

Machine Learning Approach to Detect Drowsiness on Behavioral Parameters

Mr. Aannd R^{1*}

Assistant Professor,

¹BMS Institute of Technology and Management, Yelahanka, Bangalore 560064, India.

Dr. Anil G N²

Professor & Dean Academics

²BMS Institute of Technology and Management, Yelahanka, Bangalore 560064, India.

Rishika Sankaran³, Anushruti Adhikari⁴, Kruthika Ravishankar⁵

UG Students, ^{3,4,5}BMS Institute of Technology and Management, Yelahanka, Bangalore 560064, India.

Abstract

Object detection has received a lot of research attention in recent years because of its tight association with video analysis and picture interpretation. Face detection, vehicle detection, pedestrian counting, web photos, security systems, and self-driving automobiles are all examples of object detection.

With little conscious thought, the human visual system can accomplish complicated tasks such as distinguishing multiple objects and detecting impediments. Thanks to the availability of large amounts of data, faster GPUs, and improved algorithms, we can now quickly train computers to detect and classify many elements inside a picture with high accuracy.

Our project is focused on building a single-access platform for various object detection tasks. A user-interface where the user is asked for the relevant inputs and an output based on this is generated automatically by the system. Also, accuracy and precision measures are also displayed so that the user is wary of their liability extent on the generated results.

Keywords- Face Detection, GPU, Self-Driving.

1. Introduction

Computer vision techniques for finding and categorizing items in the frame of a video sequence are referred to as object detection. Detecting and tracking moving objects or targets in real-time video is a critical and difficult operation. Despite the fact that several approaches have been used in the detecting mechanism, the level of accuracy is still not up to par.

As a result, it was the birth of the neural network method for recognizing objects in a video sequence. A deep neural network is one of them, and it elaborates the hidden layers in order to considerably raise the level of accuracy in detecting the item in a video. One of them is R-

CNN, which is built on a deep convolutional neural network and was first used for identification in 2014. Following that, newer methods such as Spp-net, fast R-CNN, faster RCNN, and R-FCN entered the field. It cannot be used to identify many real-time objects in a single frame due to their complicated network structure.

Many scenarios in today's deep learning era necessitate the completion of time-consuming manual operations. We strive to simplify these activities and turn manual involvement into supervision, ranging from large-scale stock and price predictions to the generation of large-scale datasets for use in enterprise-level projects, which include millions of attributes and instances. As a result, once trained with particular unbiased datasets and algorithms, the machine handles the detection and prediction jobs. Our aim in this project is to have an online platform that can automate these chores. Also, our goal is to integrate as many diverse tasks as possible, to provide a single platform that can handle versatile interests.

1.1 Background

COVID-19 pandemic have not only been destroying the lives of people, but have also succeeded in making a negative impact on the economy of our country. A health emergency was declared on 2020 for the entire globe by the World Health Organization (WHO). More than 3 million people by April 2020, were infected by this virus. By that time there was no antidote or vaccine for this outbreak. Certain safety rules and measures were issued by the WHO in the interest of common people. The application of information technology, particularly in fields like data science and machine learning, can aid in the fight against the epidemic. It becomes critically important to develop some early warning systems that can predict how much a sickness would harm society. Machine learning algorithms and methods are used for predicting the chance of being infected by this virus and detecting whether a person is already COVID affected or not. A trial was conducted using different algorithms, and the algorithm that gave results with the best accuracy were included in the hub.

Machine learning algorithms are used to detect drowsiness based on behavioral parameters. Faces reveal information that can be used to estimate tiredness. To evaluate the level of weariness, many facial traits can be inferred from the face. These include eye blinks, head movements, and yawning. However, developing a trustworthy and accurate sleepiness detection system is still a work in progress, and obtaining correct results is a difficult process that necessitates precise and robust algorithms. In the past, a variety of strategies for detecting driver drowsiness were investigated. The new advent of deep learning necessitates a re-evaluation of these algorithms' accuracy in detecting tiredness. The study found that while the support vector machine technique is the most widely used method for detecting drowsiness, convolutional neural networks outperformed the other two methods.

Facebook has been working and made a successful attempt in developing a system that has an incredible capacity to detect people in images with greater accuracy. In the beginning we know that if we uploaded a certain picture with someone we needed to tag them individually but now Facebook has upgraded its system and itself recognizes the person in the image deleting the

need to tag them individually. Facebook gives an accuracy of 98 percent, which is about as good as humans can get for face detection. Surveillance cameras are used in public places to capture footage and provide security. The two main advantages of this strategy over others are its uniqueness and acceptability. We want speed and precision when it comes to recognition. Detecting faces whereas is a set of issues: First, take a look at a photograph and try to identify everyone in it. Second, focus on each face, remembering that a face is still a face even if it is turned. Third, the system chooses features that can be utilized to distinguish each face, for example such as the size of the eyes, shape of the face, and many more. Finally, compare these characteristics to the data we have in order to identify the person.

1.2. Literature Survey

A background subtraction approach that works well with a method for properly detecting objects is proposed. The LIBS method (Local Illumination Based Background Subtraction) was being used. Each pixel is made up of three values to describe the backdrop: maximum, minimum, and maximum intensity variation between successive frames. The locations of these components are verified and refined using dynamic template matching. [1]

- Real-time object detection is amongst the most challenging tasks since it need quicker processing power to recognize the thing at around that precise instant. Unfortunately, most data generated only by real-time system is unlabeled, and successful training requires a large collection of tagged data. Based on a convolution neural network model, this paper offers Single Shot Multi-Box Detection (SSD), a faster detection approach in real-time item detection. As a result, in this proposed study, a low weight network model known as Mobile Net is used, which utilizes depth-wise separable convolution. The experimental results demonstrate that combining Mobile Net with the SSD model enhances the accurateness in recognizing real-time household objects. [2]

- Object recognition in a road scene has piqued the interest of academics working on self-driving cars and automatic traffic monitoring systems. On the other hand, is a regular occurrence in ordinary road scenes in Object occlusion. Due to occlusion difficulties, previous object recognition systems were restricted in their capacity to detect items accurately. In this paper, we offer an occlusion-resistant object identification network. [3]

- Future Smart Vehicle Systems (SVS) will include drowsiness detection, which will take into account yawning, drivers eye closure time, eye blinking rate, eye brow shape gestures other than break and accelerator pattern, steering movement, and continuous driving duration, vehicle speed among other things. To detect sleepiness and weariness, a variety of approaches are used. Other criteria for this may include medical parameters such as heart rate, pulse rate, and so on.[4]

- COVID-19 widespread testing and early detection are crucial in controlling the present global pandemic. The three most essential variables in any disease detection technology, notably COVID-19, are time, cost, and accuracy. To address these issues, this work offers a CNN-based method for detecting COVID-19 occurrences in patients' chest X-rays. The system is built using 330 chest X-ray pictures split into two groups: "COVID-19" as well as "Normal." Similarly, for model validation, an evenly distributed image set of 82 chest X-rays is employed. [5]

- There are two sorts of existing approaches for identifying and warning drivers who are sleepy: image processing, and 2) signal processing Which approaches are accessible for different data implementations, such as pictures from a normal camera or images from such an infrared camera, as well as the usage of bio-signals? In the first example, they utilise eyelid closure detection and yawning detection to assess if a driver is asleep. They begin by detecting the driver's eyes and lips, then utilise a camera in front of the car to compute the percentage of drowsiness. [6]

1.3. Motivation

This pandemic continues to pose a number of challenges to medical systems around the world, including increased demand for hospital beds and critical shortages of medical equipment, as well as the infection of many healthcare personnel. COVID-19 can be diagnosed quickly and efficiently with effective screening, reducing the load on healthcare systems. Human specialists used to construct deep neural network models and run experiments on them in a trial-and-error technique. This procedure necessitates a significant amount of time, expertise, and resources. To address this issue, a novel but easy approach for automatically performing optimal classification tasks using deep neural network architecture is presented. The suggested COVID detection technique is based on a convolution neural network algorithm that uses a set of neurons to convolve on a given image and extract relevant features.

Each year, drowsy driving causes about 71,000 injuries, 1,500 deaths, and \$12.5 billion in financial losses. Because of the importance of this issue, we believe it is critical to develop a sleepiness detection system, particularly in the early phases, in order to avoid accidents. We also feel that tiredness can have a negative impact on people in the workplace and in the classroom. Our approach to this problem is to create a detection system that recognizes key sleepiness characteristics and sends out an alert before it's too late.

Face detection has been the subject of a lot of research in recent decades. We can recognize a person's face without the assistance of a human. In this study, a system is developed to test the detection of human faces. It's an example of object-class detection in action. The frontal face detection will be the model's top priority. The expected human eye areas are first detected in this face-detection approach by analysing all probable valley regions in a given grey-level frame or picture. The projection of a feature on the eigen-faces determines its fitness value.

2. Overview

2.1. Model Overview: Convolutional neural network

Convolution employs a specialized type of a linear operation. Convolutional neural networks are simple architectures of neural networks with at least one layer that uses convolution instead of ordinary matrix multiplication.

These ConvNets or CNNs form a sub-category of deep neural networks specialized in visual computations, analysis and processing of the image feed given to the network. They're also known as shift invariant or space invariant artificial neural networks due to its translation invariance and shared-weights architecture (SIANN). Recommendation systems, medical diagnosis systems, image classification systems and natural language processing are some of its applications. Multilayer perceptrons are regularized variants of CNNs. They are typically completely connected networks indicating that all neurons in one layer are linked to subsequent neurons in the next layer. However, this feature makes them susceptible to overfitting.

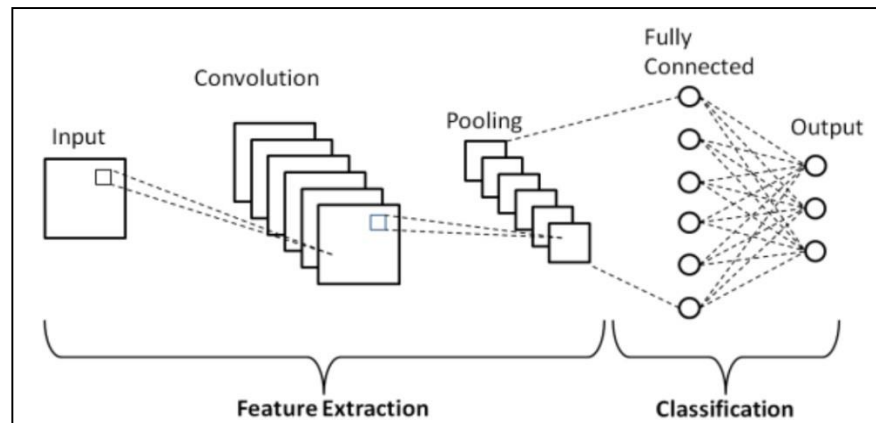


Fig 1.CNN Model

2.1.1. Viola Jones

The Viola Jones algorithm takes its name after two computer vision researchers, Paul Viola and Michael Jones, who devised the technique in 2001. Viola-Jones is a sophisticated algorithm, applied in realistic and practical face detection systems. This algorithm has a longer training period, but when trained can identify faces in real time quickly. The method focuses on multiple subregions of images (grayscale images) and works on recognition of faces by looking for certain attributes in each subregion. The sums of picture pixels within rectangular rectangles are always among the features sought by the detection framework. As a result, they bear resemblance to the Haar functions, formerly exploited in the field of object detection based on images. Viola and Jones' features, on the other hand, are often more complex since they are dependent on multiple rectangular regions. But, while being measured and evaluated against the other alternatives such as steerable filters, rectangular features like this are rudimentary. Despite being sensitive to both horizontal and vertical features, they result in a coarser feedback.

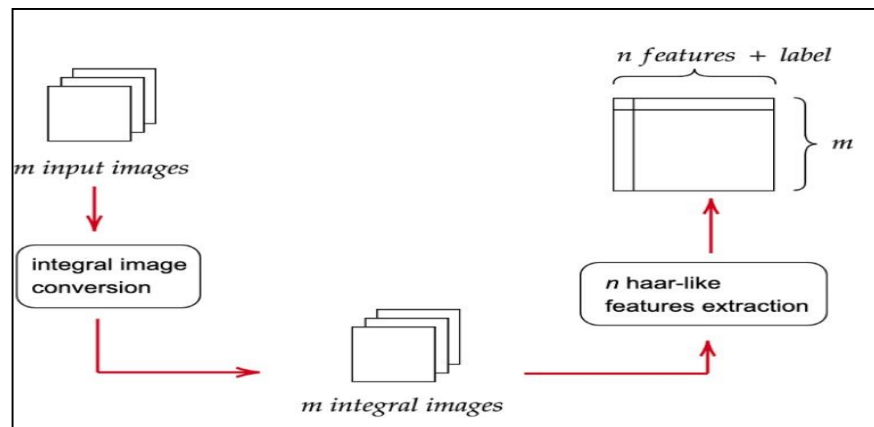


Fig2.ViolaJonesModel

2.1.1. Proposed System

We plan on having a single interface system where the user can choose amongst various tasks to be performed. There are 3 tasks namely Covid-19 Detection, Face& Eye Detection and, Drowsiness Detection that are integrated on a single platform. For example, if the user chooses to perform covid-19 detection, then CNN algorithm is automatically called to complete the task. Various kinds of detections that work on face and eye detections, drowsiness detections can be carried out.

2.1.2. System Design

Our model which is trained to perform three tasks namely Covid-19 detection, Drowsiness detection and Face&Eye detection is developed using Python programming language.

The algorithms that are used for developing the above mentioned three tasks are CNN (Convolutional Neural Networks) for implementing Covid-19 Detection and Drowsiness Detection and **Viola Jones** for implementing Face&Eye Detection.

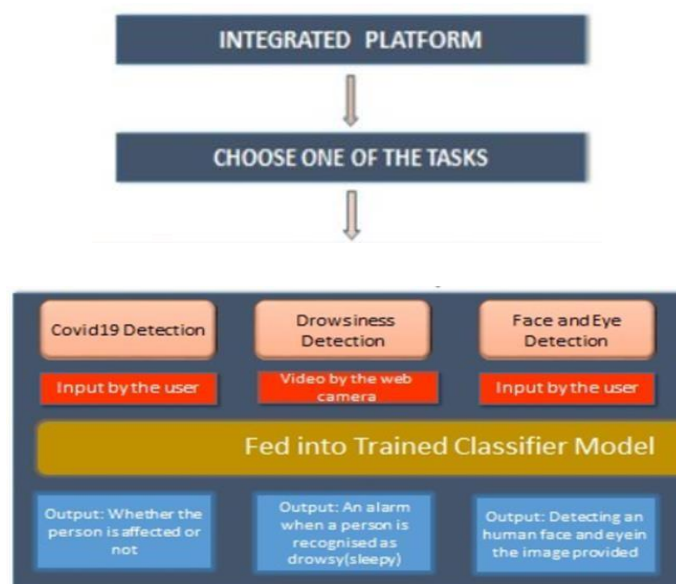


Fig3. Flow Chart

Steps:

- The user access the platform where he/she can choose any one task at a given time among the three tasks
- On choosing Covid-19 detection, the x-ray image needed to be tested is given by the user. On choosing Drowsiness detection, the live video is captured by turning on the user's webcam and on selecting Face & Eye detection, the input image is given by the user for detecting purpose
- Once the user selects any one of the task, it is fed into the classifier model and the model undergoes training where it is trained using suitable algorithms
- Finally, the output is given by the training model after successful training
- Output of covid-19 detection is whether the person is normal or covid affected, output of drowsiness detection is an alarm beeping and the output of face & eye detection is a rectangular border around the face and eye.

2.2. Architectural Design

2.2.1. Convolutional Neural Network

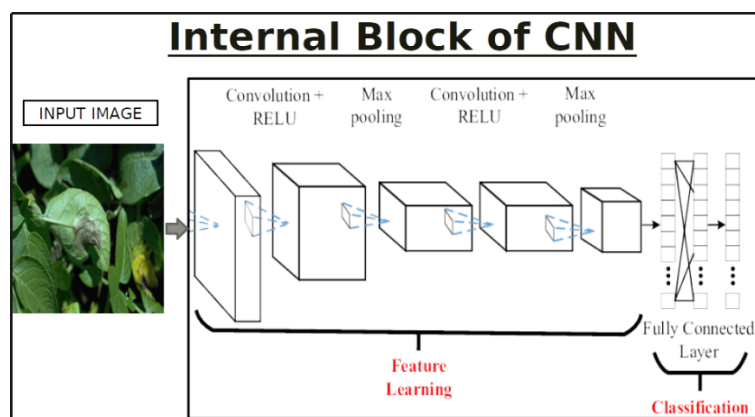


Fig 4. Architectural Design of CNN

The brief understanding of the convolutional layers in the architecture is explained as:

The first layer, convolution, extracts parameters from the image input to the network and gains an understanding of the relationship between them using kernel or filter input images. ReLU Layer- For a non-linear process, ReLU stands for Rectified Linear Unit. $f(x) = \max(0, x)$ forms the result. We utilize this to introduce CNN to non-linearity.

Pooling Layer: This layer is used to decrease the amount of parameters by down sampling and keep only the most important data for further processing. There are three types of pooling: maximum pooling (which you should choose), average pooling, and sum pooling.

Flattening: We convert our entire matrix into a vertical vector. In order for it to be transmitted to the input layer. We pass our flatten vector into the input Layer, which is fully connected. We created a model by combining these features. Finally, to classify the outputs, we use an activation function such as softmax or sigmoid.

2.2.2. Viola Jones

Viola Jones architecture consists of the following layers:

- Selecting Haar-like features

Object recognition uses digital image properties known as Haar-like features. Some universal features of the human face can be present in all human faces, such as the eyes being darker than its neighbours and the nose being brighter than the eyes. If one side is lighter than the other, it might represent the edge of a brow, or the center region of the box could be more illuminating than the neighboring pixels, indicating a nose. As a result, we can use Haar-like traits to interpret the various aspects of the face.

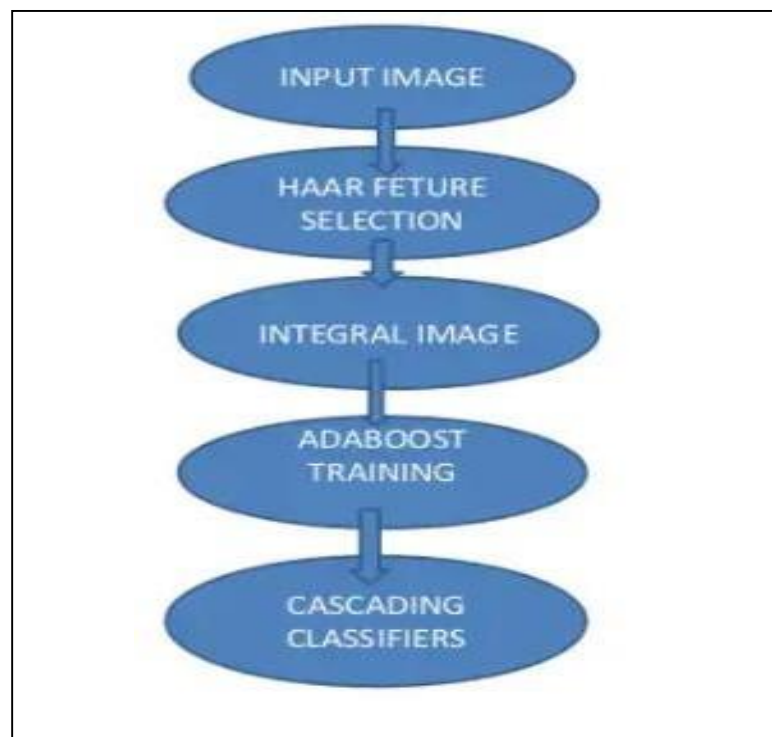


Fig5.ArchitectureDesignofViolaJones

- Creating an integral image

An integral image (also known as a summed-area table) is a representation of an image that is intermediate. It's used to add up the pixel values of an image or a rectangular section of an image. This leads to time efficiency during computations involving the summation of pixels in a rectangular region since we only have to do computations on four edges of the rectangle.

- Running Ada Boost training

Only a few of the traits displayed in the detector window are crucial for identifying a face. The AdaBoost method was used to find the best features. In the Viola-Jones technique, each Haar-like feature reflects a poor learner. AdaBoost assesses the performance of the classifiers you submit in order to select the type and size of parameter which will be contained in the final classifier. The weight or importance of the successful classifiers is enhanced. The final result is a strong classifier, otherwise called a boosted classifier.

- **Creating Cascading Classifiers**

The AdaBoost will eventually select the optimal features, but calculating these features for each region is still a time-consuming operation. The cascade's job is to quickly eliminate non-faces so that time and computations are saved. As a result, the required speed for real-time face detection has been met.

3. Implementation

3.1. Implementation of Each Module

There are 3 modules in our project namely Covid-19 Detection, Drowsiness Detection and Face & Eye Detection. User can choose any one of the task among the given 3 tasks.

3.1.1. Covid-19 Detection

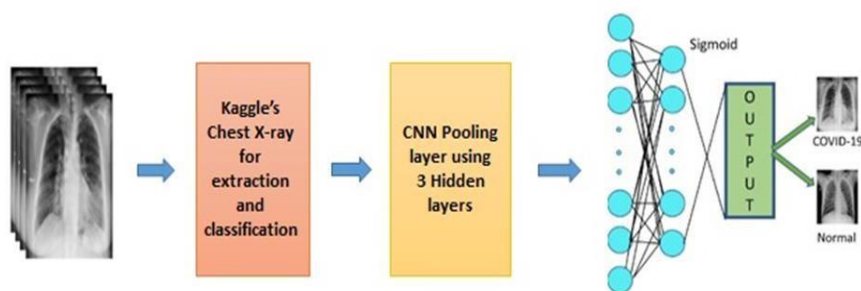


Fig 8. Covid-19DetectionArchitecture

The architecture design of Covid-19 detection is given above. The user has to provide a chest x-ray image as an input and the model is trained on Kaggle dataset using CNN algorithm. The output is whether the person is infected or not.

3.1.2. Drowsiness Detection

For Drowsiness detection, the model makes use of the webcam to detect the eye of the user. Initially the model is analyzed by viola jones to detect the face and eye of the user. Later the CNN algorithm acts upon it and once the person closes the eye for more than certain time (say 15 sec), the alarm starts up which alerts the user.

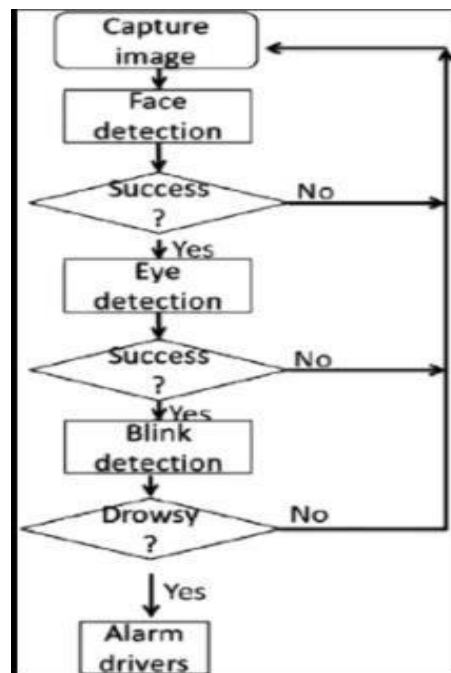


Fig9. Drowsiness Detection Architecture

3.1.3. Face & Eye Detection.

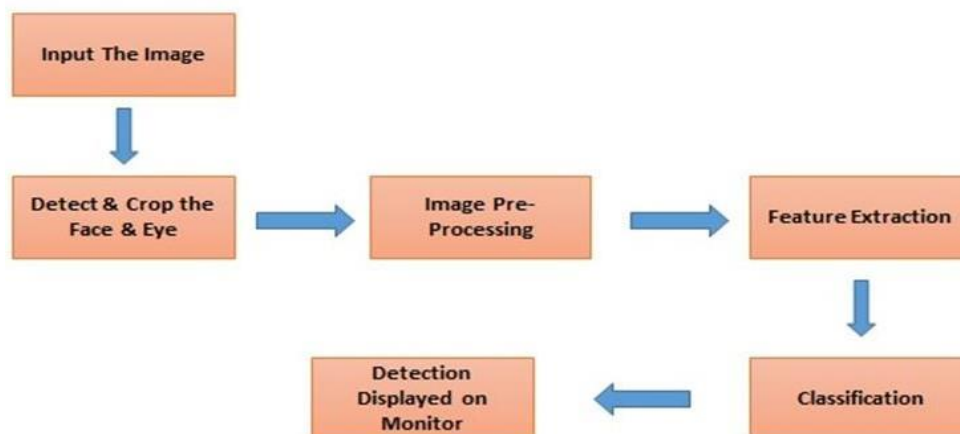


Fig10.Face&EyeDetectionArchitecture

On choosing Face & Eye Detection, the user has to provide an image as the input and preprocessing of the image takes place to improve the quality of the image for further analyzing. The next step is feature extraction where the raw data is reduced to more manageable groups for processing. Finally, the model is trained using Viola Jones algorithm to detect the face and eye of the user (draws a square around the face and eye of the image).

4. Experimental Results

Here is the result and analysis of three modules of our project.

- Covid-19 Detection analysis: When our model has completed the training assignment, the picture must be imported. The training data conducts shearing, zooming, and horizontal shifts, resulting in accuracy. These features are used to shape data for better reception.

- Drowsiness Detection analysis: To start this module, command prompt should be opened. Navigate to the directory where the main file exists and run this file. It may take a few seconds to open the webcam and once the webcam is started, the detection process starts.
- Face and Eye Detection analysis: Here an image is given as an input to detect the face and eyes. After successful detection, a rectangular box is drawn around the face and eyes of the input image.



Fig11.MLHUBHomePage.

The ML HUB website's home page which offers the user three alternatives to choose from. The options listed on the website are Covid-19 Detection, Drowsiness Detection, and Face and Eye Detection. The user can click on any one of the link to check for its working.

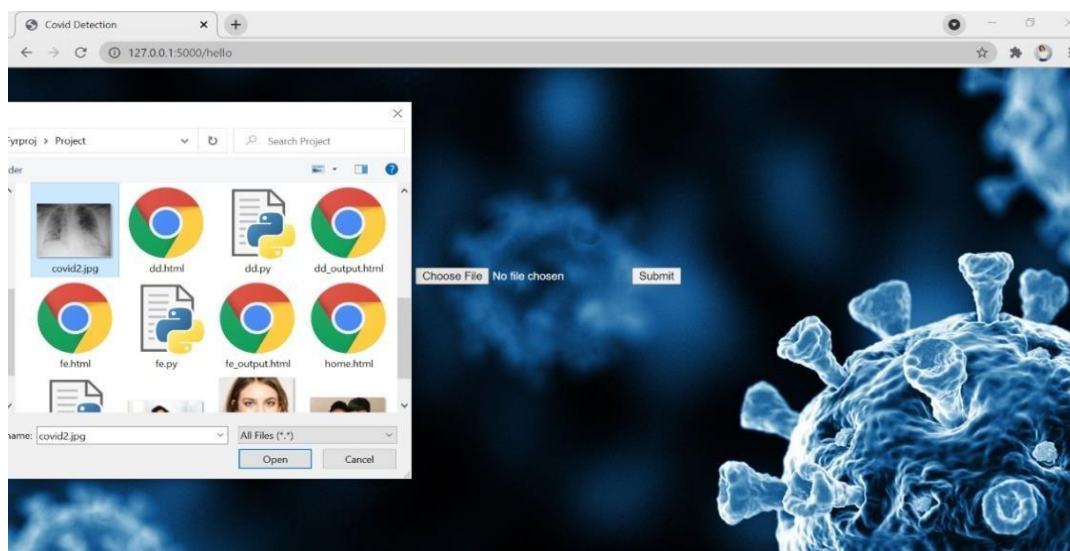


Fig12. Covid-19DetectionWindow

When the covid detection option is chosen, the user is directed to this page, where they must upload an image to check whether it is covid affected or not. On clicking submit button, the user will be taken to the final page which displays the result as covid affected or normal.

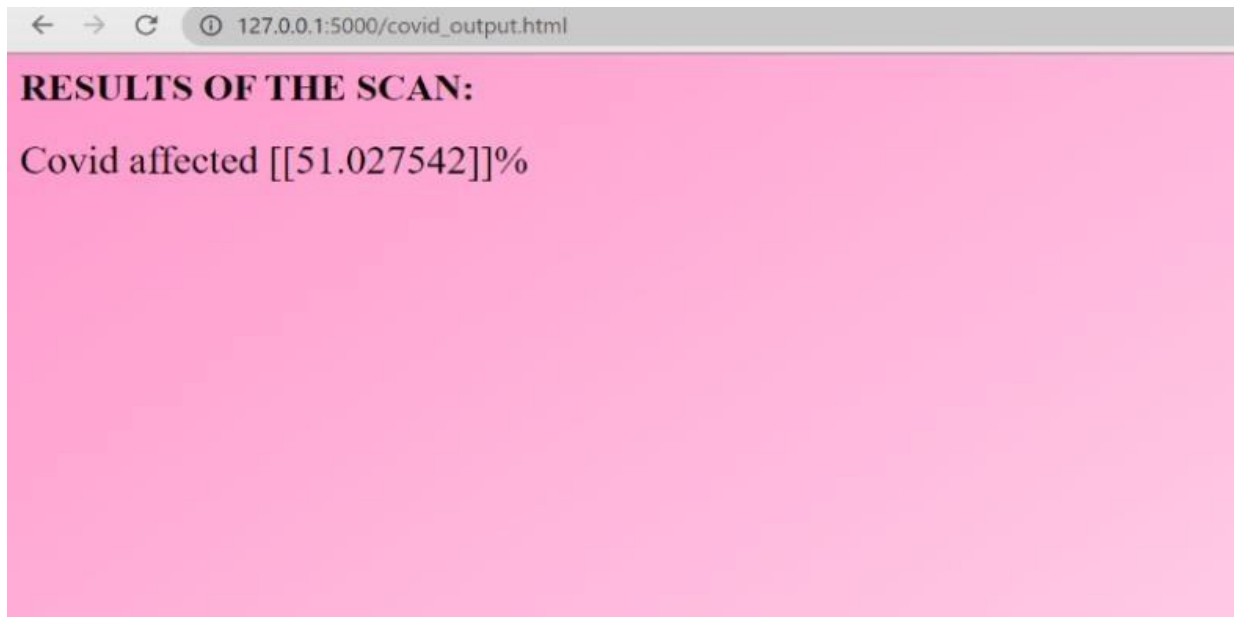


Fig13.Covid-19DetectionResult

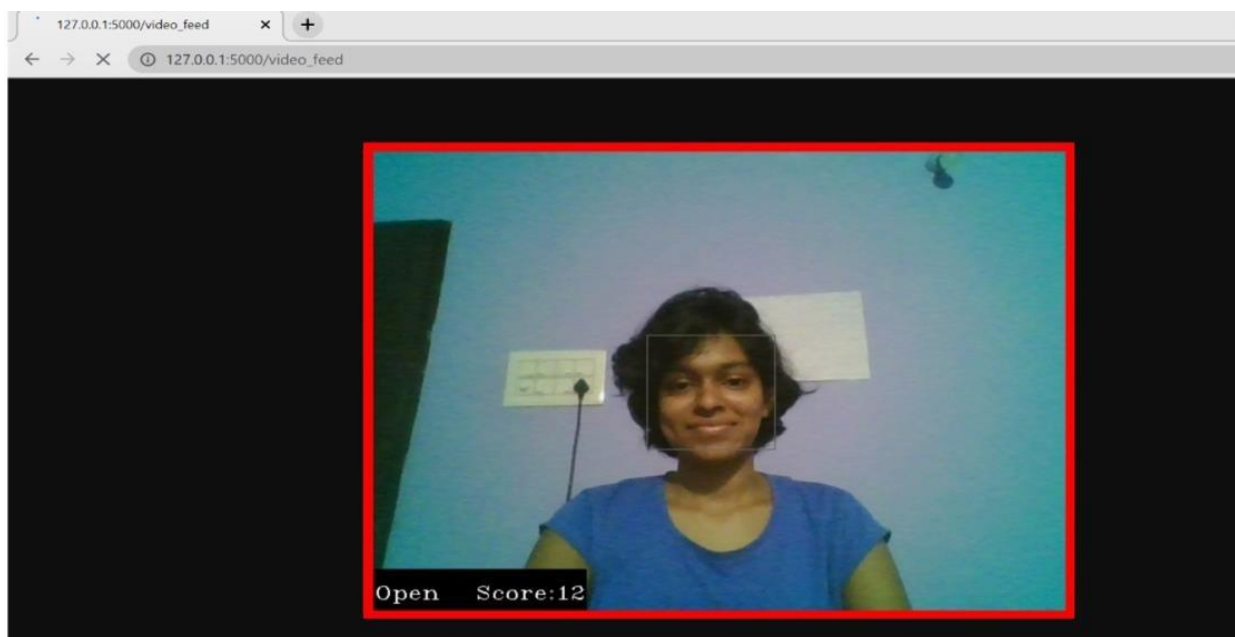


Fig 14. Drowsiness Detection Window (Open Eyes).

When the drowsiness detection option is selected, the user is sent to this web page, where the webcam is activated, displaying the user's status, including whether or not their eyes are closed and the score which keeps the track on user's eye closing rate. Once the user closes their eyes for more than certain time, an alarm goes off which alerts the user.

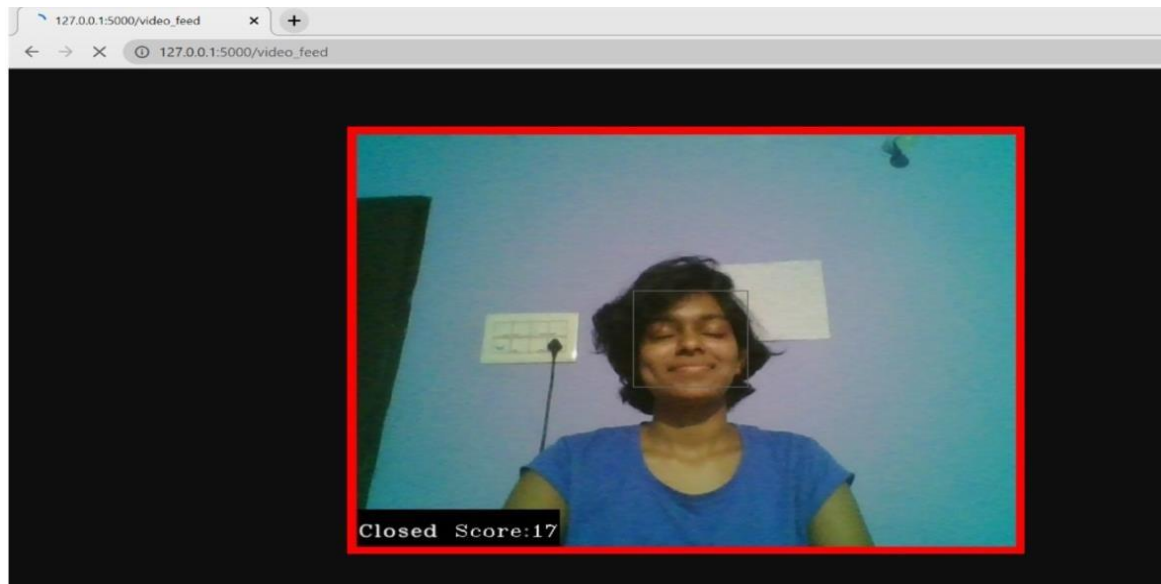


Fig15.DrowsinessDetectionWindow (Closed Eyes).

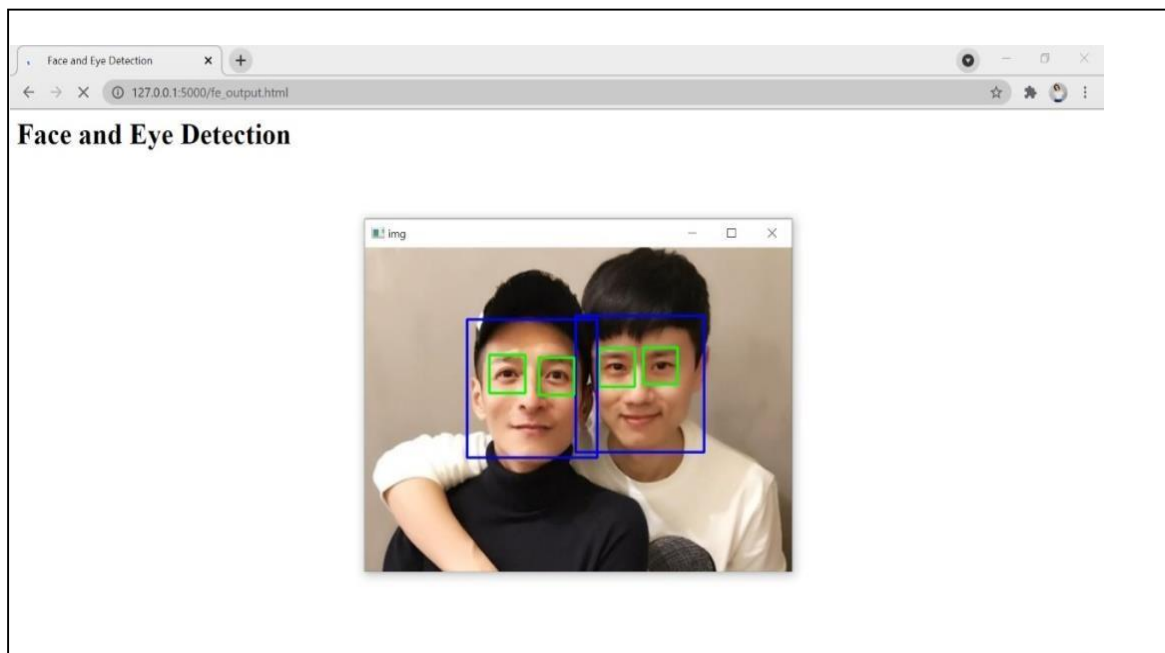


Fig16. Face & eye Detection Window

This is the face & eye detection window which asks for an input image for detection. After uploading an image, click on the submit button to view the result. After successful detection, a rectangular box is drawn around the face and eyes as shown in the below diagram.

Conclusion

In the realm of computer science, machine learning is a fast expanding topic. Machine learning has benefits in nearly every other field of study, and it is also being used commercially to tackle issues that people find too complex or time consuming to handle. In other words, machine learning utilizes a range of models to find patterns in data and generate precise forecasts based on them.

A crash course on generalization and over fitting. These difficulties are related to supervise learning, which entails feeding data into the model. When a machine learning model generalizes, it gains the ability to predict outcomes from data that it has never seen before. When a model learns the training data too well, it becomes over fit and incapable of generalization. With supervised learning, the polar opposite of over fitting can occur. Under fitting prevents the model from making correct predictions with both training and new data. Following that, a conversation about datasets ensued. In supervised learning, data is classified into three types: train, development, and test datasets. The train dataset is used to train the model. The development dataset has been used to test the model when it's being developed but not for training. When the model is ready, the test dataset is used to examine how it reacts to data that it has never seen before. It is also explained how to choose relevant fields from a dataset. Datasets should not contain information that is no longer valuable.

The first model in this series of blog posts, artificial neural networks, is then studied. There are three layers in a neural network: input, hidden, and output. Each layer is comprised of nodes. The layers are linked together using vectors. Neural networks were one of the first machine learning models to be constructed, and several types of neural networks have been studied. Take, for example, deep neural networks.

Deep neural networks have several hidden layers, whereas artificial neural networks have only one. Due to the added complexity that hidden layers provide to the model, deep neural networks outperform basic neural networks in some tasks. Yet, as their complexity grows, they become increasingly difficult to train.

Finally, we'll talk about convolutional neural networks. This is a distinct kind of fundamental neural network. A convolutional neural network has the advantage of being great at managing voice and image detection jobs. Convolutional neural networks feature a convolutional and a pooling layer instead of hidden layers. Convolutional neural networks are favored for picture and speech control because to their layers.

References

1. M. T. Schafer and K. van Es, "Social data," in *The datafied society: Studying culture through data*. Amsterdam University Press, 2017, ch. 10, pp. 147–154.
2. A. M. Kaplan and M. Haenlein, "Users of the world, unite! the challenges and opportunities of social media," *Bus. Horizons*, vol. 53, no. 1, pp. 59–68, 2019.
3. P. Gundecha and H. Liu, "Mining social media: a brief introduction," in *New Directions in Informatics, Optimization, Logistics, and Production*. Informs, 2012, pp. 1–17.
4. Z. Tufekci, "Big questions for social media big data: Representativeness, validity and other methodological pitfalls," *ICWSM*, vol. 14, pp. 505–514, 2014.
5. D. Ruths and J. Pfeffer, "Social media for large studies of behavior," *Science*, vol. 346, no. 6213, pp. 1063–1064, 2014.
6. Chen and D. B. Neill, "Human rights event detection from heterogeneous social media graphs," *Big Data*, vol. 3, no. 1, pp. 34–40, 2015.

7. M. De Choudhury, "Anorexia on it: A characterization study," in *Proceedings of the 5th International Conference on Digital Health 2015*. ACM, 2015, pp. 43–50.
8. Chengang Zhu, Guang Cheng (senior member, IEEE) and Kun Wang (senior member, IEEE), "Big data analytics for program popularity prediction in broadcast TV industries", *IEEE Access*, October'2017.
9. MohammadNoor Injadat, Fadi Salo, Ali Bou Nassif, "Data mining techniques in social media: A survey", *Neurocomputing- ScienceDirect*, June'2016.
10. Marouane Birjali, Abderrahim Beni-Hssane, Mohammed Erritali, "Analyzing social media through big data InfoSphere BigInsights and Apache Flume", *ScienceDirect*, June'2017.
11. KuaiXu ,Feng Wang, Haiyan Wang, and Bo Yang "Detecting Fake News Over Online Social Media via Domain Reputations and Content Understanding" *TSINGHUA SCIENCE AND TECHNOLOGY*,ISSN 11007 - 02140 3/ 1 4 pp 2 0 – 2 7, DOI: 10.26599/TST.2018.9010139, Volume 25, Number 1, February 2020.
12. TAO JIANG, JIAN PING LI, AMIN UL HAQ, ABDUS SABOOR, AMJAD ALI, "A Novel Stacking Approach for Accurate Detection of Fake News" *Digital Object Identifier 10.1109/ACCESS.2021.3056079*.
13. MUHAMMAD UMER, ZAINAB IMTIAZ, SALEEM ULLAH, ARIF MEHMOOD, GYU SANG CHOI, BYUNG-WON ON "Fake News Stance Detection Using Deep Learning Architecture (CNN-LSTM)" *Digital Object Identifier 10.1109/ACCESS.2020.3019735*.
14. XINYI ZHOU, REZA ZAFARANI, "A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities" *ACM Comput. Surv.* 1, 1, Article 1 (January 2020), 37 pages.
15. Shivani S Nikam, Prof. Rupali Dalvi, " Machine Learning Algorithm based model for classification of fake news on Twitter", 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), DOI: 10.1109/I-SMAC49090.2020.9243385.
16. Anand.R Pushpalatha.M Dr Rajshekhar M Patil, "A Social Networking for Sharing Infrastructure Resources in the Social Cloud Computing", *International Journal of Informative & Futuristic Research (IJIFR)*, 2016.
17. Dr. M V Vijaykumar Jagadish P, Shryavani K, Anand R, "Authorized Deduplication in Hybrid Cloud", *IJCSN International Journal of Computer Science and Network*, 2016.
18. Mr. Anand.R Priyanka Dr. Rajshekhar M. Patil, "Health Monitoring In Aerospace System", *International Journal of Informative & Futuristic Research (IJIFR)*, 2017.