A Project Report

on

Home Automation for the specially-abled

to be submitted in partial fulfilling of the requirements for the course on

Artificial Intelligence – CSE 3013

E2 + TE2

by

AJAY KRISHNAN 18BCE0652

BHAVESH GOYAL 18BCE2521

DIPTO CHAKRABARTY 18BCI0066

RISHABH JAIN 18BCI0078

under the guidance of

Dr. Harshita Patel

Fall Semester 2020-2021

TABLE OF CONTENTS

INTRODUCTION

Today every visually challenged person uses a stick to know his surroundings and move accordingly. In this process many times they lose their balance and trip. They are always asked to wear black sunglasses to prevent any discomfort or dizziness. But this also causes an awareness amongst the people near them, who look down upon them. They have a great sense of hearing which they have developed over the years, which helps them to identify the people who they already know and how they don't. But there is always a chance of human error. On top of that, they can easily be cheated by others.

The physically challenged people can barely move their bodies and have to rely on someone else to help them out in their day-to-day activities. This dependence on others becomes really tiring after some point of time and the people can do nothing but feel helpless about themselves.

The target population for the project will be all the specially-abled people of our society. The hardships faced by them is unimaginable.

As of March 4th, 2019, there were 40 million people in India, including 1.6million children, who are visually impaired due to uncorrected refractive error.

The number of physically challenged people in India count to a population of 9.5million people.

These numbers are not small as they tell us that nearly 3% of our population is either visually impaired or physically challenged.

The lives of such people are not easy. The difficulties they face on a daily basis is something no one can even think of it.

OBJECTIVE

The objective of the project is to provide aid and help make the daily tasks of specially abled people. Our aim is to make the lives of specially abled people easier by providing different functionalities like obstacle avoidance using object detection in a pathway, gesture analysis for the paralyzed, emotion and face recognition, and lie detector.

Models for each of the functionalities will be developed which will be trained on particular datasets. The end aim will be to automate or make the tasks which are performed repeatedly by specially aided people easier. This project aims to make the lives of specially abled people easier.

ABSTRACT

In order to help the specially abled we had decided to make a robust system which helps them in their day to day basis. In order to achieve the above mentioned, we decided to implement four different tasks under one combined home automation system which will help our users to a great extent.

1. Object detection for visually challenged people

2. Step Counter, so that the visually challenged person will know the number of steps ahead.

3. Emotion analysis so that we can help the user know what the person in front of him/her feels.

4. Lie detector using the eyeball movement.

We will implement different models for each of the use cases by training on datasets, the libraries used in the process are stated below. Object Detection will be used to analyze the path and warn the blind people for obstacles present in the path which till now people were dependent on physical contact using walking sticks. Facial Recognition will be used to identify closely related people and strangers for people with disabilities in their homes. Emotion analysis and gesture analysis together are aimed towards the paralyzed or blind and deaf people of the society so they can better understand and communicate their messages to their person they desire .

PRELIMINARIES

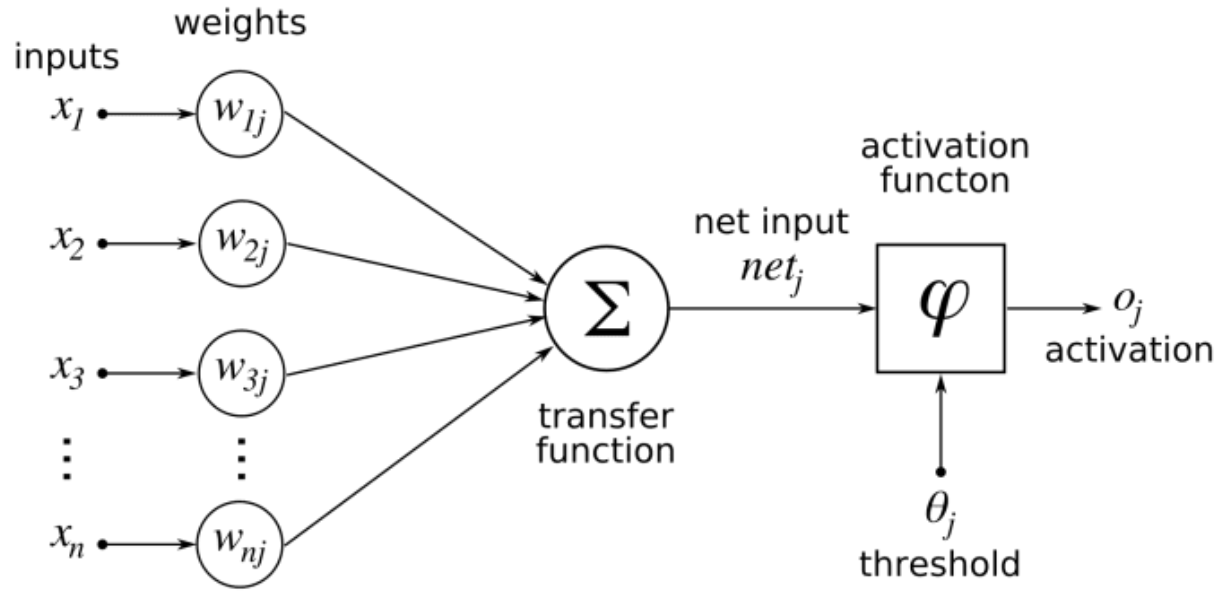
- **CNN**

Convolutional neural network (ConvNets or CNNs) is one of the main categories to do images recognition, images classifications. Objects detections, recognition faces etc., are some of the areas where CNNs are widely used.

Technically, deep learning CNN models to train and test, each input image will pass it through a series of convolution layers with filters (Kernels), Pooling, fully connected layers (FC) and apply Softmax function to classify an object with probabilistic values between 0 and 1.

Convolutional neural networks are composed of multiple layers of artificial neurons. Artificial neurons, a rough imitation of their biological counterparts, are mathematical functions that calculate the weighted sum of multiple inputs and outputs an activation value , behavior of each neuron is defined by its weights.

When fed with the pixel values, the artificial neurons of a CNN pick out various visual features.



When you input an image into a ConvNet, each of its layers generates several activation maps. Activation maps highlight the relevant features of the image. Each of the neurons takes a patch of pixels as input, multiplies their color values by its weights, sums them up, and runs them through the activation function.

- **OPENCV**

OpenCV is a cross-platform library using which we can develop real-time computer vision applications. It mainly focuses on image processing, video capture and analysis including features like face detection and object detection.

The library has more than 2500 optimized algorithms, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine

learning algorithms. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects.

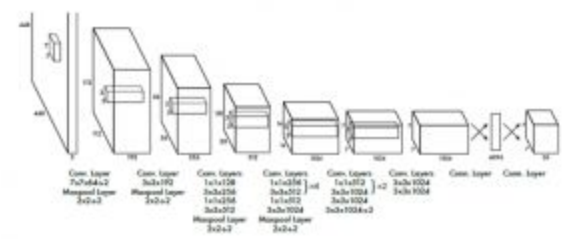
It has C++, Python, Java and MATLAB interfaces and supports Windows, Linux, Android and Mac OS. OpenCV leans mostly towards real-time vision applications and takes advantage of MMX and SSE instructions when available.

- **STREAMLIT**

Streamlit is a python based framework to create smart, robut, and interactive dashboards. It is mainly used to build or deploy machine learning and deep learning softwares and it has been used in this project to showcase all the features built.

- **YOLO**

You Only Look Once (YOLO) is a network that uses Deep Learning (DL) algorithms for object detection. YOLO performs object detection by classifying certain objects within the image and determining where they are located on it.
YOLO uses a totally different approach. YOLO is a clever convolutional neural network (CNN) for doing object detection in real-time. The algorithm applies a single neural network to the full image, and then divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted by the predicted probabilities.



With YOLO, a single CNN simultaneously predicts multiple bounding boxes and class probabilities for those boxes. YOLO trains on full images and directly optimizes detection performance.

Object detection networks such as R-CNN are harder to optimize and slower compared to YOLO. YOLO is much faster (45 frames per second) and easier to optimize than previous algorithms, as it is based on an algorithm that uses only one neural network to run all components of the task.

- **TENSORFLOW**

TensorFlow is an open source library for numerical computation and large-scale machine learning. TensorFlow bundles together a slew of machine learning and deep learning (aka neural networking) models and algorithms and makes them useful by way of a common metaphor. It uses Python to provide a convenient front-end API for building applications with the framework, while executing those applications in high-performance C++.

TensorFlow can train and run deep neural networks for handwritten digit classification, image recognition, word embeddings, recurrent neural networks, sequence-to-sequence models for machine translation, natural language processing,

and PDE (partial differential equation) based simulations. Best of all, TensorFlow supports production prediction at scale, with the same models used for training.

- **KERAS**

Keras is a high-level library that's built on top of Theano or TensorFlow. It provides a scikit-learn type API (written in Python) for building Neural Networks. Developers can use Keras to quickly build neural networks without worrying about the mathematical aspects of tensor algebra, numerical techniques, and optimization methods.

Keras offers a very easy to use as well as intuitive enough to understand API which essentially helps you test and build Deep Learning applications with least considerable effort. Keras is a high-level interface and uses Theano or Tensorflow for its backend.

Keras supports almost all the models of a neural network – fully connected, convolutional, pooling, recurrent, embedding, etc. Furthermore, these models can be combined to build more complex models.It runs smoothly on both CPU and GPU.

METHODOLOGY

•PYTHON3(The Language Used For Building Everything)

•KERAS(used for building emotion recognition and gesture analysis)

•TENSORFLOW(used for building a face recognition)

•YOLOV3(used for building object detection model)

•OPENCV(used as a pre-processing technique for training all models and building a lie detector)

Based on the dataset considered for each model we train and test each separate model. For Object Detection we use CNN model with a dataset of everyday objects which will be trained to recognise each item so that a user will be able to identify obstacles in the path. Gesture Analysis is based on training a model using different images of peoples gestures which can recognise gestures for paralzed people. Lie detector is based on the eye movement of a person based on the eye movement and distance between the eyes of a person. We are able to identify with a minimal accuracy whether a person is lying or not . Facial Recognition is based on identifying known and unknown users using convolutional models.

EMOTION RECOGNITION

DATASET USED : FER-2013

The data consists of 48x48 pixel grayscale images of faces. The faces have been automatically registered so that the face is more or less centred and occupies about the same amount of space in each image.

The task is to categorize each face based on the emotion shown in the facial expression into one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral). The training set consists of 28,709 examples and the public test set consists of 3,589 examples.

OUR CONTRIBUTION

The following architecture is adapted for building the emotion recognition model

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 46, 46, 64)        640
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 44, 44, 64)        36928
_________________________________________________________________
batch_normalization (BatchNo (None, 44, 44, 64)        256
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 22, 22, 64)        0
_________________________________________________________________
dropout (Dropout)            (None, 22, 22, 64)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 20, 20, 64)        36928
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 18, 18, 64)        36928
_________________________________________________________________
batch_normalization_1 (Batch (None, 18, 18, 64)        256
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 9, 9, 64)          0
_________________________________________________________________
dropout_1 (Dropout)          (None, 9, 9, 64)          0
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 7, 7, 128)         73856
_________________________________________________________________
```

```
conv2d_5 (Conv2D)            (None, 5, 5, 128)         147584
_________________________________________________________________
batch_normalization_2 (Batch (None, 5, 5, 128)         512
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 2, 2, 128)         0
_________________________________________________________________
flatten (Flatten)            (None, 512)               0
_________________________________________________________________
dense (Dense)                (None, 1024)              525312
_________________________________________________________________
dropout_2 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_1 (Dense)              (None, 1024)              1049600
_________________________________________________________________
dropout_3 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_2 (Dense)              (None, 7)                 7175
=================================================================
Total params: 1,915,975
Trainable params: 1,915,463
Non-trainable params: 512
_________________________________________________________________
```

LIE DETECTION

It is a well known fact that when a person is saying the truth, he/she will say it confidently without blinking much and if the same person is saying the lie he/she will blink a lot and stammer in between. For a visually challenged person, tracking the eye ball movement is not possible and this ability is provided through this project. The initial location of the eyeball is captured and using that as a reference we keep computing the movement of the eyeball of the person in front of them, if the standard deviation of the movement crosses a certain limit then the user is notified if the person is telling the truth or not.

OBJECT DETECTION

At present a visually challenged person is equipped with only a walking stick which he uses to move around, to locate objects, using our object detection we will give our users the knowledge of an object which is in front of them, so that their dependence on a moving stick will be reduced.

IMPLEMENTATION

```
import os
import cv2
import numpy as np
from tensorflow.keras import models
from keras.preprocessing import image
model = models.load_model('fer.h5')
face_haar_cascade = cv2.CascadeClassifier('haarcascade_frontalface_default.xml')
cap=cv2.VideoCapture(0)
while True:
    ret,test_img=cap.read()# captures frame and returns boolean value and captured
image
    if not ret:
        continue
    gray_img= cv2.cvtColor(test_img, cv2.COLOR_BGR2GRAY)
    faces_detected = face_haar_cascade.detectMultiScale(gray_img, 1.32, 5)
    for (x,y,w,h) in faces_detected:
        cv2.rectangle(test_img,(x,y),(x+w,y+h),(255,0,0),thickness=7)
        roi_gray=gray_img[y:y+w,x:x+h]
```

```
        roi_gray=cv2.resize(roi_gray,(48,48))
        img_pixels = image.img_to_array(roi_gray)
        img_pixels = np.expand_dims(img_pixels, axis = 0)
        img_pixels /= 255
        predictions = model.predict(img_pixels)
        max_index = np.argmax(predictions[0])
        emotions = ('angry', 'disgust', 'fear', 'happy', 'sad', 'surprise', 'neutral')
        predicted_emotion = emotions[max_index]

        cv2.putText(test_img, predicted_emotion, (int(x), int(y)),
cv2.FONT_HERSHEY_SIMPLEX, 1, (0,0,255), 2)

    resized_img = cv2.resize(test_img, (1000, 700))
    cv2.imshow('Facial emotion analysis ',resized_img)
    if cv2.waitKey(10) == ord('q')
        break
cap.release()
cv2.destroyAllWindows

import cv2
import numpy as np
import os
s=set()
s.add('person')
a=1
net = cv2.dnn.readNet("yolov3.weights", "yolov3.cfg")
```

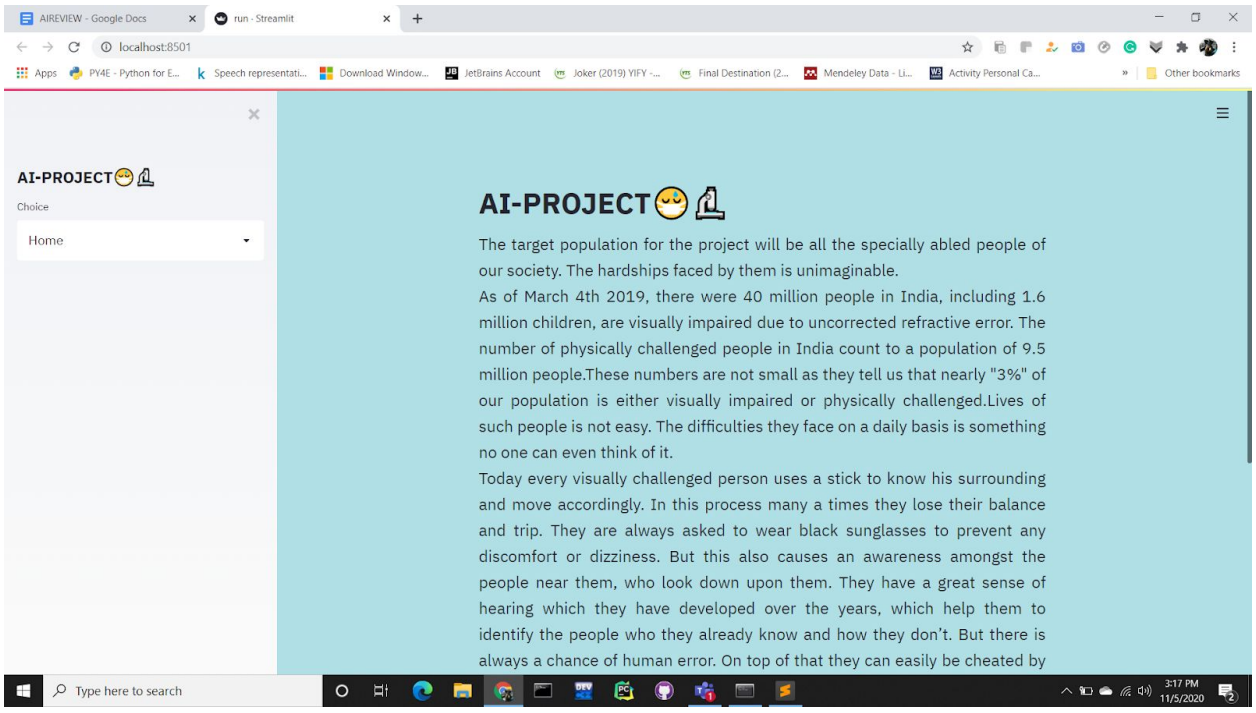
```
classes = []

with open("coco.names", "r") as f:
    classes = [line.strip() for line in f.readlines()]

layer_names = net.getLayerNames()
output_layers = [layer_names[i[0] - 1] for i in net.getUnconnectedOutLayers()]
colors = np.random.uniform(0, 255, size=(len(classes), 3))

def yolo(img):
    img = cv2.resize(img, None, fx=0.4, fy=0.4)
    height, width, channels = img.shape
    blob = cv2.dnn.blobFromImage(img, 0.00392, (416, 416), (0, 0, 0), True,
crop=False)
    net.setInput(blob)
    outs = net.forward(output_layers)
    class_ids=[]
    confidences = []
    boxes = []
    for out in outs:
        for detection in out:
            scores = detection[5:]
            class_id = np.argmax(scores)
            confidence = scores[class_id]
            if confidence > 0.5:
                center_x = int(detection[0] * width)
```

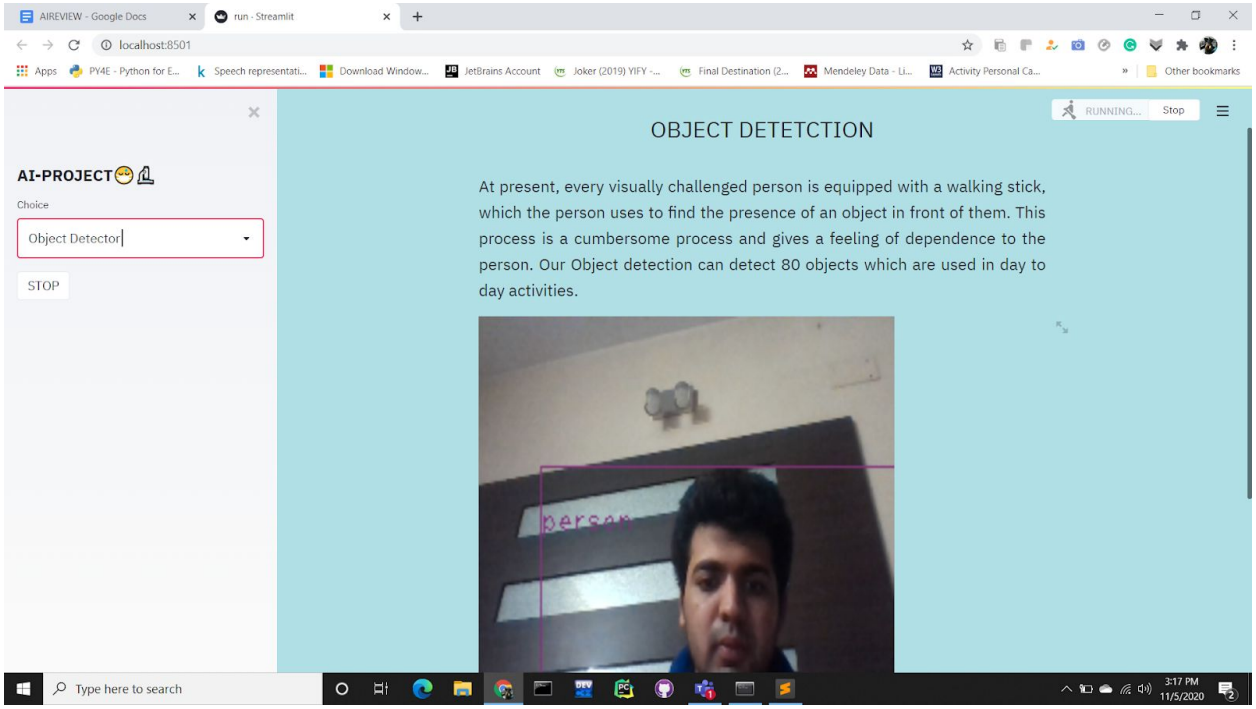
```
            center_y = int(detection[1] * height)
            w = int(detection[2] * width)
            h = int(detection[3] * height)
            x = int(center_x - w / 2)
            y = int(center_y - h / 2)
            boxes.append([x, y, w, h])
            confidences.append(float(confidence))
            class_ids.append(class_id)
    indexes = cv2.dnn.NMSBoxes(boxes, confidences, 0.5, 0.4)
    font = cv2.FONT_HERSHEY_PLAIN
    for i in range(len(boxes)):
        if i in indexes:
            x, y, w, h = boxes[i]
            label = str(classes[class_ids[i]])
            if label!='person':
                s.add(label)
                print(s)
                if len(s)>a:
                    te="There is a {} in front of you".format(label)
            color = colors[i]
            cv2.rectangle(img, (x, y), (x + w, y + h), color, 1)
            cv2.putText(img, label, (x, y + 30), font, 1, color, 1)
    img=cv2.resize(img,(512,512))
    return img


import time
```

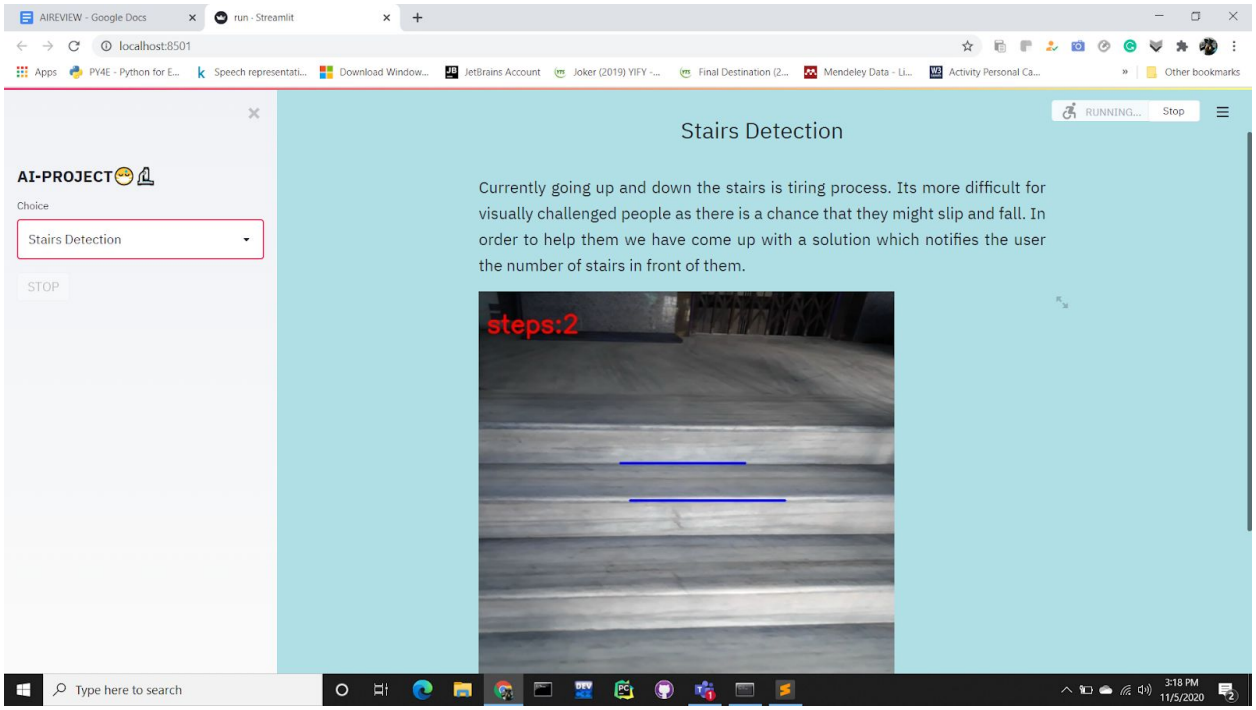
```
import numpy as np
import cv2
from gaze_tracking import GazeTracking
flag=1
gaze = GazeTracking()
#webcam = cv2.VideoCapture(0)
x=[]
y=[]
a=time.time()
def lie(frame):
    gaze.refresh(frame)

    frame = gaze.annotated_frame()
    text = ""

    if gaze.is_blinking():
        text = "Blinking"
    elif gaze.is_right():
        text = "Looking right"
    elif gaze.is_left():
        text = "Looking left"
    elif gaze.is_center():
        text = "Looking center"

    cv2.putText(frame, text, (90, 60), cv2.FONT_HERSHEY_DUPLEX, 1.6, (147,
58, 31), 2)
```

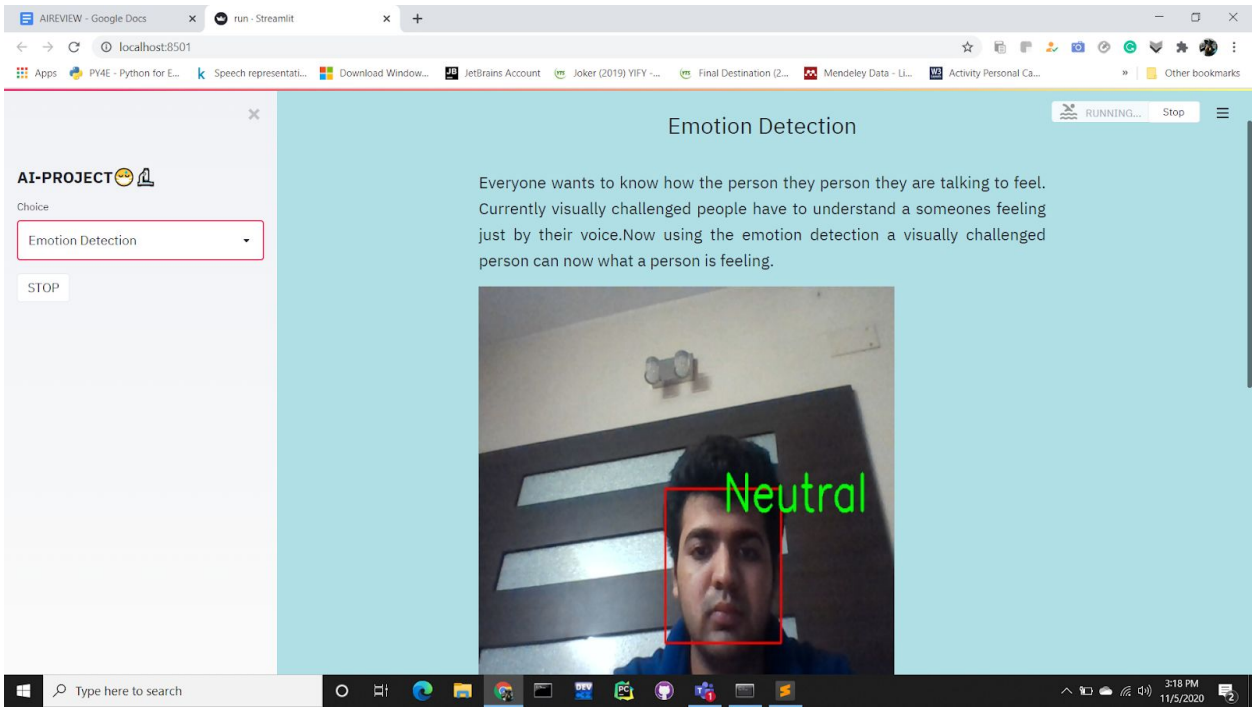
```
    left_pupil = gaze.pupil_left_coords()
    right_pupil = gaze.pupil_right_coords()
    if left_pupil!=None:
        x.append(left_pupil)
    cv2.putText(frame, "Left pupil:  " + str(left_pupil), (90, 130),
cv2.FONT_HERSHEY_DUPLEX, 0.9, (147, 58, 31), 1)
    cv2.putText(frame, "Right pupil: " + str(right_pupil), (90, 165),
cv2.FONT_HERSHEY_DUPLEX, 0.9, (147, 58, 31), 1)
    frame = cv2.resize(frame,(512,512),interpolation=cv2.INTER_AREA)
    b=time.time()
    if(b-a>15):
        print(x)
        x1=np.asarray(x)
        print(x1)
        stdx=x1.std()
        print(stdx)
        flag=0
    return frame
```

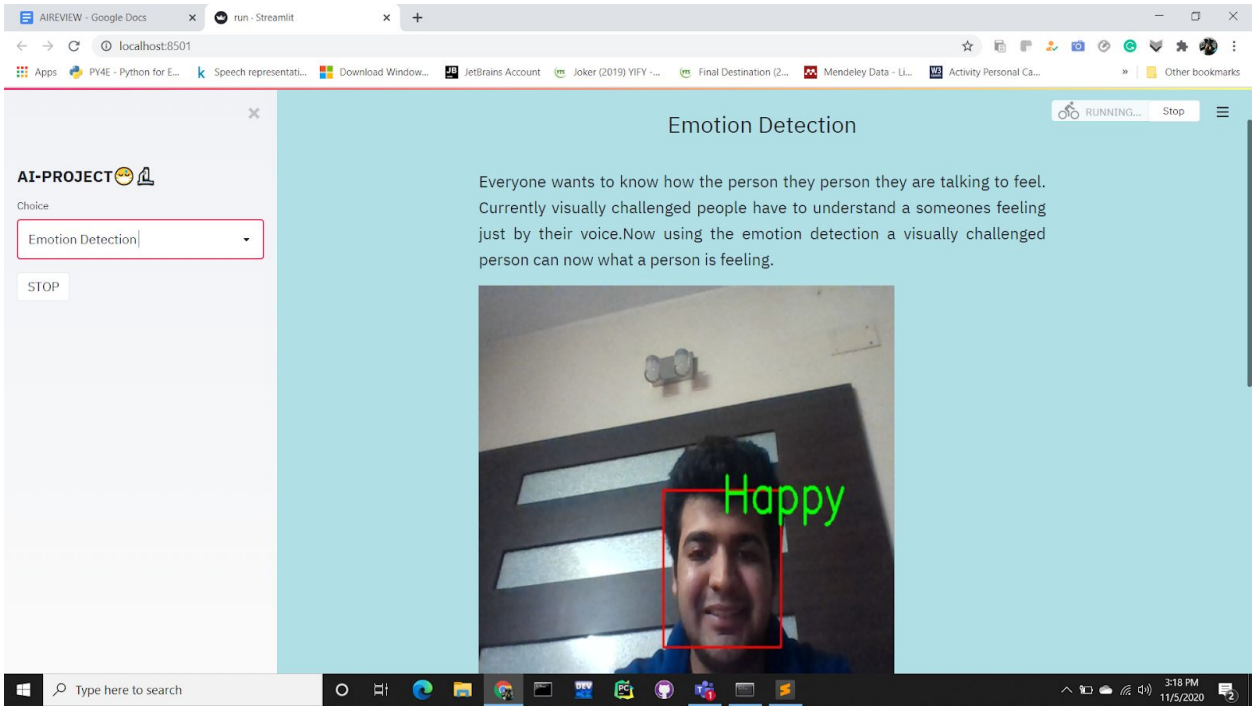
SCREENSHOTS

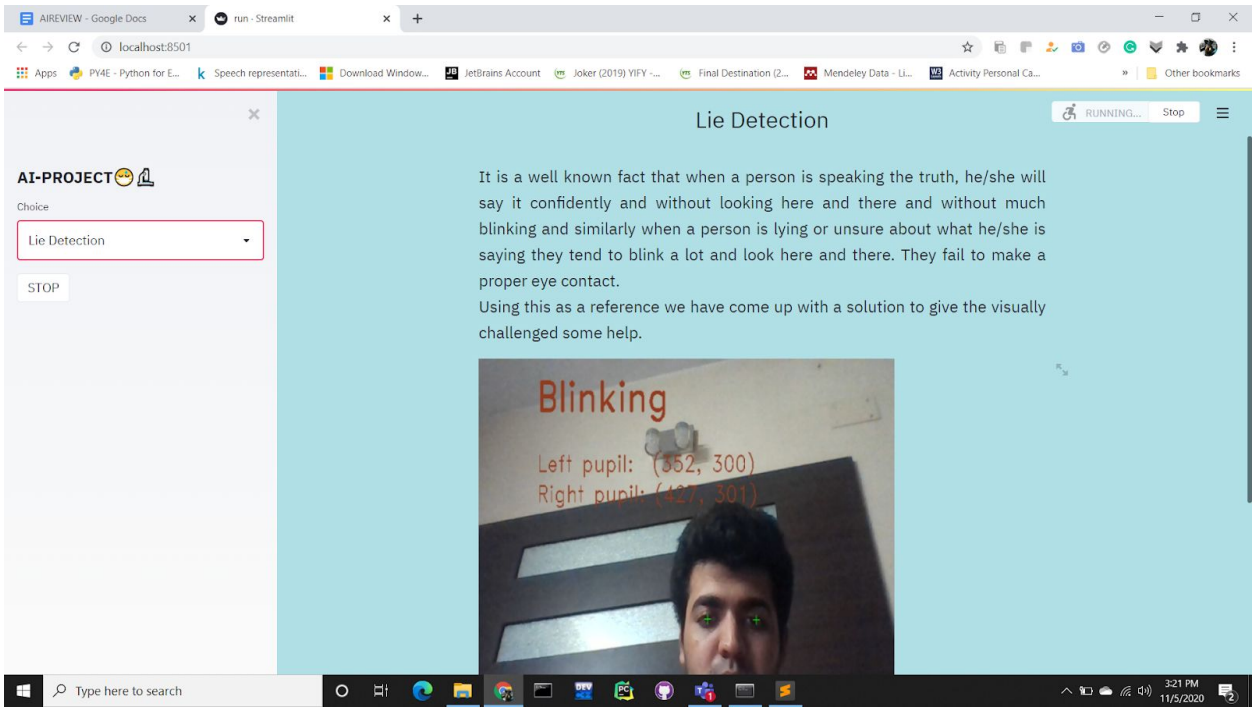


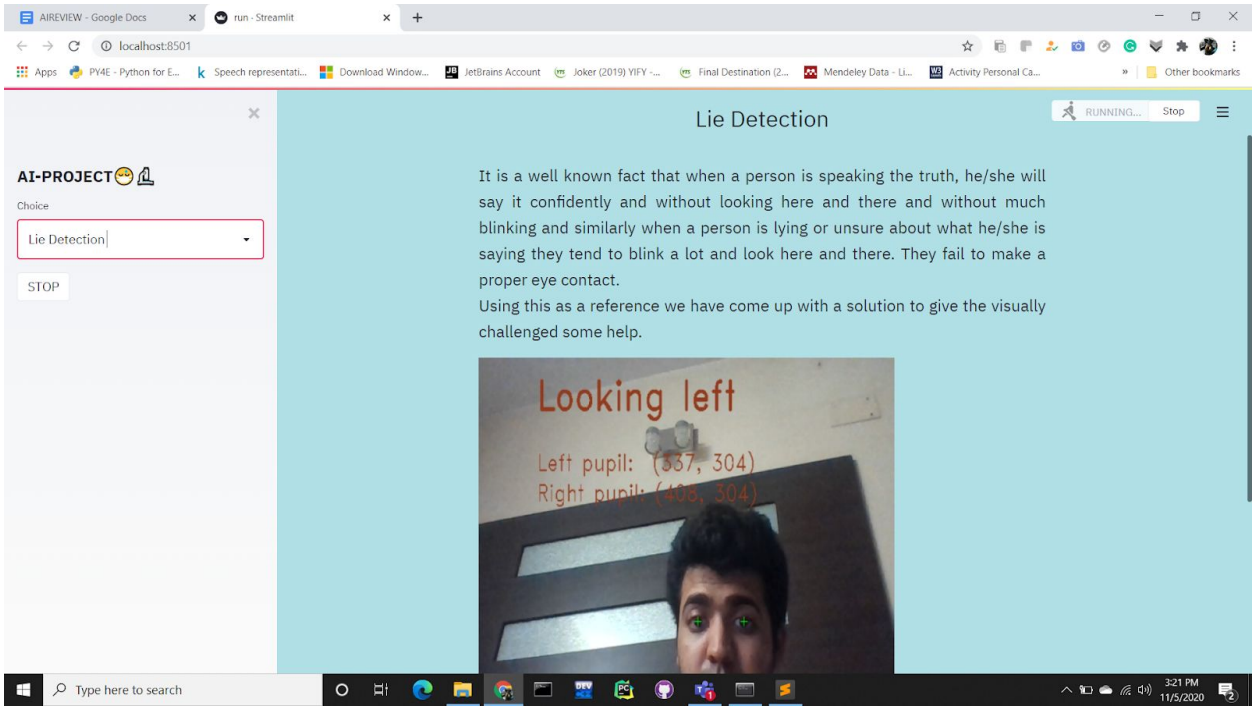



AIREVIEW - Google Docs
run - Streamlit
localhost:8501
AI-PROJECT
Choice
Stairs Detection
STOP
RUNNING...
Stop
Stairs Detection
Currently going up and down the stairs is tiring process. Its more difficult for visually challenged people as there is a chance that they might slip and fall. In order to help them we have come up with a solution which notifies the user the number of stairs in front of them.
steps:2
Type here to search
3:18 PM
11/5/2020


AIREVIEW - Google Docs
run - Streamlit
localhost:8501
AI-PROJECT
Choice
Emotion Detection
STOP
RUNNING...
Stop
Emotion Detection
Everyone wants to know how the person they person they are talking to feel. Currently visually challenged people have to understand a someones feeling just by their voice.Now using the emotion detection a visually challenged person can now what a person is feeling.
Neutral
Type here to search
3:18 PM
11/5/2020


AIREVIEW - Google Docs
run - Streamlit
localhost:8501
AI-PROJECT
Choice
Emotion Detection
STOP
Emotion Detection
RUNNING...
Stop
Everyone wants to know how the person they person they are talking to feel. Currently visually challenged people have to understand a someones feeling just by their voice.Now using the emotion detection a visually challenged person can now what a person is feeling.
Happy
Type here to search
3:18 PM
11/5/2020


AIREVIEW - Google Docs
run - Streamlit
localhost:8501
AI-PROJECT
Choice
Lie Detection
STOP
Lie Detection
RUNNING...
Stop
It is a well known fact that when a person is speaking the truth, he/she will say it confidently and without looking here and there and without much blinking and similarly when a person is lying or unsure about what he/she is saying they tend to blink a lot and look here and there. They fail to make a proper eye contact.
Using this as a reference we have come up with a solution to give the visually challenged some help.
Blinking
Left pupil: (352, 300)
Type here to search
3:21 PM
11/5/2020



LITERATURE SURVEY

1) Z. Habibah, M. Nasrun and C. Setianingsih, "Lie Detector With The Analysis Of The Change Of Diameter Pupil and The Eye Movement Use Method Gabor Wavelet Transform and Decision Tree," 2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS), Bali, 2018, pp. 214-220, doi: 10.1109/IOTAIS.2018.8600918.

Abstract-The lie is very detrimental to the fraudulent acts of many people who were cheated. The lies are common in the general population. To be able to reveal a lie we can detect through some limbs that unconsciously will show a different reaction when someone is lying. Among them, the organs of our eyes can detect if someone is lying or not. Lie detection discussed in this final Task is the eyes, namely with the object of eye tracking and eye pupil diameter changes by using the

method of Wavelet Transform to Gabor Image Processing process and afterward perform the classification to determine the answer someone is lying or not by using a Decision Tree. The existence of this lie detector is expected to be helpful for people who need to detect lies. The final test results are accurate. This research has a precision value of 97%, 94%, and recall accuracy 95% of testing has been done.

https://ieeexplore.ieee.org/document/8600918

2) Z. Labibah, M. Nasrun and C. Setianingsih, "Lie Detector With The Analysis Of The Change Of Diameter Pupil and The Eye Movement Use Method Gabor Wavelet Transform and Decision Tree," 2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS), Bali, 2018, pp. 214-220, doi: 10.1109/IOTAIS.2018.8600918.

Abstract - Every human has different attributes and behaviors. In our life, many people often lie in everything. The lie is obliged to build an understanding of others, then the understanding it establishes is wrong. Indicator of knowing people lies or not by looking at pupils, because according to the science of psychology, pupils will be dilated when the person is in a state of depression, including when lying. Eye blinking can also be an indicator to know people who lie or not by increasing the number of blinking. From the results of the research conducted lie detector. Accuracy is obtained from observations on the change in pupil dilated using a circular hough transform method and increase the number of a blink of the eye by frame difference method. In the study, a person who lies will experience pupil dilatation of 4% to 8% of the initial pupil diameter and an increase in the number of blinking of the eye up to 8 times from the initial blink of the respondent

before the question is asked. In this research the accuracy of the lie detector system is 84%.

https://ieeexplore.ieee.org/document/8226697

3) P. Shanmugavadivu and A. Kumar, "Rapid face detection and annotation with loosely face geometry," 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), Noida, 2016, pp. 594-597, DOI: 10.1109/IC3I.2016.7918032.

Abstract: Recent advancements in research paved way for face can be detected in the cluttered background in real-time. However, much attention needs to be given still in partially occluded face detection with varying degrees of face occlusion. This paper presents a strategic approach for the rapid detection and annotation of the partially occluded faces. The partially Occluded Face Detection (POFD) problem is addressed by using a combination of feature-based and part-based face detection methods with the help of a face part dictionary. In this approach, the devised algorithm aims to automatically detect face components individually and it starts from a mostly un-occluded face component called Nose. The nose is very hard to cover up without drawing suspicion. Keeping the noise component as a reference, the algorithm searches the surrounding area for other main facial features if any. Once face parts qualify facial geometry, they are normalized (scale and rotational) and tag with annotation about each facial feature so that the partial face recognition algorithm can be adapted accordingly with the test image.

https://ieeexplore.ieee.org/document/7918032

4) Quan you Zhao and Shujun Zhang, "A face detection method based on corner verifying," 2011 International Conference on Computer Science and Service System (CSSS), Nanjing, 2011, pp. 2854-2857, doi: 10.1109/CSSS.2011.5974784.

Abstract-Face detection is the first step for automatic face recognition system and many surveillance systems, it has been widely applied in many fields. A face detection method based on corner verifying is presented. Firstly, most of the background area is quickly filtered out by skin color detection. Then faces are detected by the AdaBoost face detection algorithm combined with skin color detection, but there are some non-faces to be detected. Finally, the faces are verified by corner distribution of human face after Harris corner detection to further reduce the error rate. Experimental results show that this method can efficiently and accurately detect the faces with multi-feature fusion.

https://ieeexplore.ieee.org/document/5974784

5) K. Dang and S. Sharma, "Review and comparison of face detection algorithms," 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, Noida, 2017, pp. 629-633, doi: 10.1109/CONFLUENCE.2017.7943228.

Abstract-With the tremendous increase in video and image databases, there is a great need of automatic understanding and examination of data by the intelligent systems as manually it is becoming out of reach. Narrowing it down to one specific domain, one of the most specific objects that can be traced in the images are people i.e. faces. Face detection is becoming a challenge by its increasing use in a number of applications. It is the first step for face recognition, face analysis and detection of other features of face. In this paper, various face detection algorithms are

discussed and analyzed like Viola-Jones, SMQT features & SNOW Classifier, Neural Network-Based Face Detection and Support Vector Machine-Based face detection. All these face detection methods are compared based on the precision and recall value calculated using a DetEval Software which deals with precised values of the bounding boxes around the faces to give accurate results.

https://ieeexplore.ieee.org/document/7943228

6) A. Raghunandan, Mohana, P. Raghav and H. V. R. Aradhya, "Object Detection Algorithms for Video Surveillance Applications," 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, 2018, pp. 0563-0568, doi: 10.1109/ICCSP.2018.8524461.

Abstract-Object Detection algorithms find application in various fields such as defence, security, and healthcare. In this paper various Object Detection Algorithms such as face detection, skin detection, colour detection, shape detection, target detection are simulated and implemented using MATLAB 2017b to detect various types of objects for video surveillance applications with improved accuracy. Further, various challenges and applications of Object Detection methods are elaborated.

https://ieeexplore.ieee.org/document/8524461

7) S. Kanimozhi, G. Gayathri and T. Mala, "Multiple Real-time object identification using Single shot Multi-Box detection," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019, pp. 1-5, doi: 10.1109/ICCIDS.2019.8862041.

Abstract-Real time object detection is one of the challenging task as it need faster computation power in identifying the object at that time. However the data generated by any real time system are unlabelled data which often need large set of labeled data for effective training purpose. This paper proposed a faster detection method for real time object detection based on convolution neural network model called as Single Shot Multi-Box Detector(SSD).This work eliminates the feature resampling stage and combines all calculated results as a single component. Still there is a need for a lightweight network model for the places which lack computational power like mobile devices(eg: laptop, mobile phones, etc). Thus a lightweight network model which uses depth-wise separable convolution called MobileNet is used in this proposed work. Experimental results reveal that use of MobileNet along with SSD models increase the accuracy level in identifying the real time household objects.

https://ieeexplore.ieee.org/document/8862041

8) T. Ju, W. Lu, K. Chen and J. Guo, "Vision-based moving objects detection for intelligent automobiles and a robustness enhancing method," 2014 IEEE International Conference on Consumer Electronics - Taiwan, Taipei, 2014, pp. 75-76, doi: 10.1109/ICCE-TW.2014.6904109.

Abstract-This paper presents a vision-based moving objects detection work which attracts much attention in intelligent automobile applications recently. Vision-based objects detection provides object behavior information of objects and is an intuitive detection method similar to human visual perception. Besides, vision-based objects detection methods are much low-cost compared with detection methods such as RADAR (Radio Detection And Ranging), or LiDAR

(Light Detection And Ranging). However, current vision-based objects detection methods still suffer from several challenges such as high false alarms and unstable detection rate which limit their value in practical applications. Accordingly, this paper presents a robustness enhancing method for vision-based moving objects detection.

https://ieeexplore.ieee.org/document/6904109

9) M. R. Malgireddy, J. J. Corso, S. Setlur, V. Govindaraju and D. Mandalapu, "A Framework for Hand Gesture Recognition and Spotting Using Sub-gesture Modeling," 2010 20th International Conference on Pattern Recognition, Istanbul, 2010, pp. 3780-3783, doi: 10.1109/ICPR.2010.921.

Abstract-Hand gesture interpretation is an open research problem in Human Computer Interaction (HCI), which involves locating gesture boundaries (Gesture Spotting) in a continuous video sequence and recognizing the gesture. Existing techniques model each gesture as a temporal sequence of visual features extracted from individual frames which is not efficient due to the large variability of frames at different timestamps. In this paper, we propose a new sub-gesture modeling approach which represents each gesture as a sequence of fixed sub-gestures (a group of consecutive frames with locally coherent context) and provides a robust modeling of the visual features. We further extend this approach to the task of gesture spotting where the gesture boundaries are identified using a filler model and gesture completion model. Experimental results show that the proposed method outperforms state-of-the-art Hidden Conditional Random Fields (HCRF) based methods and baseline gesture spotting techniques.

https://ieeexplore.ieee.org/document/5597566

10) C. Nagadeepa, N. Balaji and V. Padmaja, "An Efficient Framework for 2-Dimensional Gesture Based Telugu Character Recognition," 2016 IEEE 6th International Conference on Advanced Computing (IACC), Bhimavaram, 2016, pp. 446-450, doi: 10.1109/IACC.2016.89.

Abstract-Gesture identification plays a vital role in today's human-computer interaction. In this paper, we proposed a sensor based gesture recognition system which makes the teacher write in Telugu language on a digital board from anywhere within the classroom. Various classification algorithms k-Nearest Neighbor (KNN), Support Vector Machine (SVM) and Decision tree are individually used for hand gesture based Telugu character recognition. Here, we assess the performance of three classification algorithms which are compared with 16 different Telugu character vowel gestures. Each gesture is collected by using an inertial sensor based embedded device. The dataset contains 16 gestures, each gesture repeated for eleven times from three different people. The gesture identification accuracy for k-Nearest Neighbor classification is 97.2%, SVM is 92.8% and Decision tree is 86.5%.

https://ieeexplore.ieee.org/document/7544878

11) M. Elmezain, A. Al-Hamadi and B. Michaelis, "A Robust Method for Hand Gesture Segmentation and Recognition Using Forward Spotting Scheme in Conditional Random Fields," 2010 20th International Conference on Pattern Recognition, Istanbul, 2010, pp. 3850-3853, doi: 10.1109/ICPR.2010.938.

Abstract—This paper proposes a forward spotting method that handles hand gesture segmentation and recognition simultaneously without time delay. To spot meaningful gestures of numbers (0-9) accurately, a stochastic method for designing

a non-gesture model using Conditional Random Fields (CRFs) is proposed without training data. The non-gesture model provides confidence measures that are used as an adaptive threshold to find the start and the end point of meaningful gestures. Experimental results show that the proposed method can successfully recognize isolated gestures with 96.51% and meaningful gestures with 90.49% reliability.

https://ieeexplore.ieee.org/document/5597659

12) A Novel Approach for Lie Detection Based on F-Score and Extreme Learning Machine by Junfeng Gao,Zhao Wang,Yong Yang,Wenjia Zhang Published: June 3, 2013

https://doi.org/10.1371/journal.pone.0064704

Abstract - Though there exists different methods of detecting lies, machine learning can be used to create a reliable and more efficient system to detect lies. This thesis proposes a method of using nonverbal human behaviors to detect lies using machine learning. This lie detection system is based on micro-expressions of human beings which uses Facial Landmark Detection System and Azure Machine Learning. Movements of individual facial muscles are recorded while a person answers some certain questions. By using the two algorithms TwoClass Support Vector Machine and Linear Regression, we attempted to create a machine that can detect lies. We reached an accuracy of approximately 76.2%.

13) F. G. Rodrıguez-Telles, R. Perez-Alcocer, A. Maldonado-Ramirez,L. Abril Torres-Mendez, B. B. Dey, and E. A. Martinez-Garcia, “Vision-based reactive autonomous navigation with obstacle avoidance: Towards A non-invasive and

cautious exploration of marine habitat," in IEEE International Conference on Robotics and Automation (ICRA), 2014,pp. 3813–3818

Abstract - This paper describes a vision-based obstacle avoidance strategy using Deep Learning for Autonomous Underwater Vehicles (AUVs) equipped with simple colored monocular cameras. For each input image, our method uses a deep neural network to compute a transmission map that can be understood as a relative depth map. The transmission map is estimated for each patch of the image to determine the obstacles nearby. This map enables us to identify the most appropriate Region of Interest (RoI) and to find a direction of escape. This direction allows the robot to avoid obstacles by performing a control action. We evaluate our approach in two underwater video sequences. The results show the approach is able to successfully find a RoI that avoids coral reefs, fish, the seafloor and any other object present in the scene.

14) B.J.A. Krose, J.W.M van Dam: "Learning to avoid collisions: a reinforcement learning paradigm for mobile robot navigation", Proceedings of IFAC/IFIP/IMACS Symposium on Artificial Intelligence in Real-Time Control,pp. 295-30, 1992

Abstract - One of the basic issues in the navigation of autonomous mobile robots is the obstacle avoidance task that is commonly achieved using a reactive control paradigm where a local mapping from perceived states to actions is acquired. A control strategy with learning capabilities in an unknown environment can be obtained using reinforcement learning where the learning agent is given only sparse reward information. This credit assignment problem includes both temporal and structural aspects. While the temporal credit assignment problem is solved

using core elements of the reinforcement learning agent, solution of the structural credit assignment problem requires an appropriate internal state space representation of the environment. In this paper, a discrete coding of the input space using a neural network structure is presented as opposed to the commonly used continuous internal representation. This enables a faster and more efficient convergence of the reinforcement learning process.

15) High Speed Obstacle Avoidance using Monocular Vision and Reinforcement Learning , Computer Science Department, Stanford University, Stanford, CA 94305 USA

Abstract - The task of driving a remote control car at high speeds through unstructured outdoor environments. It is to present an approach in which supervised learning is first used to estimate depths from single monocular images. The learning algorithm can be trained either on real camera images labeled with ground-truth distances to the closest obstacles, or on a training set consisting of synthetic graphics images. The resulting algorithm is able to learn monocular vision cues that accurately estimate the relative depths of obstacles in a scene. Reinforcement learning/policy search is then applied within a simulator that renders synthetic scenes. This learns a control policy that selects a steering direction as a function of the vision system's output. We present results evaluating the predictive ability of the algorithm both on held out test data, and in actual autonomous driving experiments.

16) An improved face recognition algorithm and its application in attendance management system by Serign Modou Bah FangMin ,https://doi.org/10.1016/j.array.2019.100014 , Under a Creative Commons license

Abstract - Face Recognition is a computer application that is capable of detecting, tracking, identifying or verifying human faces from an image or video captured using a digital camera. Although a lot of progress has been made in the domain of face detection and recognition for security, identification and attendance purpose, there are still issues hindering the progress to reach or surpass human level accuracy. These issues are variations in human facial appearance such as; varying lighting condition, noise in face images, scale, pose etc. This research paper presents a new method using Local Binary Pattern (LBP) algorithm combined with advanced image processing techniques such as Contrast Adjustment, Bilateral Filter, Histogram Equalization and Image Blending to address some of the issues hampering face recognition accuracy so as to improve the LBP codes, thus improve the accuracy of the overall face recognition system. Our experiment results show that our method is very accurate, reliable and robust for face recognition systems that can be practically implemented in real-life environments.

17) FaceNet: A Unified Embedding for Face Recognition and Clustering Published / Last Updated – June 17th, 2015 , Authors and Contributors – Florian Schroff, Dmitry Kalenichenko, and James Philbin, from Google Inc.

In this paper, the authors present a face recognition system called FaceNet. This system uses a deep convolutional neural network which optimizes the embedding, rather than using an intermediate bottleneck layer. The authors state that the most important aspect of this method is the end-to-end learning of the system.

The team trained the convolutional neural network on a CPU cluster for 1,000 to 2,000 hours. They then evaluated their method on four datasets. Notably, FaceNet attained an accuracy of 99.63% on the famous Labeled Faces in the Wild (LFW) dataset, and 95.12% on the Youtube Faces Database.

18) Hand gesture recognition , Published in**:** 2017 7th International Conference onCommunication Systems and Network Technologies (CSNT)

The idea of gesture recognition is to develop a technique which can identify particular human gestures and use them to transfer information. In gesture recognition methods, a camera displays the human body movements and communicates the data to a computer that uses the gestures as input to control gadgets or applications. The idea of developing hand gesture recognition procedure is to develop an interaction between human and computer and the recognized gestures are used to control the meaningful information [1]. The principal components of a hand gesture recognition process are data acquisition, hand localization, hand characteristics recognition and gesture identification. Gesture is an indication of physical behavior or emotional expression. It consists of body gesture and hand gesture [2]. It divides into two categories: static gesture and dynamic gesture. Static hand gesture recognition performed without any additional devices which are used to give instructions to the machine. A person commands the machine using his bare hands, the person hand's images are captured. and analyzed to determine the meaning of hand gesture. To recognize static gestures a common classifier or template matcher is used. A dynamic movement is supposed to be a route between the first phase and last phase. Dynamic gesture recognition characterizes the hand movements. It has four features: - velocity, shape, location

and orientation. Gesture recognition has many uses in coming times like: - in medical application, entertainment application, automated systems, better life for the disabled people.

19) G. R. S. Murthy, R. S. Jadon. (2009). "A Review of Vision Based Hand Gestures Recognition,"International Journal of Information Technology and Knowledge Management, vol. 2(2), pp. 405-410.

Hand gesture recognition system received great attention in the last few years because of its manifoldness applications and the ability to interact with machines efficiently through human computer interaction. In this paper a survey of recent hand gesture recognition systems is presented. Key issues of hand gesture recognition systems are presented with challenges of gesture systems. Review methods of recent postures and gestures recognition system presented as well. Summary of research results of hand gesture methods, databases, and comparison between main gesture recognition phases are also given. Advantages and drawbacks of the discussed systems are explained finally.

20) A systematic review on hand gesture recognition techniques, challenges and applications Mais Yasen and Shaidah Jusoh Department of Computer Science, Princess Sumaya University for Technology, Amman, Jordan ,Published 16 September 2019 Corresponding author Mais Yasen, mai20130045@std.psut.edu.jo

With the development of today's technology, and as humans tend to naturally use hand gestures in their communication process to clarify their intentions, hand gesture recognition is considered to be an important part of Human Computer Interaction (HCI), which gives computers the ability of capturing and interpreting hand gestures, and executing commands afterwards. The aim of this study is to

perform a systematic literature review for identifying the most prominent techniques, applications and challenges in hand gesture recognition.

21) Collision Avoidance Algorithm Using Deep Learning Type Artificial Intelligence for a Mobile Robot , Takeaki Takiguchi, Jae Hoon Lee, Member, IAENG, and Shingo Okamoto, Member, IAENG , Proceedings of the International MultiConference of Engineers and Computer Scientists 2018 Vol I IMECS 2018, March 14-16, 2018, Hong Kong

This paper presents a savvy route framework to create safe movement for a versatile robot. It comprises two primary modules of perceiving deterrence and settling on choice with man-made brainpower, separately. Right off the bat, an acknowledgement calculation utilising laser range locate (LRF) and robot odometer is created to distinguish objects close to the robot. The proposed acknowledgement calculation gives both position and speed data of hindrances by utilising range information aggregated for a certain time-frame. At that point, the outcome is utilised for figuring a safe moving heading of the robot to maintain a strategic distance from impact with objects.

For that, a computerised reasoning calculation of multi-layered neural networks was planned and prepared by profound learning technique with numerous informational collections of data including sets of sensor information and its answer of movement order for different circumstances.

22) P. Long, W. Liu, J. Pan, “Deep-Learned Collision Avoidance Policy for Distributed Multiagent Navigation,” IEEE Robotics and Automation Letters, Vol.2, No.2, pp. 656-663, 2017.

Rapid, low-idleness deterrent shirking that is coldhearted toward sensor commotion is basic for empowering various decentralized robots to work dependably in jumbled and dynamic conditions. While other dispersed multiagent impact evasion frameworks exist, these frameworks require online mathematical streamlining where monotonous boundary tuning and amazing detecting are vital. We present a novel start to finish system to create a receptive impact evasion strategy for effectively dispersed multiagent route. Our technique defines an operator's route procedure as a profound neural organization planning from the watched boisterous sensor estimations to the specialist's directing orders regarding development speed. We train the organization on countless edges of crash evasion information gathered by over and again running a multiagent test system with various boundary settings. We approve the scholarly profound neural organization strategy in a bunch of reproduced and genuine situations with loud estimations and exhibit that our technique can produce a hearty route procedure that is unfeeling toward defective detecting and works dependably in all circumstances.

23) L. Tai, S. Li, M. Liu, “A Deep-network Solution Towards Model-less Obstacle Avoidance,” Proceedings of IROS, 2016.

Obstacle avoidance is the center issue for mobilerobots. Its goal is to permit portable robots to investigate an unknown climate without crashing into different articles. It is the reason for different undertakings, for example reconnaissance and salvage, etc.Previous approaches principally centered around mathematical models (such as developing neighborhood cost-maps) which could be respected as low-level insight with no intellectual cycle. Recently,deep learning has made incredible discoveries in computer vision, particularly for acknowledgment and intellectual assignments. It takes advantage of the various leveled models

motivated by human brain structures. Nonetheless, profound learning, up till now,has rarely been utilized for controlling and dynamic.

24) Robot Path Planning with Avoiding Obstacles in Known Environment Using Free Segments and Turning Points Algorithm , Volume 2018 |Article ID 2163278

Presently, the path planning issue is one of the most investigated themes in self-governing advanced mechanics. That is the reason finding a protected way in a jumbled climate for a portable robot is a significant necessity for the achievement of any such versatile robot venture. In this work, a created calculation dependent on free sections and a defining moment technique for tackling the issue of robot way arranging in a static climate is introduced. The point of the defining moment approach is to scan a protected way for the portable robot, to make the robot move from a beginning situation to an objective situation without hitting obstructions. This proposed calculation handles two distinct targets which are the way security and the way length. Also, a powerful control law which is called sliding mode control is proposed to control the adjustment of a self-governing versatile robot to follow an ideal direction. At long last, recreation results show that the created approach is a decent choice to acquire the sufficient way and exhibit the productivity of the proposed control law for vigorous following of the portable robot.

25) AUTONOMOUS MOTION PLANNING USING A PREDICTIVE TEMPORAL METHOD By ERIC L. THORN, JR , UNIVERSITY OF FLORIDA 2009

This new strategy gives a serious way to deal with the issue of creating arrangement directions in unique conditions by richly associating the undertakings

of snag identification and forecast, climate planning and movement arranging. The dynamic ecological portrayal appears as a regular lattice which is stretched out into the time measurement by adding worldly layers to the lattice structure. The layers of this transient network speak to unmistakable time-ventures into what's to come. These time-steps are controlled by thinking about how the movement arranging calculation figures its discrete control orders. Hindrance movement expectation is consolidated into the fleeting matrix by assessing future places of moving hindrances and showing these evaluations in the layer of the fleeting framework related with the forecast occasions. The new movement arranging strategy at that point can utilize this prescient worldly lattice to examine potential control input arrangements to produce an ideal direction to accomplish its objective. As the calculation assesses potential control orders at different time-steps later on, it does as such by investigating the different transient layers of the new matrix structure compared to these unmistakable control times. By thinking about the assessed future movements of any hindrances, the movement arranging calculation can all the more keenly compute its control successions to dodge these articles in an effective way.

26) Reinforcement Learning Neural Network to the Problem of Autonomous Mobile Robot Obstacle Avoidance,Bing-Qiang Huang; Guang-Yi Cao; Min Guo,2005 International Conference on Machine Learning and Cybernetics

A way to deal with the issue of self-sufficient versatile robot hindrance evasion utilizing support learning neural organization is proposed in this paper. Q-learning is one sort of fortification learning technique that is like unique programming and the neural organization has an amazing capacity to store the qualities. We incorporate these two techniques with the expectation to guarantee independent robot conduct in convoluted eccentric climates. The reproduction results show that

the reenacted robot utilizing the support learning neural organization can improve its learning capacity clearly and can complete the given undertaking in an unpredictable climate.

27) Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods, A.M.Mutawa ,M.Murugappan, AyaHassouneh

Constant emotion recognition has been a functioning field of examination in the course of recent many years. This work plans to order genuinely handicapped individuals (hard of hearing, moronic, and disabled) and Autism kids' enthusiastic articulations dependent on facial tourist spots and electroencephalograph (EEG) signals utilizing a convolutional neural organization (CNN) and long transient memory (LSTM) classifiers by building up a calculation for continuous feeling acknowledgment utilizing virtual markers through an optical stream calculation that works successfully in lopsided lightning and subject head turn (up to 25°), various foundations, and different skin tones. The separation between the focal point of the subject's face and every marker position is utilized as an element for outward appearance arrangement. This separation include is factually approved utilizing a single direction investigation of change with a hugeness level of $p < 0.01$. Moreover, the fourteen signs gathered from the EEG signal peruser (EPOC+) channels are utilized as highlights for enthusiastic grouping utilizing EEG signals. At long last, the highlights are cross-approved utilizing fivefold cross-approval and given to the LSTM and CNN classifiers. We accomplished a most extreme acknowledgment pace of 99.81% utilizing CNN for feeling identification utilizing facial tourist spots. Be that as it may, the most emotion recognition rate accomplished utilizing the LSTM classifier is 87.25% for feeling recognition utilizing EEG signals.

28) Emotion Recognition from Facial Expression using Deep Learning , International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8 Issue-6S, August 2019

Facial expression recognition which is increasing more significance and need for it increments massively. Despite the fact that there are techniques to distinguish articulations utilizing AI and Artificial Knowledge strategies, this work endeavors to utilize profound learning furthermore, picture arrangement strategy to perceive articulations and characterize the articulations as per the pictures. Conversely the paper center to overview and survey different facial extraction highlights, passionate information bases, classifier calculations, etc. This paper is composed as follows. Segment 2 portrays foundation data about articulation acknowledgment, feeling acknowledgment framework and uses of feeling acknowledgment. Area 3 clarifies the Feature determination strategies and Image enhancement. Area 4 analyzes different Facial passionate information base. Segment 5 locations different classifier calculations for characterizing pictures as indicated by the articulation distinction. The paper is finished up in Section 6.

29) Group Emotion Recognition Using Machine Learning , School of Computer Science The University of Manchester United Kingdom ,Samanyou Garg Supervisor: Prof. Angelo Cangelosi

Automatic facial emotion recognition is a difficult assignment that has increased critical logical interest in the course of recent years yet the issue of feeling acknowledgement for a gathering of individuals has been less broadly examined. In any case, it is gradually picking up prevalence because of the monstrous measure of information accessible on interpersonal interaction destinations containing

pictures of gatherings of individuals partaking in different get-togethers. Gathering facial emotion recognition is a difficult issue because of hindrances like head and body present varieties, impediments, variable lighting conditions, change of entertainers, differed indoor and outside settings and picture quality. The goal of this errand is to arrange a gathering's apparent feeling as Positive, Neutral or then again Negative.

30) Using machine learning to detect events in eye-tracking data,February 2017 Behavior Research Methods 50(1),DOI: 10.3758/s13428-017-0860-3

Event detection is a challenging stage in eye movement data analysis. A major drawback of current event detection methods is that parameters have to be adjusted based on eye movement data quality. Here we show that a fully automated classification of raw gaze samples as belonging to fixations, saccades, or other oculomotor events can be achieved using a machine-learning approach. Any already manually or algorithmically detected events can be used to train a classifier to produce similar classification of other data without the need for a user to set parameters. In this study, we explore the application of random forest machine-learning technique for the detection of fixations, saccades, and post-saccadic oscillations (PSOs). In an effort to show practical utility of the proposed method to the applications that employ eye movement classification algorithms, we provide an example where the method is employed in an eye movement-driven biometric application. We conclude that machine-learning techniques lead to superior detection compared to current state-of-the-art event detection algorithms and can reach the performance of manual coding.

31) Emotion recognition using facial expressions.Tarnowski, Paweł & Kołodziej, Marcin & Majkowski, Andrzej & Rak, Remigiusz. (2017). Procedia Computer

Science. International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland

Facial expressions play an important role in recognition of emotions and are used in the process of non-verbal communication, as well as to identify people. They are very important in daily emotional communication, just next to the tone of voice.In the article there are presented the results of recognition of seven emotional states (neutral, joy, sadness, surprise, anger, fear, disgust) based on facial expressions. Since face is the most important to detect emotions they are using Microsoft Kinect for 3D face modeling.Kinect has small scanning resolution, but a relatively high rate of image registering (30 frames/s). And then classification of features were performed using k-NN classifier and MLP neural network.

32)An optimised algorithm for accurate steps counting from smart-phone accelerometry. Salvi, Dario & Velardo, Carmelo & Brynes, Jamieson & Tarassenko, L.. (2018).Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society.

Step counting from smart-phones allows a wide range of applications related to fitness and health. Estimating steps from phones' accelerometers is challenging because of the multitude of ways a smart-phone can be carried. We focus our work on the windowed peak detection algorithm, which has previously been shown to be accurate and efficient and thus suitable for mobile devices. We explore and optimise further the algorithm and its parameters making use of data collected by three volunteers holding the phone in six different positions. In order to simplify the analysis of the data, we also built a novel device for the detection of the ground truth steps. Over the collected data set, the algorithm reaches 95% average accuracy. We implemented the algorithm for the Android OS and released it as an open source project. A separate dataset was collected with the algorithm running

on the smart-phone for further validation. The validation confirms the accuracy of the algorithm in real-time conditions.

33)An accurate step detection algorithm using unconstrained smartphones.Yang, Xiaokun & Huang, Baoqi. (2015).

In this paper,they propose an accurate step detection method for users who carry on a smartphone in an unconstrained manner. In the first step, a rotation matrix from the device reference frame to the earth reference frame is constructed by fusing the signals from the accelerometer, magnetometer and gyroscope of a smartphone in real time based on a Kalman filter. Then, the acceleration in the device reference frame is transformed to the counterpart in the earth reference frame, and its vertical component is used to realize the peak detection algorithm. So by detecting the peaks they are able to accurately detect and count steps .

34)Maglogiannis, I., Vouyioukas, D. & Aggelopoulos, C. Face detection and recognition of natural human emotion using Markov random fields. *Pers Ubiquit Comput* 13, 95–101 (2009)

This paper presents an integrated system for emotion detection. In this research effort, we have taken into account the fact that emotions are most widely represented with eye and mouth expressions. The proposed system uses color images and it consists of three modules. The first module implements skin detection, using Markov random fields models for image segmentation and skin detection. A set of several colored images with human faces have been considered as the training set. A second module is responsible for eye and mouth detection and extraction. The specific module uses the HLV color space of the specified eye and mouth region. The third module detects the emotions pictured in the eyes and mouth, using edge detection and measuring the gradient of eyes' and mouth's

region figure. The paper provides results from the system application, along with proposals for further research.

35)Emotion Detection Algorithm Using Frontal Face Image Moon Hwan Kim, Young Hoon Joo, and Jin Bae Park.

An emotion detection algorithm using frontal facial image is presented in this paper. The algorithm is composed of three main stages: image processing stage and facial feature extraction stage, and emotion detection stage. In the image processing stage, the face region and facial component is extracted by using a fuzzy color filter, virtual face model, and histogram analysis method. The features for emotion detection are extracted from the facial component in the facial feature extraction stage. In the emotion detection stage, the fuzzy classifier is adopted to recognize emotion from extracted features. It is shown by experiment results that the proposed algorithm can detect emotion well.

36)Walczyk, J. J., Griffith, D. A., Yates, R., Visconte, S. R., Simoneaux, B., & Harris, L. L. (2012). *LIE Detection by Inducing Cognitive Load. Criminal Justice and Behavior*

This paper shows research on the accuracy of eyewitness testimony has expanded dramatically in recent years. Most of it concerns the issue of mistaken identification, not the problem of uncovering deceptive accounts of witnesses, which is the focus of this research. In the literature, a technique for lie detection has been proposed that induces cognitive load on liars by averting their rehearsal of deception: Time Restricted Integrity-Confirmation. The current authors tested it by instructing “witnesses" of actual crime videos to lie or tell the truth to related questions. Each of 145 adults was randomly assigned to a truth telling, an unrehearsed lying, or a rehearsed lying condition. The cognitive cues were response time, answer consistency, eye movements, and pupil dilation. Eye data

were gathered with an infrared eye tracker. Truth tellers had the quickest response times and the fewest inconsistencies. Moreover, they generally had more eye movements, suggesting low cognitive loads. Discriminant analyses classified rehearsed liars, unrehearsed liars, and truth tellers up to 69% accurately, with few false positives. Further refinement is warranted.So this paper tells us that eyes are most used for finding if the person is lying or not .

37)S. Cheemalapati, M. Gubanov, M. Del Vale and A. Pyayt, "A real-time classification algorithm for emotion detection using portable EEG," 2013 IEEE 14th International Conference on Information Reuse & Integration (IRI), San Francisco

Military personnel, airplane pilots, and bus drivers often operate in stressful conditions when something unexpected can happen and cause dangerous consequences if they do not respond properly. Additionally, stress adversely affects human decision making abilities, therefore prompt, preferably real-time detection of fear is very important. Based on previous studies for non-portable multi-electrode electroencephalography (EEG) systems the ratio of the power of the slow waves to that of the fast waves increases when a person is relaxed and decreases when s/he is scared. In this paper they test small portable EEG and develop algorithms for real time detection of the stressful condition - fear. During the experiment they compare EEG signals of subjects in relaxed state with those in stressed state while they are watching a scene from a scary movie. The ratio of the slow/fast wave powers was measured and the observed pattern was similar to one obtained using a multi-electrode system. We integrate stream-processing algorithms in the system to ensure real-time detection of any changes in mental condition and timely generate the alarm event.

38) Z. Labibah, M. Nasrun and C. Setianingsih, "Lie Detector With The Analysis Of The Change Of Diameter Pupil and The Eye Movement Use Method Gabor

Wavelet Transform and Decision Tree," 2018 IEEE International Conference on Internet of Things and Intelligence System

The lie is very detrimental to the fraudulent acts of many people who were cheated. The lies are common in the general population. To be able to reveal a lie we can detect through some limbs that unconsciously will show a different reaction when someone is lying. Among them, through organs of our eyes can detect someone is lying or not.Lie detection discussed in this final Task is the eyes, namely with the object of eye tracking and eye pupil diameter changes by using method of Wavelet Transform to Gabor Image Processing process and afterwards perform the classification to determine the answer someone is lying or not by using a Decision Tree. The existence of this lie detector is expected to be helpful for people who need to detect lies. With the final test results are accurate. This research has the precision value of 97%, 94%, and recall accuracy 95% of testing has been done.

39) Object Detection through Modified YOLO Neural Network

Volume 2020 |Article ID 8403262 | https://doi.org/10.1155/2020/8403262

Academic Editor: Rahman Ali

In the field of object detection, recently, tremendous success is achieved, but still it is a very challenging task to detect and identify objects accurately with fast speed. Human beings can detect and recognize multiple objects in images or videos with ease regardless of the object’s appearance, but for computers it is challenging to identify and distinguish between things. In this paper, a modified YOLOv1 based neural network is proposed for object detection. The new neural network model has been improved in the following ways. Firstly, modification is made to the loss function of the YOLOv1 network. The improved model replaces the margin style with proportion style. Compared to the old loss function, the new is more flexible and more reasonable in optimizing the network error. Secondly, a spatial pyramid

pooling layer is added; thirdly, an inception model with a convolution kernel of 1 1 is added, which reduced the number of weight parameters of the layers.

40) Region-based Convolutional Networks for Accurate Object Detection and Segmentation

Also proposed in 2013, R-CNN is a bit late compared with OverFeat. However, this region-based approach eventually led to a big wave of object detection research with its two-stage framework, i.e, region proposal stage, and region classification and refinement stage.

Region proposal from selective search highly depends on the similarity assumption, so it can only provide a rough estimate of location. To further improve localization accuracy, R-CNN borrowed an idea from “Deep Neural Networks for Object Detection” (aka DetectorNet), and introduced an additional bounding box regression to predict the center coordinates, width and height of a box. This regressor is widely used in the future object detectors.