AI-BASED SMART GLASSES FOR BLIND PEOPLE

Dr.CHEN CHOU

Department of Computer Engineering, Dr. D. Y. Patil School of Engineering, Lohegaon, Savitribai Phule Pune University, Pune, Maharashtra, India

Sharan.inamdar@gmail.com, Shubhamksharma2001@gmail.com, goswamivikesh709@gmail.com, vivekmahajan1940@gmail.com

Abstract - Motivated by the recent success of supervised and weakly supervised common object discovery, in this work we move forward one step further to tackle common object discovery in a fully unsupervised way. Generally, object co-localization aims at simultaneously localizing objects of the same class across a group of images. Traditional object localization/ detection usually trains specific object detectors which require bounding box annotations of object instances, or at least image-level labels to indicate the presence/absence of objects in an image. Given a collection of images without annotations, our proposed fully unsupervised method is to simultaneously discover images containing common objects and localize common objects in corresponding images. Without requiring knowing the total number of common objects, we formulate this unsupervised object discovery as a sub-graph mining problem from a weighted graph of object proposals, where nodes correspond to object proposals and edges represent the similarities between neighboring proposals. The positive images and common objects are jointly discovered by finding sub-graphs of strongly connected nodes, with each sub-graph capturing one object pattern. Human communication is based on speech and text. So visually impaired people can gather information from voice. With the help of this project, visually impaired people can read the text present in the captured image. In this Project, we use Raspberry Pi Camera, this help to take pictures and that picture is converted into scan image for further processing by using Image magic software. We use TTS (Text to Speech) engine to transform the text into speech. Experimental results show that the analysis of different captured images will be more helpful to blind people.

Keywords: Object detection, Text to Speech, Raspberry pi, Camera Module.

I. INTRODUCTION

In this project, we will understand what object detection is and look at a few approaches to solve problems in this space. Then we will deep dive into building our object detection system in Python.

Today, there are about 2-3 percent of people of the world's population are blind and have low vision paired with people. Visually impaired people or blind people are those people who lack in their visual perception and they are unable to see the object. We all know that blind people have their script language known as Braille language, which is slightly difficult to learn. But they have hearing capability and this capability makes them hear from the environment. This method is very difficult to perform and requires more practice. To overcome the problem of visually impaired people or blind people reading text and documents. So, we have to design a device which is reading the text or document and

the output of the device is the sound signal. The sound can be easily heard by blind people and they can understand the text in the source of the sound. This system will help most of the people in the world who face the problem of vision loss.

II. LITERATURE SURVEY

Paper name: Object Detection Using Image Processing Author: Fares Jelled, ´ Moscow Institute of Physics & Technology [1]The main objective of this article is to develop an OpenCV-Python code using the Haar Cascade algorithm for object and face detection. Currently, UAVs are used for detecting and attacking infiltrated ground targets. The main drawback of this type of UAV is that sometimes the object is not properly detected, which thereby causes the object to hit the UAV. This project aims to avoid such unwanted collisions and damage to UAVs. UAV is also used for surveillance that uses the Voila-jones algorithm to detect and track humans. This algorithm uses cascade object detector function and vision.

Paper name: Edge Preserving and Multi-Scale Contextual Neural Network for Salient Object Detection. Author: Xiang Wang, Huimin Ma, Member IEEE, Xiaozhi Chen, and Shaodi You.[2]In this paper, we propose a novel edge-preserving and multi-scale contextual neural network for salient object detection. The proposed framework is aiming to address two limits of the existing CNN-based methods. First, region-based CNN methods lack sufficient context to accurately locate salient objects since they deal with each region independently. Second, pixel-based CNN methods suffer from blurry boundaries due to the presence of convolutional and pooling layers. Motivated by these, we first propose an end-to-end edge-preserved neural network based on the Fast R-CNN framework (named RegionNet) to efficiently generate a saliency map with sharp object boundaries. The proposed framework achieves both clear detection boundary and multiscale contextual robustness simultaneously for the first time, and thus achieves an optimized performance. Experiments on six RGB and two RGB-D benchmark datasets demonstrate that the proposed method achieves state-of-the-art performance.

paper name: 3D Object Proposals using Stereo Imagery for Accurate Object Class Detection. Author: Xiaozhi Chen∗ , Kaustav Kundu∗ , Yukun Zhu, Huimin Ma, Sanja Fidler and Raquel Urtasun.[3]In this paper, we propose a novel 3D object detection approach that exploits stereo imagery and contextual information specific to the domain of autonomous driving. We propose a 3D object proposal method that goes beyond 2D bounding boxes and is capable of generating high-quality 3D bounding box proposals. We make use of the 3D information estimated from a stereo camera pair by placing 3D candidate boxes on the ground plane and scoring them via 3D point cloud features. In particular, our scoring function encodes several depth-informed features such as point densities inside a candidate box, free space, visibility, as well as object size priors and height above the ground plane. The inference process is very efficient as all the features can be computed in constant time via 3D integral images.

Paper name: Scalable Object Detection using Deep Neural Networks Author: Christian Szegedy, Dumitru Erhan, Alexander Toshkov Toshev

Deep convolutional neural networks have recently achieved state-of-the-art performance on several image recognition benchmarks, including the ImageNet Large-Scale Visual Recognition Challenge

(ILSVRC-2012). The winning model on the localization sub-task was a network that predicts a single bounding box and a confidence score for each object category in the image. Such a model captures the whole-image context around the objects but cannot handle multiple instances of the same object in the image without naively replicating the number of outputs for each instance. In this work, we propose a saliency-inspired neural network model for detection, which predicts a set of class-agnostic bounding boxes along with a single score for each box, corresponding to its likelihood of containing any object of interest. The model naturally handles a variable number of instances for each class and allows for crossclass generalization at the highest levels of the network. We can obtain competitive recognition performance on VOC2007 and ILSVRC2012 while using only the top few predicted locations in each image and a small number of neural network evaluations.

III. SYSTEM ARCHITECTURE

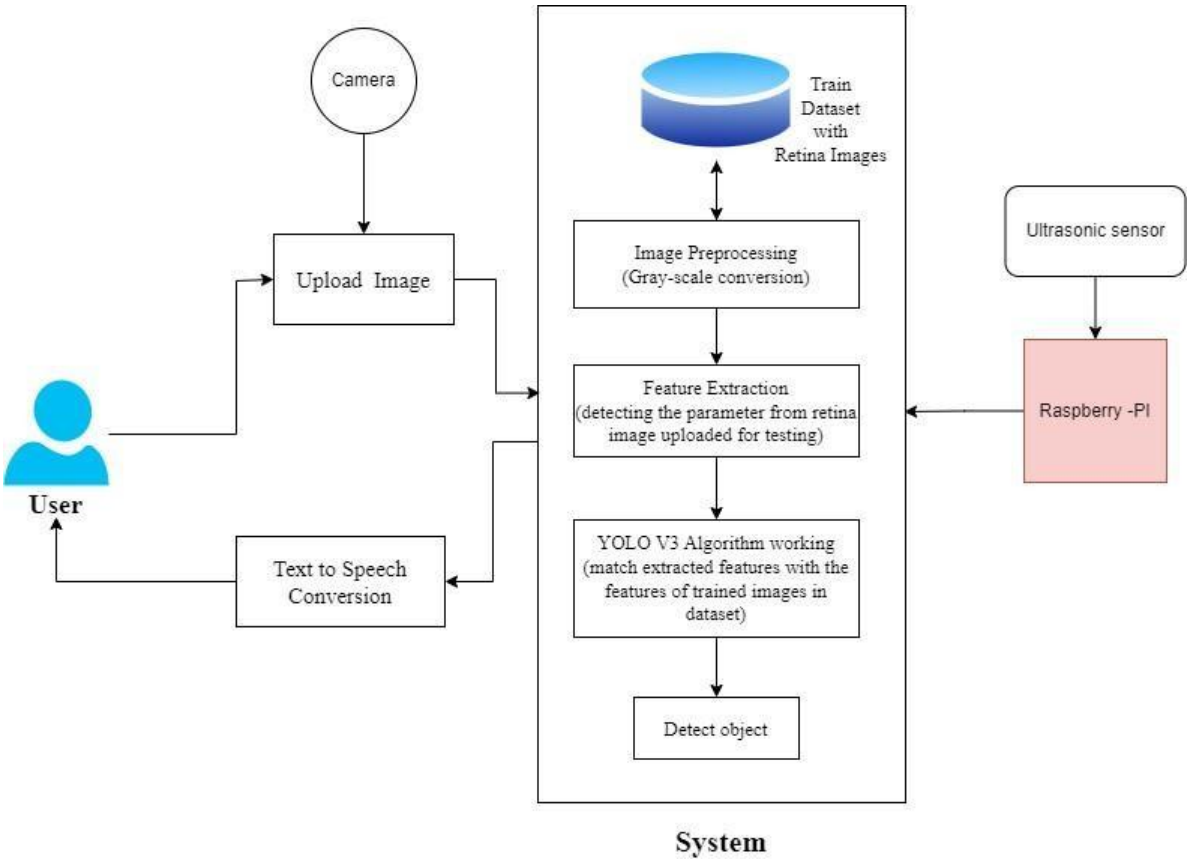
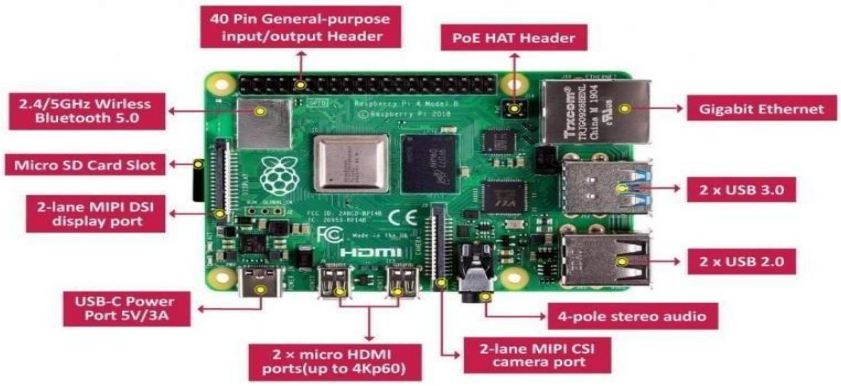


Fig. Block Diagram

III. METHODOLOGY

Raspberry Pi:

Raspberry Pi is defined as a minicomputer the size of a credit card that is interoperable with any input and output hardware device like a monitor, a television, a mouse, or a keyboard – effectively converting the set-up into a full-fledged PC at a low cost **Diagram**



Ultrasonic Sensor module

Ultrasonic Sensor:

An ultrasonic sensor is a device that uses sound waves at frequencies above the range of human hearing (ultrasonic waves) to measure distance, detect objects, and measure the speed of objects. Ultrasonic sensors are widely used in a variety of applications, such as robotics, industrial automation, and automotive systems.

Ultrasonic sensor modules typically consist of a transducer that generates and receives ultrasonic waves, a signal processing circuit, and an interface for connecting to a microcontroller or other device. The sensor emits a high-frequency sound wave, which reflects off of an object and returns to the sensor. The sensor then measures the time it takes for the sound wave to travel to the object and back and calculates the distance to the object using the speed of sound.

Diagram:

Camera Module:

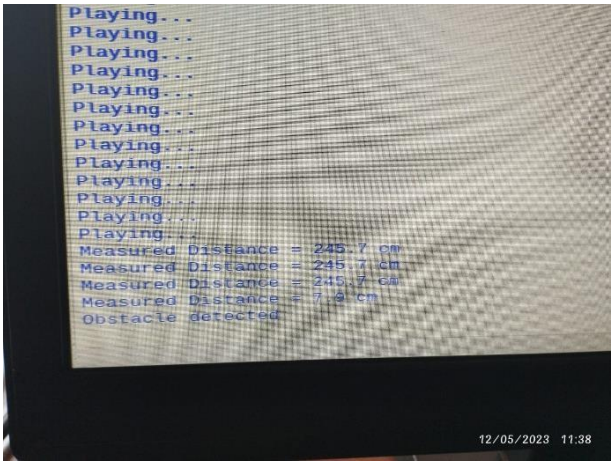
A webcam is a video camera that is designed to record or stream to a computer or computer network. A webcam is a digital video device commonly built into a computer. Its main function is to transmit pictures over the Internet. It is popularly used with instant messaging services and for recording images**.**

WORKING:

In this project, we present a unique smart glass for visually impaired people to overcome traveling difficulties. It can detect the obstacle and measure the distance perfectly using the ultrasonic sensor and a microcontroller (raspberry pi). After receiving information from the environment, it passes to the blind person through a headphone. The ultrasonic sensor will measure the distance between the object and the person and information to the user. When an obstacle comes near the blind within the range of 5 meters, the sensors measure the obstacle distance and send the value to the Raspberry Pi. The SD card module can communicate with the microcontroller. The voices are recorded in the SD card module to instruct the visually impaired users, which is played through a headphone according to the ultrasonic data.
The camera module will use to detect and read the object and will inform the microcontroller.

Result images:

Object recognition



ultrasonic sensor image

VI. CONCLUSION:

We propose a framework for common object discovery and localization in wild images. Like most previous methods which are based on the assumption that there is only one object contained in each positive image, inspired by min-cut/max-flow algorithms, we then present a constrained sub-graph mining algorithm to optimize the two models. We can classify and detect the object by using a neural network correctly. To analyze images and extract high-level information, image enhancement, motion detection, object tracking and behavior understanding research have been studied. In this paper, we have studied and presented different methods of moving object detection, used in video surveillance. We have studied detection techniques into various categories, here, we also discuss the related issues, to the moving object detection technique. The drawback of temporal differencing is that it fails to extract all relevant pixels of a foreground object especially when the object has uniform texture or moves slowly. When a foreground object stops moving, the temporal differencing method fails in detecting a change between consecutive frames and loses the track of the object. This article gives valuable insight into this important research topic and encourages new research in the area of moving object detection as well as in the field of computer vision. In the kernel tracking approach, various estimating methods are used to find the corresponding region to the target object. Now a day, the most preferred and popular kernel tracking techniques are based on Mean-shift tracking and particle filter. Contour tracking can be divided into the state space method and energy function minimization method according to the way of evolving contours.

REFERENCES

[1] Dumitru Erhan, Christian Szegedy, Alexander Toshev, and Dragomir Anguelov Google, Inc. 1600 Amphitheatre Parkway, Mountain View (CA), 94043, USA.

[2] Fares Jalled, ´ Moscow Institute of Physics & Technology, Department of Radio Engineering & Cybernetics Ilia Voronkov, Moscow Institute of Physics & Technology, Department of Radio Engineering & Cybernetics

[3] Xiaozhi Chen∗ , Kaustav Kundu∗ , Yukun Zhu, Huimin Ma, Sanja Fidler and Raquel Urtasun.

[4] Ross Girshick1 Jeff Donahue1,2 Trevor Darrell1,2 Jitendra Malik1 1UC Berkeley and 2 ICSI {rbg,jdonahue,trevor,malik}@eecs.berkeley.edu.

[5] Xiang Wang, Huimin Ma, Member IEEE, Xiaozhi Chen, and Shaodi You