTURE ENHANCEMENT

In future system the implementation can be used in the smart phone system where the blind people can easily access using their smart phone system. Some limitations to this system are that the smart phone on which this application will be used will have to be switched on and should have enough battery. It has to be carried by the user with him/her all the time. A wearable device is more convenient as our hands become free in that case

APPENDICES

Main.py

```
import cv2 import

numpy as np


# load the class labels from disk import pyttsx3 as

pyttsx3 engine = pyttsx3 .init() rows =

open("synset_words.txt").read().strip().split("\n")

classes = r[r.find(" ") + 1:].split(",")[0] for r in rows]


# load our serialized model from disk net = cv2.dnn.readNetFromCaffe(

"bvlc_googlenet.prototxt" , "bvlc_googlenet.caffemodel ")

##############testarun

cap = cv2.VideoCapture(O)

while I • ret, img = capread()

    #cv2.imshow("Image", img)

    image = (img)
```

```
resized = cv2.resize(image, (224, 224)) blob =
cv2.dnn.b10bFrom1mage(resized, 1, (224, 224), (104, 117, 123))
net.setlnput(blob)

preds = net-forward()
idx =np.argsort(preds[0])[::-1][0]
text = "Label: { , {:.2f}%".format(classes[idx],preds[0][idx]* 100)
textl = " ( ) " format(classes[idx])
 cv2.putText(image, text, (5, 25), cv2.FONT_HERSHEY_SIMPLEX,0.7,(0, O, 255), 2)
cv2.imshow("Image1" ,img)
#print(textl)
engine.say(textl)
engine.runAndWait() k =
cv2.waitKey(30) & Oxff

if k == 27:
        break
cap.release()
cv2.destroyA11Windows()
```

```
Test,py import mysql.connector conn = mysql.connector.connect(user='root', password='',
host='localhost•, database='test') cursor = conn.cursor()
cursor.execute("INSERTINTO userVALUES('1','2','3','4')")
conn.commit()
conn.close()

#!E:\yogesh\2019\Python\Object_detection\venv\Scripts\python.exe
# EASY-INSTALL-ENTRY-SCRIPT:'setuptools==39.1.0','console_scripts','easy_install-3.6'
```

```
_requires_= 'setuptools==39 1.0' import re

import sys from pkg_resources import

load_entry_point


if name       main ':sys.argv[O] =

   e.sub(r'(-script\.pyw?|\.exe)?$','', sys.argv[O])

   sys .exit( load_entry_point('setuptools==39.l.O', 'console_scripts',

      'easy_ins


import os

import sys

import

shutil


def get_next_cache_dir():

   work_dir os.getcwd() try:

      # change working directory to avoid import from local folder

      #     during     installation     process     os.chdir(os

   .path.dirname(sys.executable)) import comtypes.client

   return    comtypes.client._code_cache._find_gen_dir()

   except ImportError:

      return None

   finally:

   os.chdir(work_dir)
```

```
def _remove(directory):
   shutil.rmtree(directory) print(Removed
   directory "%s"' % directory)


def emove_directory(directory, silent):
   if directory:
      if silent:
         _remove(directory)
      else:
         try:
            confirm = raw_input(Remove comtypes cache directories? (y/n): ')
         except NameError: confirm = input('Remove comtypes cache
         directories? (y/n):
         ') if confirm-lower() == 'y':_remove(directory) else:

            print(Directory "%s" NOT removed' % directory)
   return False return True


if len(sys.argv) > 1 and "-y" in sys-
argv[l:]: silent =True else:

   silent = False
#First iteration may get folder with restricted rights.
# Second iteration always gets temp cache folder (writable for all).
directory get_next_cache_dir() removed =
remove_directory(directory, silent)
```

```
if removed: directory =
    get_next_cache_dir()

    # do not request the second confirmation # if
    the first folder was already removed
    remove_directory(directory, silent=removed)

#!E:\yogesh\2019\Python\Object_detection\venv\Scripts\python.exe
#EAS Y-INSTALL-ENTR Y-SCRIPT: 'setupt001s=39.1
.0','console_scripts','easy_install-3.6'__requires_ = 'setuptools==39. I .0' import re import sys
from pkg_resources import load_entry_point

if name       main ':sys.argvlO] =
    e.sub(r'(-script\.pyw?|\.exe)?$',", sys.argvlO])
    sys-exit(
        load_entry_point('setuptools==39.1.0','console_scripts', 'easy_install-3.6')()
              #!E:\yogesh\2019\Python\Object_detection\venv\Scripts\python.exe
# EASY-INSTALL-ENTRY-SCRIPT:'setuptools==39.1.0','console_scripts','easy_install'__
requires_ = 'setuptools==39. I .0'
import    re    import    sys    from
pkg_resources                  import
load_entry_point

if _name      main __':sys.argv10] =
    e.sub(r'(-script\.pyw?|\.exe)?$',", sys.argv[0])
    sys-exit(
        load_entry_point('setuptools==39.1.0','console_scripts','easy_ins
```

B. SCREENSHOT

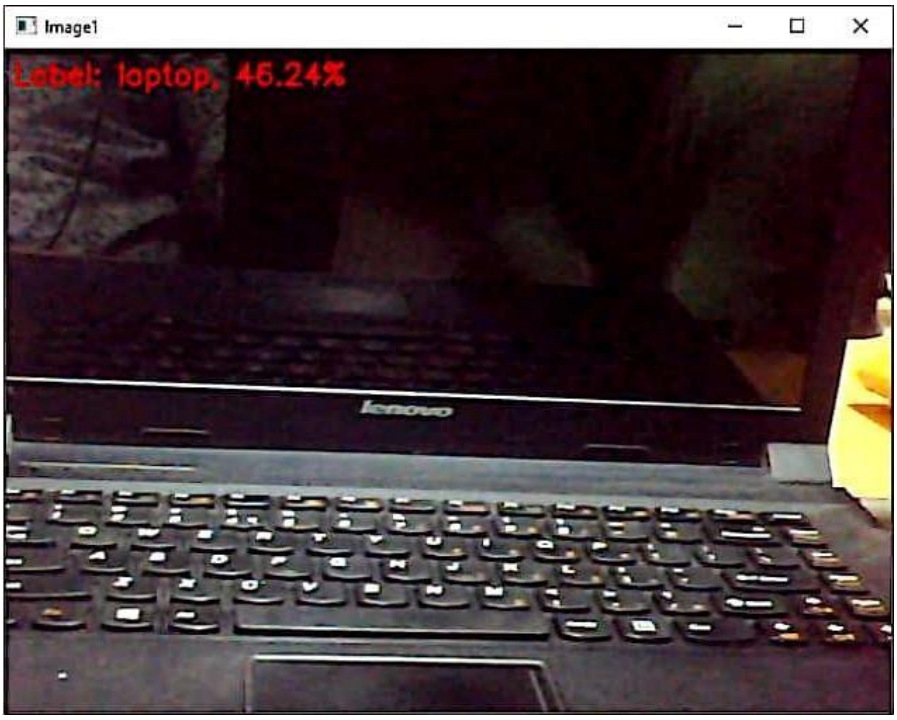


Fig 1.1 Object detection system

Fig 2 detecting the watch with voice recognition

REFERENCES

[1] QibinHou ; Ming-Ming Cheng ;XiaoweiHu ; Ali Borji ; ZhuowenTu ; PhilipH. S. Torr, Deeply Supervised Salient Object Detection with Short Connections, IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 41 , Issue: 4 , April 1 2019)

[2] Shuze Du ; Shifeng Chen, Salient Object Detection via Random Forest, IEEE Signal Processing Letters (Volume: 21 , Issue: I , Jan. 2014)

13] Xiaozhi Chen ; Kaustav Kundu ; Yukun Thu ; Huimin Ma ; Sanja Fidler ; Raquel Urtasun, 3D Object Proposals Using Stereo Imagery for Accurate Object Class Detection, IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 40 , Issue: 5 , May 1 2018)

[4]Yunhang Shen ; Rongrong Ji ; Changhu Wang ; Xi Li ; Xuelong Li , Weakly Supervised Object Detection via Object-Specific Pixel Gradient, IEEE Transactions on Neural Networks and Learning Systems (Volume: 29 , Issue: 12 , Dec. 2018)

[5]Yuxing Tang ; Josiah Wang ; Boyang Gao ; Emmanuel Dellandréa ; Robert Gaizauskas ; Liming Chen, Large Scale Semi-Supervised Object Detection Using Visual and Semantic Knowledge Transfer, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016