KANTIPUR ENGINEERING COLLEGE

(Affiliated to Tribhuvan University)

Dhapakhel, Lalitpur

[Subject Code: EX707]

A MAJOR PROJECT MID-TERM REPORT ON

AUTOMATED DRIVING LICENSE TEST

Submitted by:

Nitesh Poudel	**[KAN075BEI03]**
Pawan Bhandari	**[KAN075BEI04]**
Prajjwol Shrestha	**[KAN075BEI05]**
Ujjwal Naral	**[KAN075BEI09]**

A MAJOR PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR IN ELECTRONICS, COMMUNICATION AND INFORMATION TECHNOLOGY ENGINEERING

Submitted to:

Department of Computer and Electronics Engineering

December, 2022

AUTOMATED DRIVING LICENSE TEST

Submitted by:

Nitesh Poudel [KAN075BEI03]

Pawan Bhandari [KAN075BEI04]

Prajjwol Shrestha [KAN075BEI05]

Ujjwal Naral [KAN075BEI09]

A MAJOR PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR IN ELECTRONICS, COMMUNICATION AND INFORMATION TECHNOLOGY ENGINEERING

Submitted to:

Department of Computer and Electronics Engineering

Kantipur Engineering College

Dhapakhel, Lalitpur

December, 2022

ABSTRACT

In the traditional method of administering driving license exams, a human evaluator observes the candidate's performance and manually assesses whether they have passed or failed. This process can be time-consuming and prone to human error.

The "Automated Driving License Test" project aims to address these issues by replacing manual supervision with a software solution. The software utilizes various Python libraries to capture frames of the exam and processes these frames to determine the candidate's performance. This allows for a more efficient and accurate evaluation of the candidate's driving skills.

One of the key benefits of this project is the increased efficiency it brings to the testing process. With the automated system, exams can be conducted more quickly and with less human involvement, freeing up time for other tasks. In addition, the use of a software system for evaluation reduces the potential for human error, providing a more objective assessment of the candidate's performance.

Overall, the "Automated Driving License Test" project has the potential to revolutionize the way driving license exams are conducted. By replacing manual supervision with a software solution, this project brings increased efficiency, accuracy, and objectivity to the testing process. It could be implemented in driving schools and licensing agencies, helping to ensure that only safe and competent drivers are granted a license to operate a motor vehicle.

Keywords *– Driving test automation,object detection, image processing*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Background

Obtaining a driver's license is a crucial milestone for many individuals, as it allows them to operate a motor vehicle and provides a sense of independence and freedom. In order to obtain a driver's license, individuals must pass a driving exam that tests their knowledge of traffic laws and their ability to safely operate a vehicle. According to the Department of Transport Management (DoTM) the no of applications to obtain driving license were over 25000 per week. A report from NADA shows

Traditionally, driving license exams have been administered by a human evaluator who observes the candidate's performance and assesses whether they have passed or failed. However, this process can be time-consuming and subject to human error. There is also the potential for bias or inconsistency in the evaluation process, as different evaluators may have different standards for what constitutes a pass or fail. Here are some problems with the traditional model of administering driving license tests:

- **Time-consuming**: The process of evaluating a candidate's performance during a driving license test can be time-consuming, as the evaluator must observe the candidate's performance for the entire duration of the exam.
- **Prone to human error**: Human evaluators are subject to error, and there is the potential for mistakes or inconsistencies in the evaluation process.
- **Potential for bias**: Different evaluators may have different standards for what constitutes a pass or fail, leading to the potential for bias in the evaluation process.
- **Inefficient**: The traditional model of administering driving license tests can be inefficient, as it requires the presence of a human evaluator and may involve multiple test sessions if the candidate does not pass on the first attempt.
- **Subjective**: The traditional model relies on subjective evaluations by human evaluators, which may not provide a consistent or objective assessment of the candidate's performance.

In an effort to address these issues and bring increased efficiency, accuracy, and objectivity to the process of administering driving license exams, the ”Automated Driving License Test” project was developed. This project replaces manual supervision with a software solution that utilizes various Python libraries to capture and process frames of the exam to determine the candidate’s performance.

1.2 Problem Definition

The traditional method of administering driving license exams involves manual supervision by a human evaluator, which can be time-consuming, prone to human error, and subject to bias. In addition, the subjective nature of the evaluation process may lead to inconsistency in the assessment of candidates’ performance. As a result, there is a need for a more efficient, accurate, and objective solution for administering driving license exams. The ”Automated Driving License Test” project aims to address these issues by replacing manual supervision with a software solution that utilizes various Python libraries to capture and process frames of the exam to determine the candidate’s performance. By automating the evaluation process, the project aims to increase efficiency, accuracy, and objectivity in the testing process, while reducing the potential for human error and bias.

1.3 Objectives

The primary objectives of this projects is as follows:

1. To develop a software solution for administering driving license exams that is more efficient, accurate, and objective than the traditional method of manual

The specific objectives of this project are as follow:

1. To utilize various Python libraries to capture and process frames of the exam to determine the candidate’s performance.
2. To reduce the potential for human error and bias in the evaluation process by automating the assessment of candidates’ performance.
3. To increase the efficiency and accuracy of the testing process by automating the evaluation process.

4. To provide a consistent and objective assessment of candidates' performance by relying on software-based evaluation rather than subjective human evaluations.
5. To potentially improve road safety by ensuring that only competent and qualified drivers are granted a license to operate a motor vehicle.

1.4 Project Features

The project will be able to accomplish following:

- **Automated evaluation**: The software solution utilizes various Python libraries to capture and process frames of the exam to determine the candidate's performance, allowing for an automated evaluation of the candidate's driving skills.
- **Increased efficiency**: The automation of the evaluation process allows for a more efficient testing process, as exams can be conducted more quickly and with less human involvement.
- **Improved accuracy**: The use of a software system for evaluation reduces the potential for human error, providing a more accurate assessment of the candidate's performance..
- **Consistent and objective assessment**: The software-based evaluation system provides a consistent and objective assessment of candidates' performance, rather than relying on subjective human evaluations.

1.5 Project Application

- **Driving schools**: The software could be used to evaluate the performance of students during their driving lessons and provide feedback on areas where they need to improve. This could help to ensure that students are fully prepared to pass the driving license exam and become safe, competent drivers.
- **Licensing agencies**: The software could be used to conduct driving license exams for individuals seeking to obtain or renew their license. The automation of the evaluation process would increase the efficiency and accuracy of the testing process, while also reducing the potential for human error and bias.

1.6 System Requirement

1.6.1 Software Requirements

The software requirements are as follows:

1. Windows/Linux/Mac
2. Rasbian OS
3. Arduino IDE

1.6.2 Hardware Requirements

The hardware requirements are as follows:

1. Raspberry Pi
2. Pi camera/USB camera
3. Arduino Nano
4. DC motor
5. Motor mount
6. Wheel
7. Motor Driver
8. Bluetooth module
9. Wire

1.7 Project Feasibility

The feasibility analysis of the system has been done from various aspects such as technical, operational and economical viewpoint.

The present technology is sufficient to meet the requirements of the system, the required algorithm exists and the device to input the data to the system is also present. The system is believed to work well when developed and installed. The requirements for the project have been accounted for and the system is built on available resources to meet the requirements. The detailed feasibility study is as follows

1.7.1 Technical Feasibility

Our project satisfies technical feasibility needs. The existing network protocols and operating system services of various operating systems would allow for feasible implementation of this application. As this service satisfies technological hardware and software capabilities of present day available personal devices, the proposed project was decided to be technically feasible.

1.7.2 Operational Feasibility

The operation of the system requires only a modern computer, the user interface will be simple and intuitive. The solution proposed for the project is operationally workable and user-friendly to end users.

1.7.3 Economic Feasibility

Economic feasibility analysis is the most commonly used method for determining the efficiency of a project. It is also known as cost analysis. It helps in identifying profit against investment expected from a project. Cost and time are the most essential factors involved in this field of study. Developed system is economically feasible. It can be developed on a simple PC which can be available at an affordable cost. System is built on open-source language, so development is free of cost. All the references and resources are freely available on the internet. So, we can say that the developed system is economically feasible.

CHAPTER 2
LITERATURE REVIEW

2.1 Previous Works

In the last few years with the rise in affordable compute power and abundance of data various machine learning algorithms are being used to solve problems in everyday life which previously were thought to be immensely difficult in the field of computer science.

Computer vision is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. Computer vision tasks include methods for acquiring, processing, analyzing and understanding digital images, and extraction of high-dimensional data from the real world in order to produce numerical or symbolic information.

The task of automating driving license test can be accomplished in two different ways. A hardware based driving license test by interfacing Arduino UNO board with number of sensors can be implement at low cost. This automated system is done by interfacing Arduino UNO board with number of sensors, these sensors are kept on the test track to identify the errors of the candidate while he is taking the test. Arduino UNO microcontroller reads the data collected by sensors, processes them and send the result information to the candidate and authorities of Regional Transport Office (RTO) whether he/she is passed or failed to get a driving license[1]. Main drawback of using a hardware based automation system is the need to align sensor every time a tester fails the test by colliding with the sensor.

Indoor Positioning System can be used to track mobile devices indoor. IPS works indoor giving accurate position where satellite positioning system such as GPS, Galielo, Glonass fail to provide accurate location of the subject in question.[2]. Indoor Positioning System technology can also be used for outdoor positioning, although the same technology used inside and outside may give different results[3]. An IPS consists of transmitters located on the observed subject and a wireless signal reader that detects and reads the received signals. The location of transmitters and receivers can be swapped, which depends on the type of application[2].Various wireless technolo-

gies are used to communicate between transmitters and readers. The most popular are Ultra-Wideband (UWB), Wireless Fidelity (Wi-Fi), Bluetooth, light, and active/passive RFID systems[4]. Visible Light Communication (VLC) is gaining increasing attention, since its high data rate, security, no interference with Radio Frequency (RF) spectrum and high resolution [5] make it an appropriate choice for localization application. Apart from their support in VLC-based or infrared-based positioning, vision techniques are also used in camera-based IPS without any supporting light framework apart from those common to any modern building[4]. Even though Simultaneous Localization and Mapping (SLAM) is mainly based on cameras for sensory input presently, one should not enclose SLAM into vision-based methods. SLAM is possible using, for example, laser scanners, sonars or odometric data provided by wheel encoders [6].
A software based system that uses the frames captured by a simple camera can be implement to indentify the position of vehicle. The position of vehicle can be then compared with the position of markings to determine whether the examinee has made any mistake or not. The software uses simple image processing library such as YOLO and image processing methods to identify the vehicle and track. Chen, Weina et al.[2017] [7] gave a lane marking detection algorithm both geometry and modified Hough Transform. The algorithm divided the image captured into road part and non road part by using camera geometry information. The problems of hardware implementation can be solved by implementing a software based testing system. Pole mounted camera along the entire length of test track can be deployed to determine the position of vehicle. A smartphone mounted on the front windscreen can be deployed to use its camera and sensors to determine the position of vehicle. A monocular vision system that could locate the positions of the road lane in real time was proposed for lane detection using single camera. The algorithm worked by initially using edge detection to find all present edges from road image as road line required was included in it and canny approach was used to achieve the edge map from the road image for its accurate edge detection. [8] HAMS, in its general incarnation, uses the smartphone's front and rear cameras, and other sensors, to monitor the driver (for instance, their gaze) and the road scene in front (for instance, the distance to the vehicle in front), simultaneously. It employs advanced Artificial Intelligence (AI) models, which the team has developed for efficient and robust. operation[9].

2.2 Related Theory

2.2.1 Artificial Neural Network(ANN)

Artificial neural networks (ANN) are computing systems vaguely inspired by the biological neural networks that constitute animal brains. Such systems ”learn” to perform tasks by considering examples, generally without being programmed with task-specific rules[9].

Neurons

ANNs are composed of artificial neurons which retain the biological concept of neurons, which receive input, combine the input with their internal state (activation) and an optional threshold using an activation function, and produce output using an output function. The important characteristic of the activation function is that it provides a smooth transition as input values change, i.e., a small change in input produces a small change in output.[4]

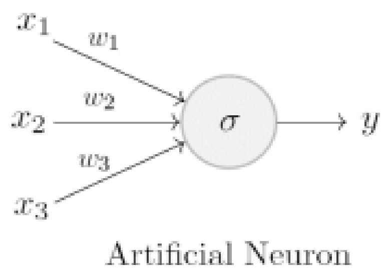


Figure 2.1: Artificial Neuron

The network consists of connections, each connection providing the output of one neuron as an input to another neuron. Each connection is assigned a weight that represents its relative importance. A given neuron can have multiple input and output connections. The neurons are typically organized into multiple layers, especially in deep learning. Neurons of one layer connect only to neurons of the immediately preceding and immediately following layers. The layer that receives external data is the input layer. The layer that produces the ultimate result is the output layer. In between them are zero or more hidden layers.

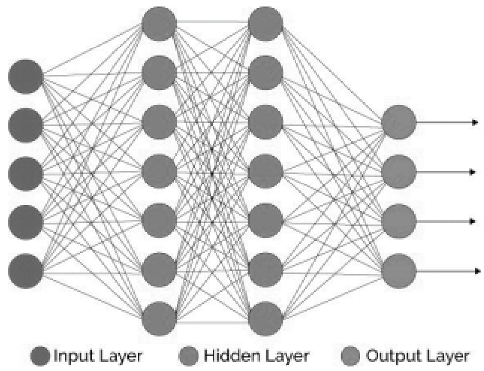


Figure 2.2: Structure of Neural Network

2.2.2 Convolutional Neural Network(CNN)

CNN is a kind of neural network, its weight sharing network structure makes it more similar to the biological neural network, reduces the complexity of the network model and the number of weights [5].CNNs were inspired by biological processes. The name "convolutional neural network" indicates that the network employs a mathematical operation called convolution. Convolution is a specialized kind of linear operation. They use convolution in place of general matrix multiplication in at least one of their layers.

Design

A convolutional neural network consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of a series of convolutional layers that convolve with a multiplication or other dot product.

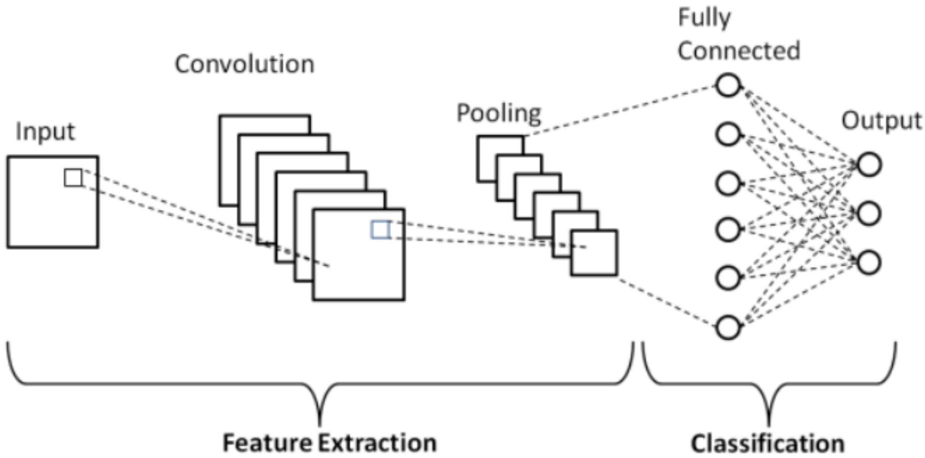


Figure 2.3: Basic schematic of Convolutional Neural Network(CNN)

A CNN typically has three layers:

- Convolutional layer:This layer performs a dot product between two matrices, where one matrix is the set of learnable parameters otherwise known as a kernel, and the other matrix is the restricted portion of the receptive field. The kernel is spatially smaller than an image but is more in-depth. This means that, if the image is composed of three (RGB) channels, the kernel height and width will be spatially small, but the depth extends up to all three channels.During the forward pass, the kernel slides across the height and width of the image-producing the image representation of that receptive region. This produces a two-dimensional representation of the image known as an activation map that gives the response of the kernel at each spatial position of the image. The sliding size of the kernel is called a stride.

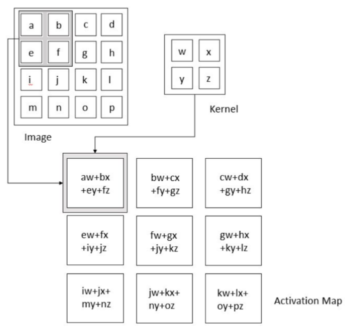


Figure 2.4: Activation Map

- Pooling layer: The pooling layer replaces the output of the network at certain locations by deriving a summary statistic of the nearby outputs. This helps in reducing the spatial size of the representation, which decreases the required amount of computation and weights. The pooling operation is processed on every slice of the representation individually.

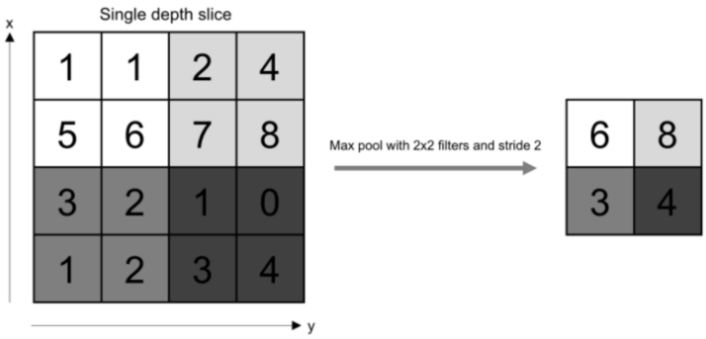


Figure 2.5: Pooling Operation

- Fully connected layer: Neurons in this layer have full connectivity with all neurons in the preceding and succeeding layer as seen in regular Forward Convolutional Neural Network(FCNN). This is why it can be computed as usual by a matrix multiplication followed by a bias effect.

2.2.3 Hough Transform

The Hough transform (HT) was firstly proposed in for machine analysis of bubble chamber photographs. It parametrizes straight lines with slope-offset, leading to an unbounded transform space (since the slope can be infinity). extended HT by using angle-radius rather than slope-offset parameters, and is conceptually similar to two-dimensional Radom transform. Then generalized the idea of HT to localize arbitrary shapes, e.g. ellipses and circles, from digital images. For example, by parameterizing with angle and radius, line detection can be performed by voting edge evidence and finding peak response in the finite parametric space. Typically, with the edge detectors such as Canny and Sobel , the detected lines are the maximal local response points in the transformed parametric space. The core idea of HT is used in two recent works which parameterize the outputs of CNNs with offsets and orientations to predict surface meshes or convex decomposition of 3D shapes. Despite the success of HT on line detection, it suffers from high computational costs and unstable performance. To accelerate the voting of HT proposed the "probabilistic Hough transform" to randomly pick sample points from a line, while using the gradient direction of images to decide the voting points. Meanwhile, the work of employed kernel-based Hough transform to perform hough voting by using the elliptical-Gaussian kernel on collinear pixels to boost the original HT. Besides partitioned the input image into hierarchical image patches, and then applied HT independently to these patches. Use a coarse-to-fine accumulation and search strategy to identify significant peaks in the Hough parametric spaces. tackled line detection within a regularized framework, to suppress the effect of noise and clutter corresponding to nonlinear image features. The Hough voting scheme is also used in many other tasks such as detecting centroid of 3D shapes in point cloud and finding image correspondence.

CHAPTER 3
METHODOLOGY

Automatic driving license consists of two parts. Object detection and lane detection. Object detection will be used to determine the position of vehicle and the position obtained will be compared with lane. If the position of vehicle is determined to be touching the lane then test will be automatically stopped prompting a message.

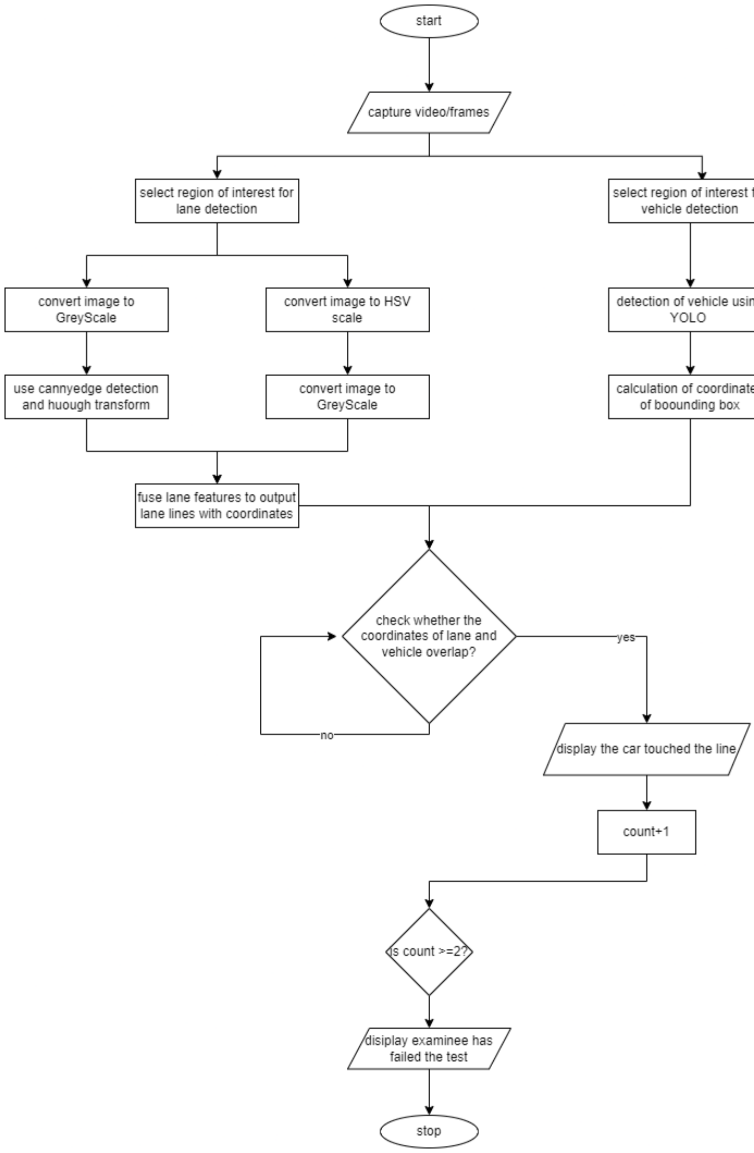


Figure 3.1: Flow chart of System

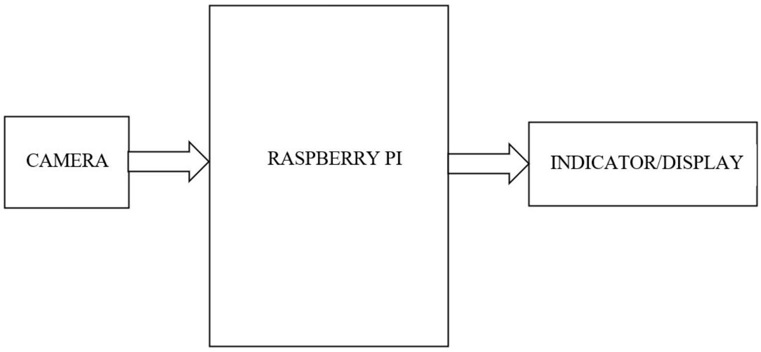


Figure 3.2: Block Diagram of System

3.1 Object Detection

The Object Detection part focuses on detecting the car body precisely using You Only Look Once(YOLO) algorithm.

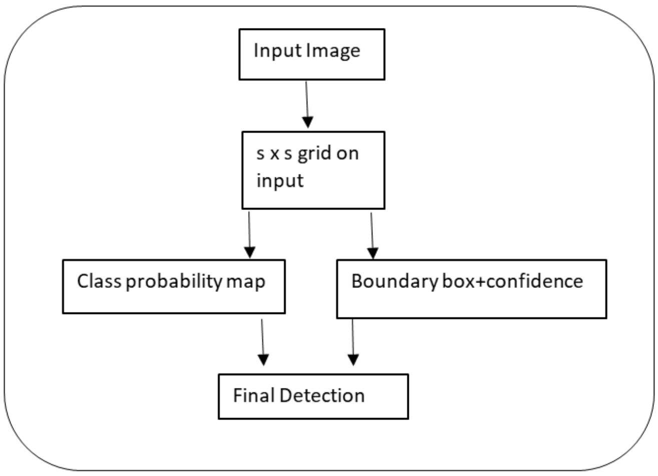


Figure 3.3: Object detection using YOLO

You Only Look Once(YOLO) algorithm works using the following three techniques:

- Residual blocks:First, the image is divided into various grids. Each grid has a dimension of S x S.Every grid cell will detect objects that appear within them.
- Bounding box regression: A bounding box is an outline that highlights an object

in an image.Every bounding box in the image consists of the attributes, width (bw), height (bh), class, bounding box center (bx,by).

- Intersection Over Union (IOU):Intersection over union (IOU) is a phenomenon in object detection that describes how boxes overlap. YOLO uses IOU to provide an output box that surrounds the objects perfectly.

3.2 Lane Detection

Hough transform is a feature extraction method used in image analysis. The purpose of the technique is to find imperfect instances of objects within a certain class of shapes by a voting procedure. This voting procedure is carried out in a parameter space, from which object candidates are obtained as local maxima in a so-called accumulator space that is explicitly constructed by the algorithm for computing the Hough transform. Hough transform can be used to isolate features of any regular curve like lines, circles, ellipses, etc. Hough transform in its simplest from can be used to detect straight lines in an image. A generalized Hough transform can be used in applications where simple analytic description of features is not possible.

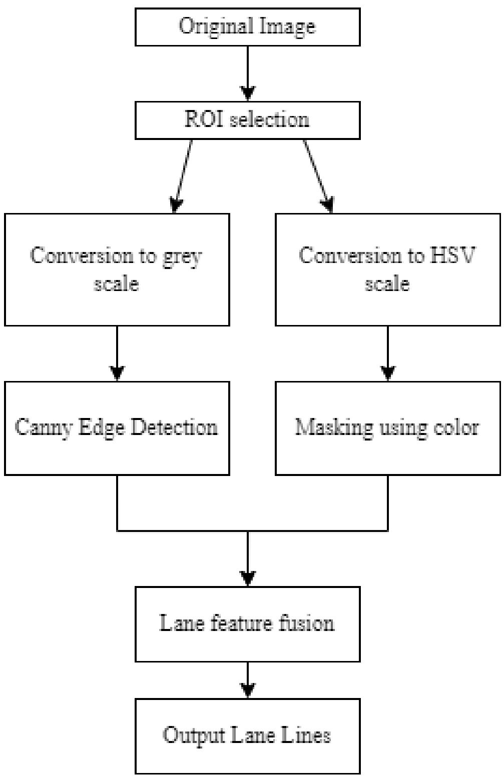


Figure 3.4: Lane Detection using feature selection

3.2.1 Testing

Effectiveness of the software will be tested by building a scaled downed model. The scaled downed car is in ratio 0f 1:17 with real car.

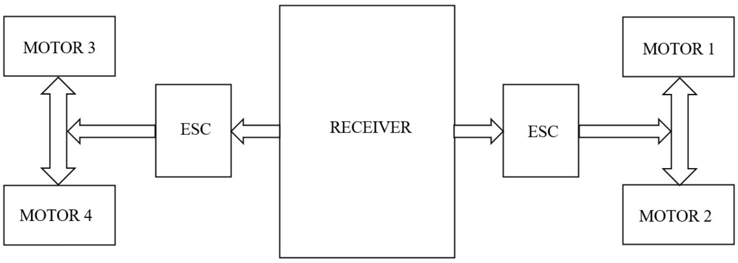


Figure 3.5: Block Diagram of Car

Four motor are used in the scaled down model of the car. The motors are controlled using Electronics Speed Controller(ESC) and control signal is provided to the Electronic Speed Controller through a 2.4 Ghz receiver. A 2.4 Ghz transmitter is used by controller to transmit control signal wirelessly.

CHAPTER 4
EPILOGUE

4.1 Result

4.2 Conclusion

4.3 Future Implementation

4.4 Work Completed

As part of the ”Automated Driving License Test” project, we have made significant progress in developing a software solution for administering driving license exams. The following tasks and milestones have been completed:

1. Scaled downed car: A scaled down version of test car with four wheels is constructed to test the software’s performance.
2. Custom object detection: Software if able to identify scaled down test car with accuracy of 0.9.
3. Lane Detection: An algorithm for the detection of road lines is developed and is in the process of refinement to increase it’s accuracy and reliability.
4. Pass or Fail determination: CNN is able to identify if the car has touched the line or not and is able to determine whether the examinee has passed or failed the exam.

4.5 Work in Process

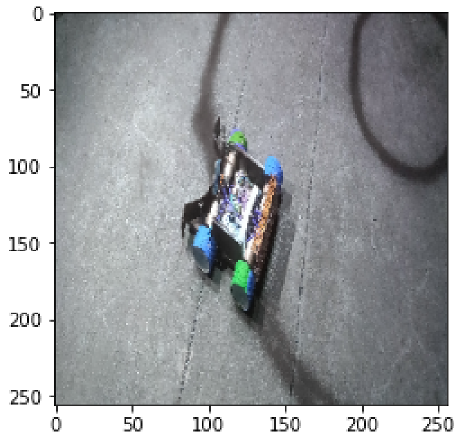
1. Lane detection is on process of refinement.
2. Data model training.
3. Scaled downed Track.
4. Linking all algorithms into one program.

4.6 Work Remaining

1. Scaled downed Track.
2. Linkning all algorithms into one program

4.7 Result

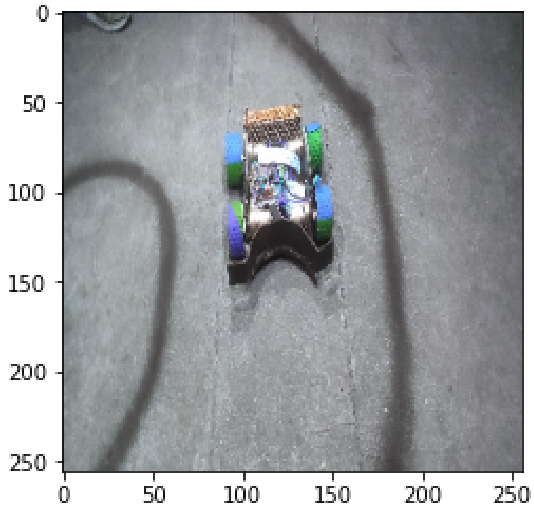
1. The system predicts that the examinee has failed their exam when the robot touches the line



```
Predicted class is Fail
```

Figure 4.1: example of result prediction

2. The system predicts that the examinee has passed their exam when the robot does not touch the line



Predicted class is Pass

Figure 4.2: example of result prediction

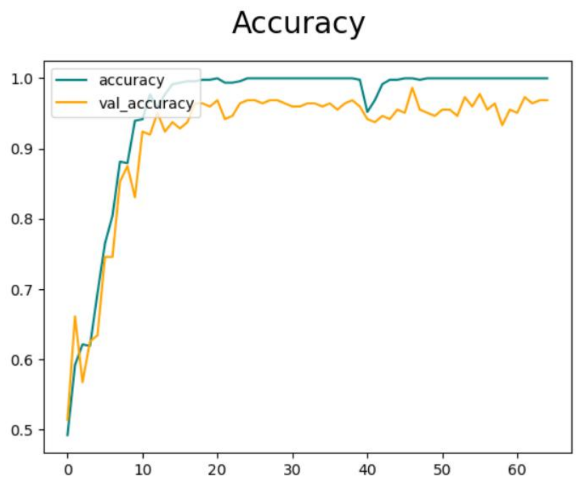
3. 

Figure 4.3: Accuracy

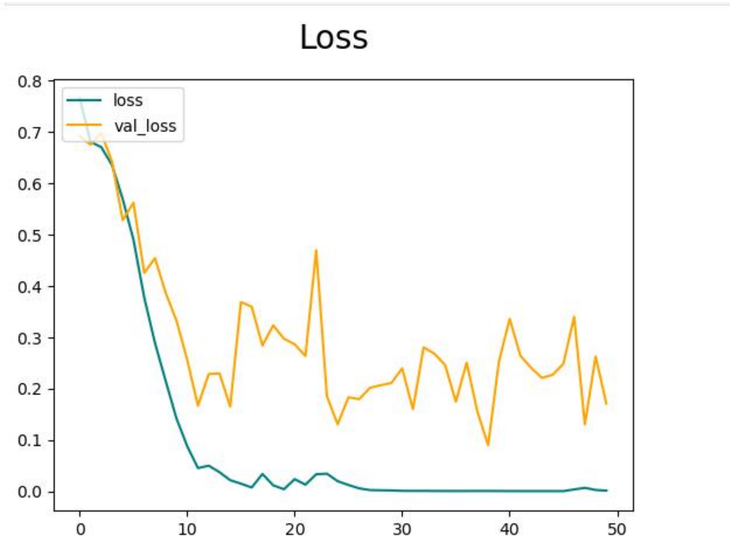
4. 

Figure 4.4: Loss

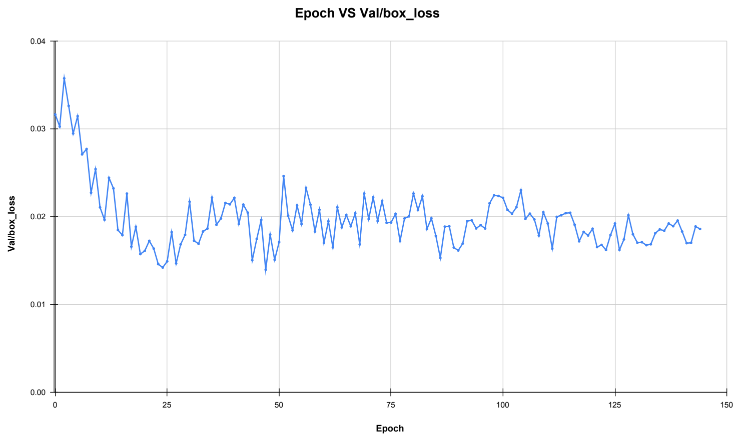
5. 

Figure 4.5: Epoch VS Val box loss for YOLO

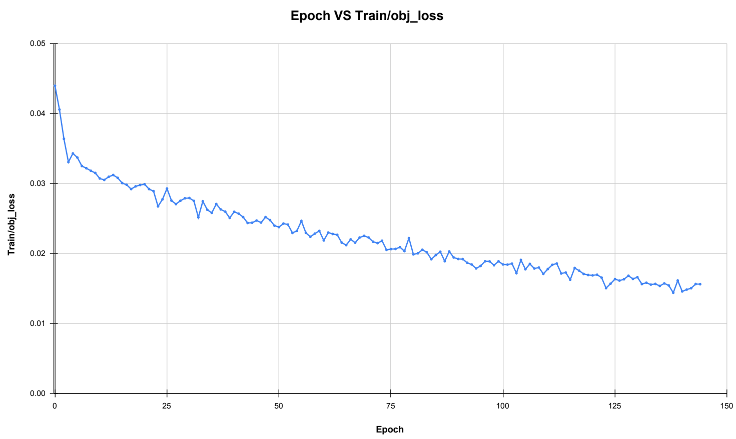
6. 

Figure 4.6: Epoch VS Train obj loss for YOLO

4.8 Work Scheduled

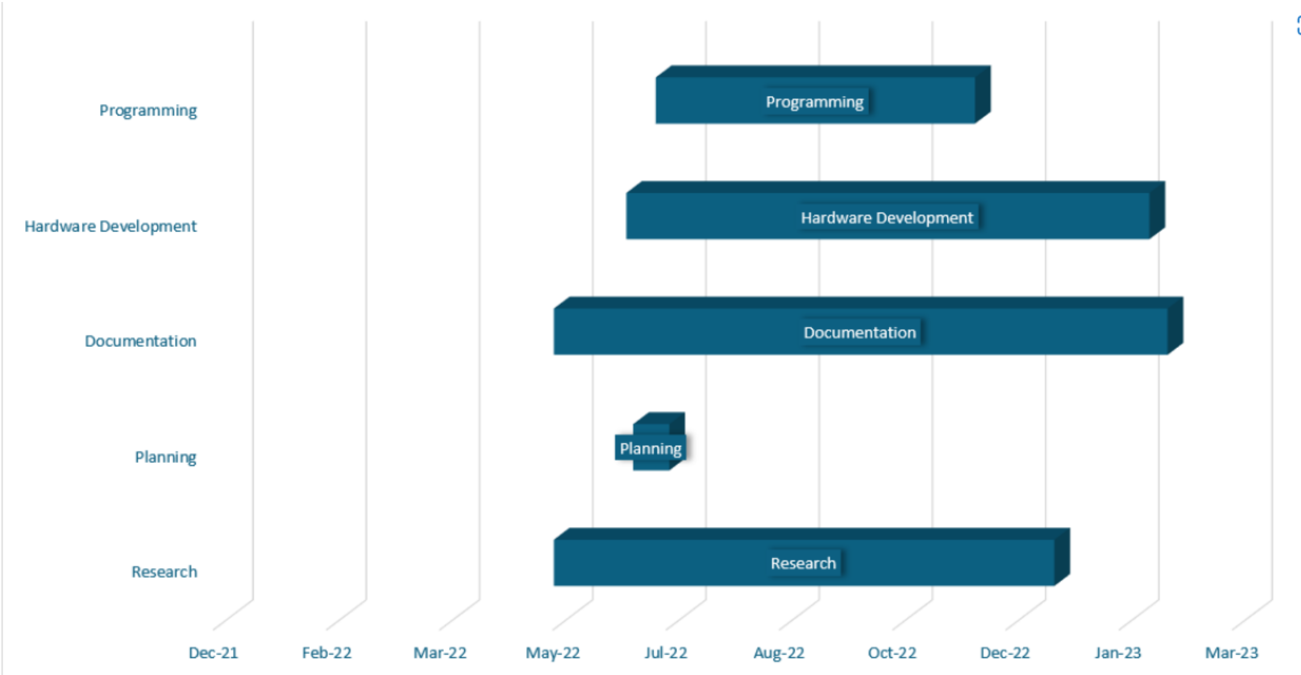
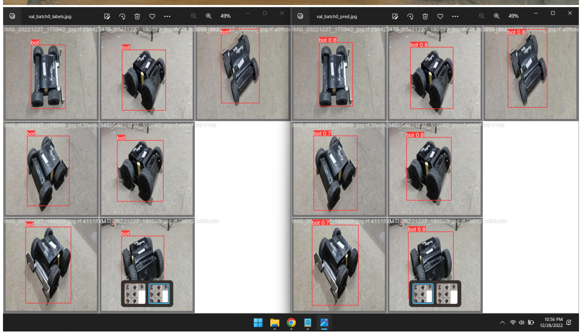


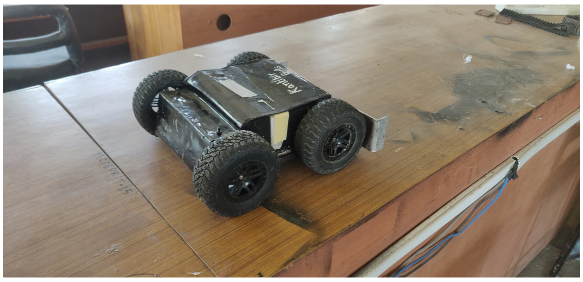
Figure 4.7: Gantt Chart

4.9 Cost Estimation

Items	No. of Item	Unit Price(Rs.)	Total Price(Rs.)
Raspberry Pi 4	2	4,500/-	9,000/-
200 rpm metal gear dc motor	4	1000/-	4000/-
Arduino Nano	1	800/-	800/-
Motor Driver Module	2	1000/-	2000/-
RTC module	1	200/-	200/-
Pi camera/ USB camera	2	1500/-	3000/-
Bluetooth Module	1	800/-	800/-
Total			21,000/-

Table 4.1: Cost Estimation



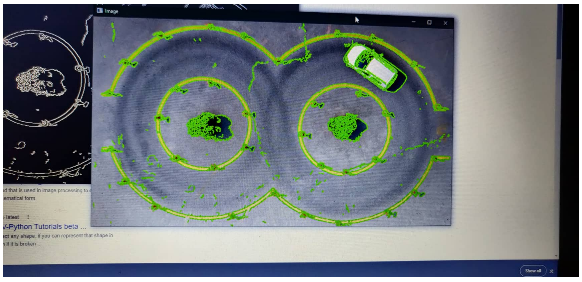


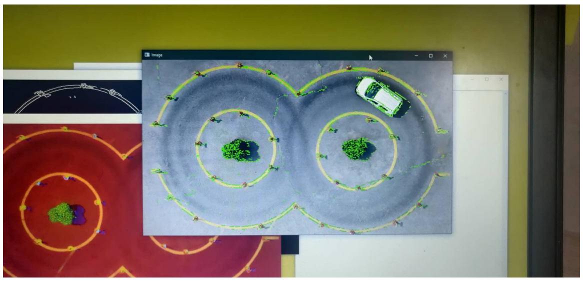
%begincenter

Snapshots

Work Done

Work in Process





BIBLIOGRAPHY

[1] A. Chaudhari and C. S. R. Murthy, "Femto-to-femto (f2f) communication: The next evolution step in 5g wireless backhauling," in *2017 15th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pp. 1–8, May 2017.

[2] P. Bencak, D. Hercog, and T. Lerher, "Indoor positioning system based on bluetooth low energy technology and a nature-inspired optimization algorithm," *Electronics*, vol. 11, no. 3, 2022.

[3] M. A. Fadzilla, A. Harun, and A. B. Shahriman, "Localization assessment for asset tracking deployment by comparing an indoor localization system with a possible outdoor localization system," in *2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA)*, pp. 1–6, 2018.

[4] G. M. Mendoza-Silva, J. Torres-Sospedra, and J. Huerta, "A meta-review of indoor positioning systems," *Sensors*, vol. 19, no. 20, 2019.

[5] L. U. Khan, "Visible light communication: Applications, architecture, standardization and research challenges," *Digital Communications and Networks*, vol. 3, no. 2, pp. 78–88, 2017.

[6] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1309–1332, 2016.

[7] W. Chen, Q. Zeng, J. Liu, and H. Wang, "Seamless autonomous navigation based on the motion constraint of the mobile robot," *Industrial Robot: An International Journal*, vol. 44, pp. 178–188, 03 2017.

[8] R. Liu and J. Mao, "Research on improved canny edge detection algorithm," *MATEC Web of Conferences*, vol. 232, p. 03053, 01 2018.

[9] B. Bordia, N. Nishanth, S. Patel, M. Anand Kumar, and B. Rudra, “Automated traffic light signal violation detection system using convolutional neural network,” in *Soft Computing: Theories and Applications*, pp. 579–592, Springer, 2020.