1-The project idea in details:

a) Overview

Object detection technique helps in the detection, and localization of multiple visual instances of objects in an image. It provides a much better understanding of the object as a whole, rather than just basic object classification. This method can be used to count the number of instances of unique objects and mark their precise locations, along with labeling. With time, the performance of this process has also improved significantly, helping us with real-time use cases. All in all, it answers the question: "What object is where and how much of it is there?".

Object detection is commonly confused with image recognition, so before we proceed, it's important that we clarify the distinctions between them.

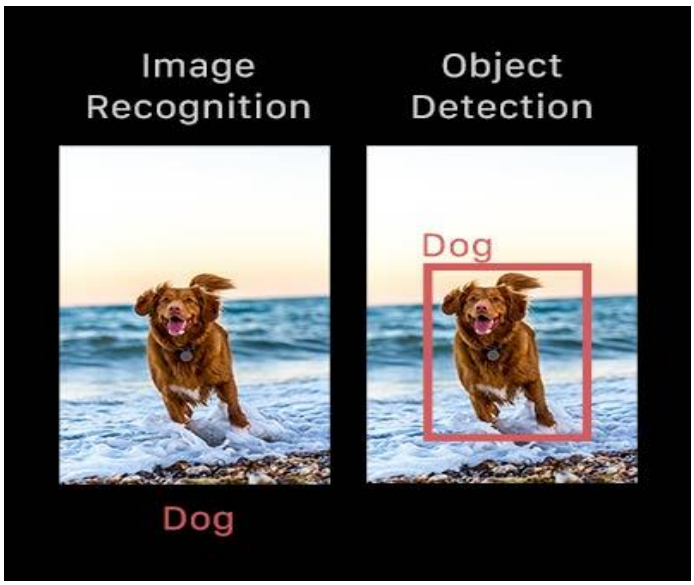
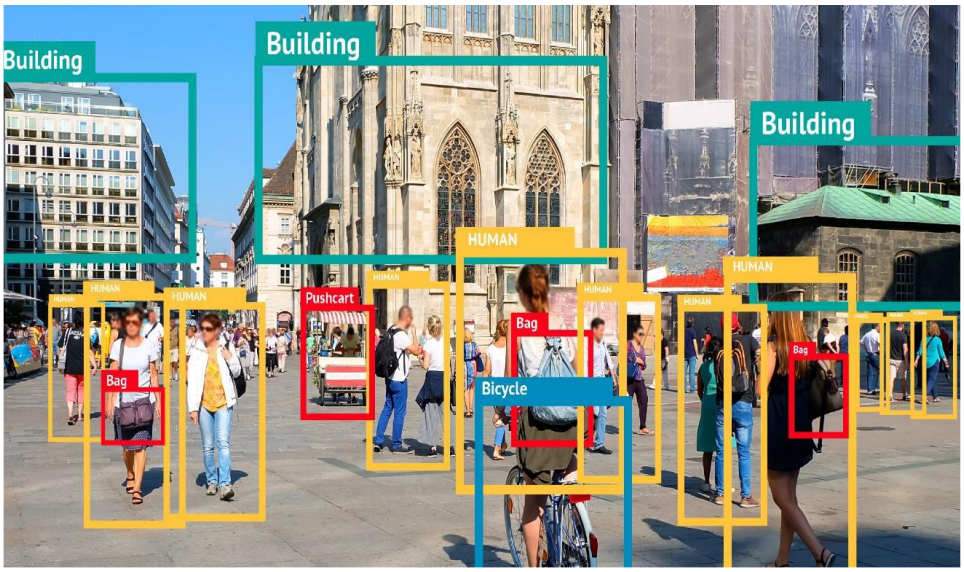


Image recognition assigns a label to an image. A picture of a dog receives the label "dog". A picture of two dogs, still receives the label "dog". Object detection, on the other hand, draws a box around each dog and labels the box "dog". The model predicts where each object is and what label should be applied. In that way, object detection provides more information about an image than recognition.



b) How it works

The main concept behind object detection is that every object will have its features. These features can help us to separate objects from the other ones. **Object detection** methodology uses these features to classify the objects.

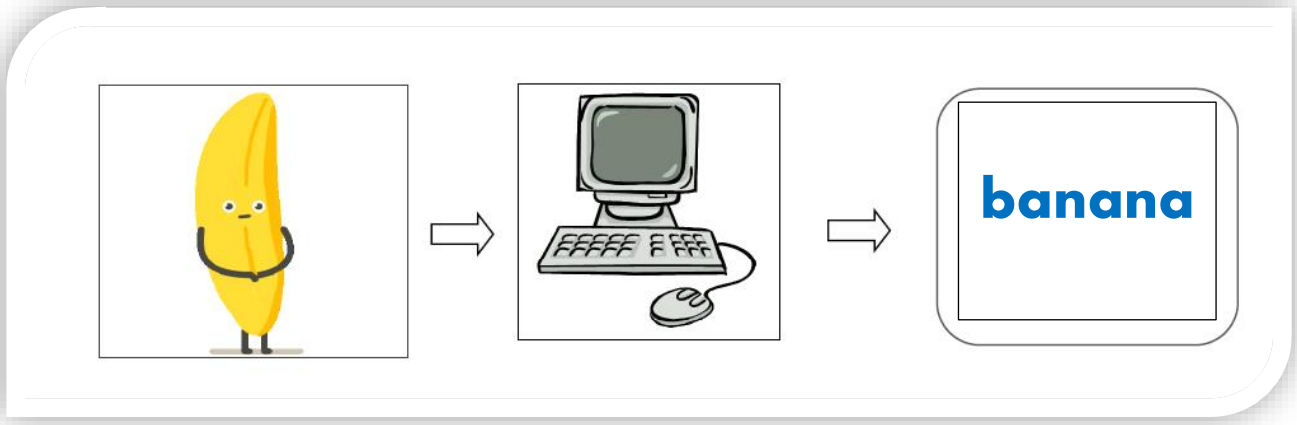
In order to achieve this, we can have a variety of approaches, but there are two main approaches- a machine learning approach and a deep learning approach. In more traditional ML-based approaches, computer vision techniques are used to look at various features of an image, such as the color histogram or edges, to identify groups of pixels that may belong to an object. These features are then fed into a regression model that predicts the location of the object along with its label.

one of the machine learning approach (which requires the features to be defined explicitly) algorithms is ANN, we make an architecture of artificial neural networks (which are brain-inspired mathematical models intended to replicate the way that humans learn) and train it with some open source dataset to make a real-time output in order to be able to detect any given object in real time and thus making the output with the name of the object around it.

As a user view: we get a set of data images for different objects and code our system to learn these images by knowing the features of every object that could be used to differentiate it with other objects.

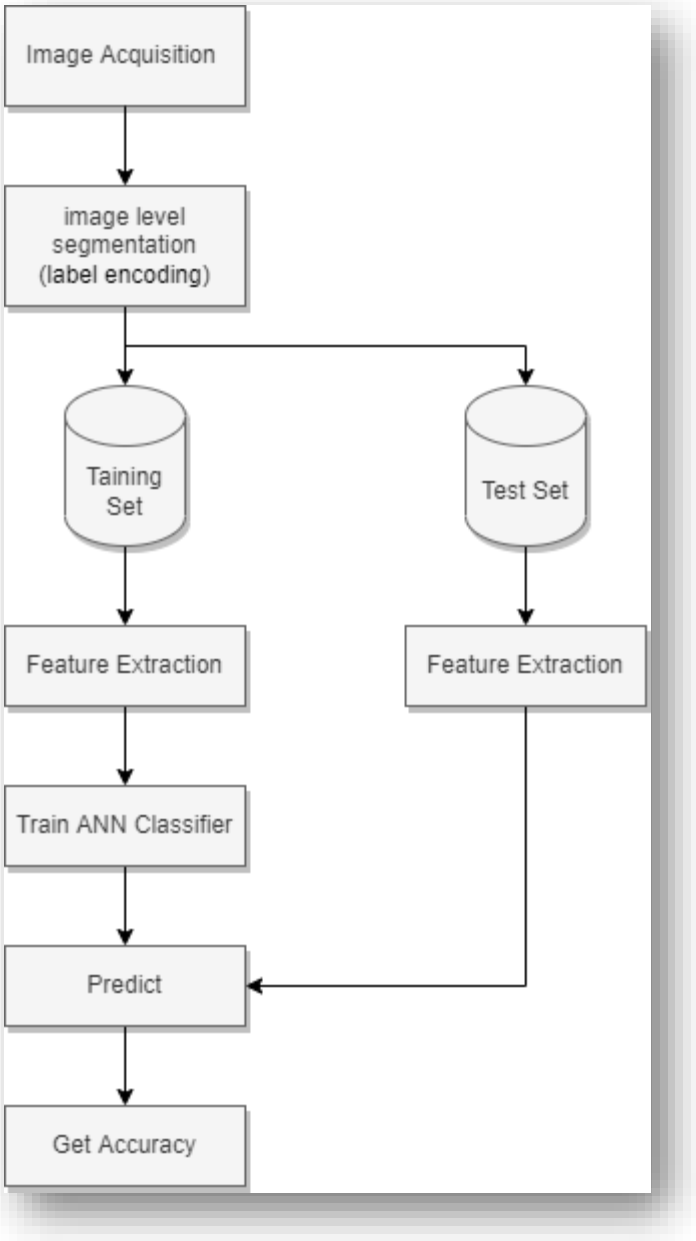
Ex: a car has some features like shape of tiers, window, color … etc.

After that the computer now the features of every object so, we pass the source of an image or video (input pc& mobile camera in our case) and then through an algorithm our Ai system can clearly detect the object in this image.



Implementation view:

Block Diagram



2- Main functionalities:

Get Data

First, we get data from dataset to detect on it to review the out-put.

OS library

Python OS module allows us to use the operating system dependent functionalities and to interact with the underlying operating system in several different ways.

CV2

OpenCV is an open-source Python library, which used to understand the content of the digital image.

Preprocessing

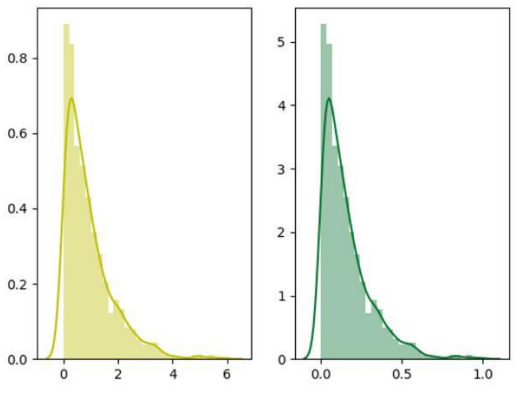
Then we get image from dataset we are doing some processing on it to simulate the algorithm.

Resize

now we resize image to make sure that all images will get the same size to provide the algorithm.

Normalization

refers to rescaling real-valued numeric attributes into a [0 ,1] range. is used in machine learning to make model training less sensitive to the scale of features.



Left: Original Data, Right: Normalized Data

#Label encoding

Label Encoding refers to converting the labels into a numeric form so as to convert them into the machine-readable form. Machine learning algorithms can then decide in a better way how those labels must be operated. It is an important pre-processing step for the structured dataset in supervised learning.

#Test and split (train_test_split)

That we get the dataset and divide it in two part like 80 % training and 20 % testing.

#Feature extract

#VGG16

VGG16 is a convolution neural net (CNN) architecture which was used to win ILSVR(Imagenet)

One very important area of application is image processing, in which algorithms are used to detect and isolate various desired portions or shapes (features) of a digitized image or video stream. It is particularly important in the area of optical character recognition.

Artificial Neural Networks

#Fit

Given a Dataset comprising of a group of points, find the best fit representing the Data.

We often have a dataset comprising of data following a general path, but each data has a standard deviation which makes them scattered across the line of best fit.

#Predict

When you get the image and get the Feature extraction from testing images, we predict the image

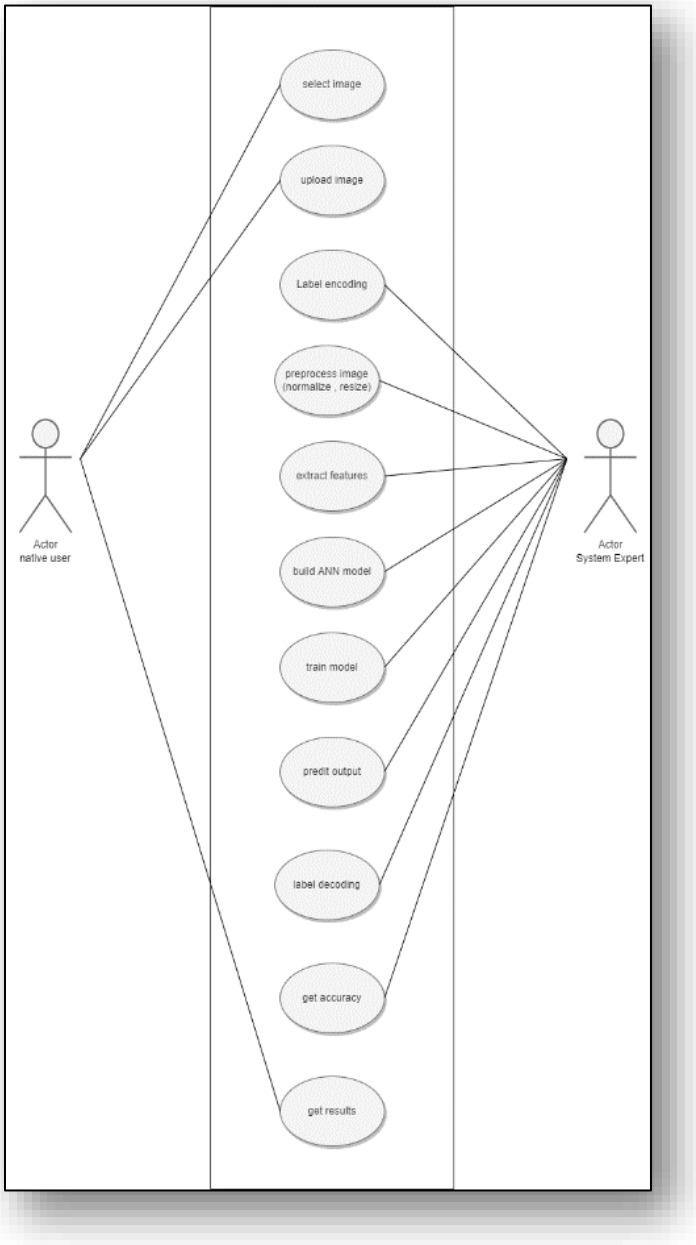
#Get accuracy

Accuracy is a measure for the closeness of the measurements to a specific value, while precision is the closeness of the measurements to each other, i.e. not necessarily to a specific value.

For example: how many times that the program is doing the algorithm as percentage.

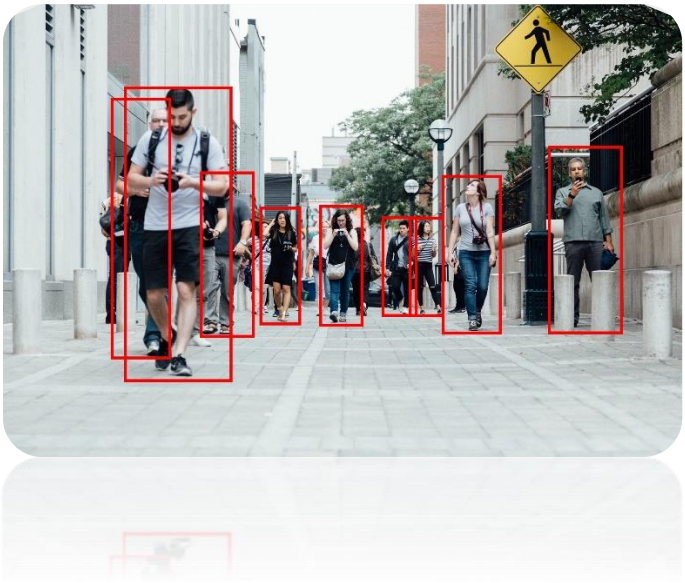
#Decoding

Use-Case Diagram

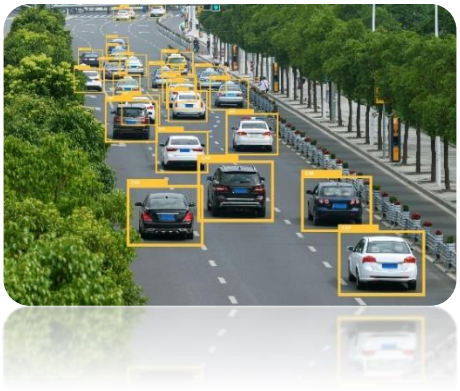


3-Applications in the market :

Person detection: **Person detection is undertakings of Computer vision frameworks for finding and following individuals. Person detection is the task of finding all examples of individuals present in a picture, and it has been most broadly achieved via looking through all areas in the picture, at all potential scales, and contrasting a little region at every area with known layouts or examples of individuals.**



Vehicle Detection: **Vehicle Detection is one of the most important part in our daily life. As the world is moving faster and the numbers of cars keep on increasing day by day, Vehicle detection is very important. By using the Vehicle Detection technique we can detect the number plate of a speeding car or accident-affected car. This also enables for the security of society and decreases the number of crimes done by car. By using Vehicle Detection Technology Pixel Solutions have successfully detected the speed of the vehicle and we have also detected the number plate of the car using Optical Character Recognition (OCR).**



Self-Driving Cars: In order for a vehicle to drive autonomously in its environment, it must know the location of all the objects surrounding it accurately to be able to decide the next move and execute it safely.

Carmakers are integrating more and more sensors like cameras, lidar, and radar into their cars, and algorithms like object detection to detect cars, pedestrians, traffic signals, etc, and estimate their locations … to reach the ultimate goal and deliver the vehicle of the future.

Police and forensic: Object detection can track and locate specific objects such as a person, vehicle, or backpack from frame to frame. It allows police officers and forensic professionals to inspect every nook and corner of a crime site to collect evidentiary proofs. However, due to the presence of a large volume of data, the process of object detection is a bit tricky and requires hours of footage to identify what can aid in the success of a case.

Parking system: **Pre-integrated visual detectors in cars can detect open parking spaces in surface lots or parking garages. It can also provide the driver with a front and rear view of the parking space and other vehicles to securely park the car.**

Similar applications in object detection project:

AI Dielmo:

MADE EASY

Computer vision & object detection

https://ai.dielmo.com/

alwaysAI:

Computer Vision with Object Detection

https://learn.alwaysai.co/object-detection-ga?utm_campaign=CV%20Services%20-%20GA&utm_source=ppc&utm_medium=CV%20Services&utm_term=Object%20Detection&utm_campaign=CV+Services&utm_medium=ppc&utm_source=adwords&utm_term=object%20detection&hsa_mt=b&hsa_net=adwords&hsa_ver=3&hsa_kw=object%20detection&hsa_acc=9308646013&hsa_grp=92640778004&hsa_tgt=kwd-297297792206&hsa_src=g&hsa_ad=413433580518&hsa_cam=9005147173

Pixel Solutionz:

https://www.pixelsolutionz.com/application-object-detection-real-life/

Ui Path:

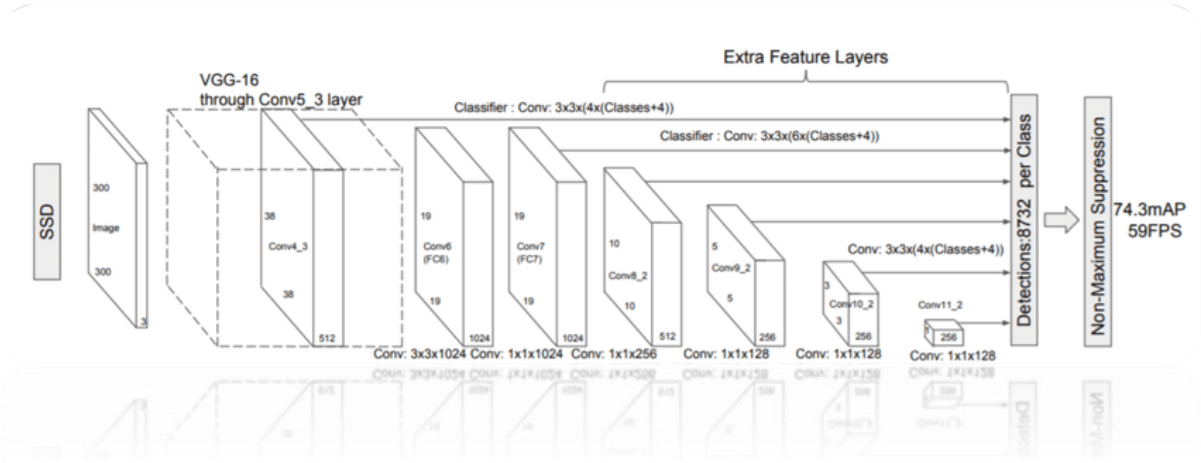
https://www.uipath.com/

The initial literature review of Academic publications :

I. Face Mask Detector

After the breakout of the worldwide pandemic COVID-19, there arises a severe need of protection mechanisms, face mask being the primary one. The basic aim of the project is to detect the presence of a face mask on human faces on live streaming video as well as on images. We have used deep learning to develop our face detector model. The architecture used for the object detection purpose is Single Shot Detector (SSD) because of its good performance accuracy and high speed. Alongside this, we have used basic concepts of transfer learning in neural networks to finally output presence or absence of a face mask in an image or a video stream. Experimental results show that our model performs well on the test data with 100% and 99% precision and recall, respectively.

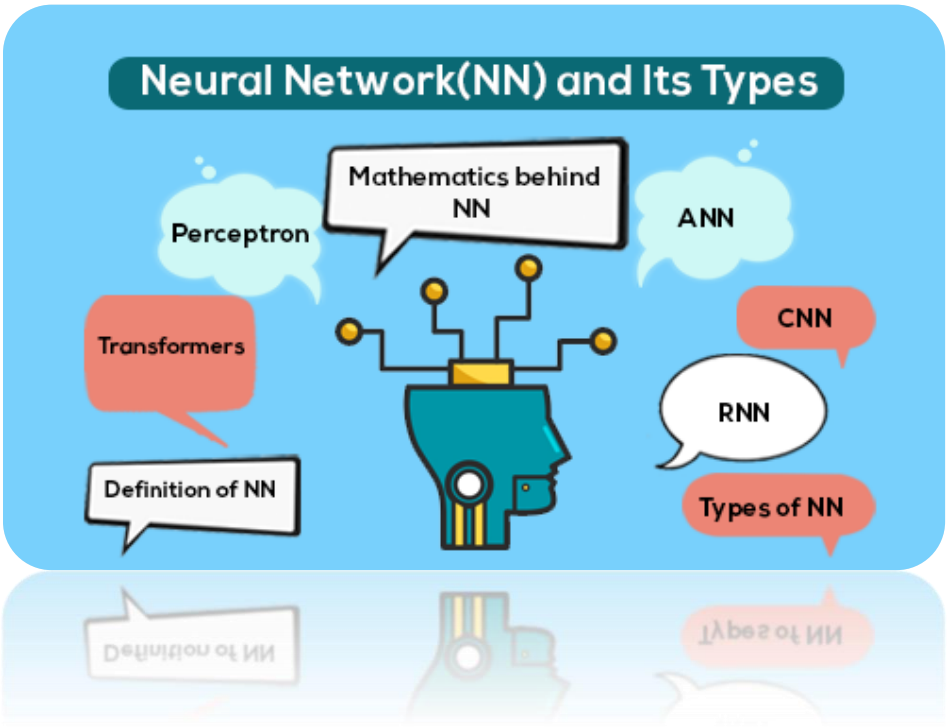
https://www.researchgate.net/publication/344173985_Face_Mask_Detector



II. A Literature Survey: Neural Networks for object detection

Humans have a great capability to distinguish objects by their vision. But, for machines object detection is an issue. Thus, Neural Networks have been introduced in the field of computer science. Neural Networks are also called as 'Artificial Neural Networks' . Artificial Neural Networks are computational models of the brain which helps in object detection and recognition. This paper describes and demonstrates the different types of Neural Networks such as ANN, KNN, FASTER R-CNN, 3D-CNN, RNN etc. with their accuracies.
From the study of various research papers, the accuracies of different Neural Networks are discussed and compared and it can be concluded that in the given test cases, the ANN gives the best accuracy for the object detection.

http://www.viva-technology.org/New/IJRI/2018/9.pdf



III. Driving Dataset Literature Review

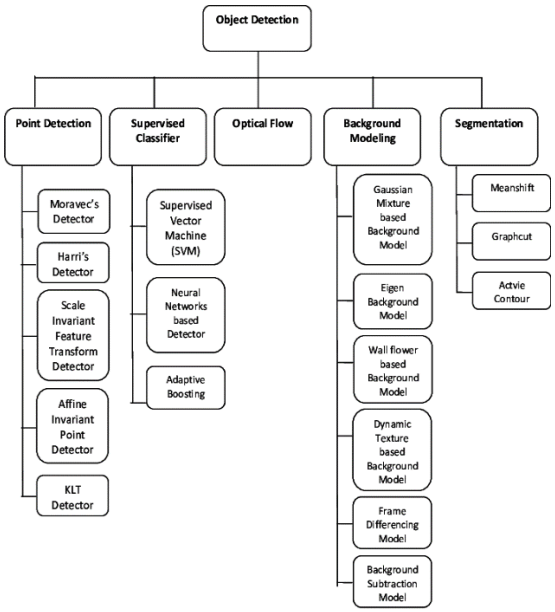
For the last decade, the progress made in the autonomous driving scientific community and industry has been exceptional. With the rise of deep learning and better hardware, algorithms embodying the different aspects of driving ,such as lane following, obstacle detection, semantic segmentation, tracking, and motion estimation have reached unprecedented performance. Although there are still no SAE Level 4 self-driving vehicles as of yet, recent developments in robotics and machine learning could soon make this aspiration a reality.

https://www.researchgate.net/publication/336869619_Driving_Datasets_Litera-ture_Review

IV. A survey on Real time Object Detection and Tracking Algorithms

Real time object detection and tracking are important and challenging tasks in many computer vision applications such as video surveillance, robot navigation and vehicle navigation. Object detection involves detecting the object in sequence of videos. Every tracking mechanism requires object detection mechanism either in each frame or when an object appears newly on the video sequence. Object tracking is the process of locating an object or multiple objects using either static or dynamic camera. The availability of high powered computers, high quality and inexpensive video cameras will increase need for automated video analysis. It has generated a great deal of interest in object detection and tracking algorithms. Even though high powered computers are used for object detection and tracking algorithm, most of the object detection algorithms such as background subtraction, temporal difference, foreground extraction and simple differencing requires long time to detect object ,requires more storage space and no robustness against illumination changes. Recently computer vision research has to address the Multiple object detection and tracking in dynamic environment. In this paper, we address the various object detection and tracking algorithms.

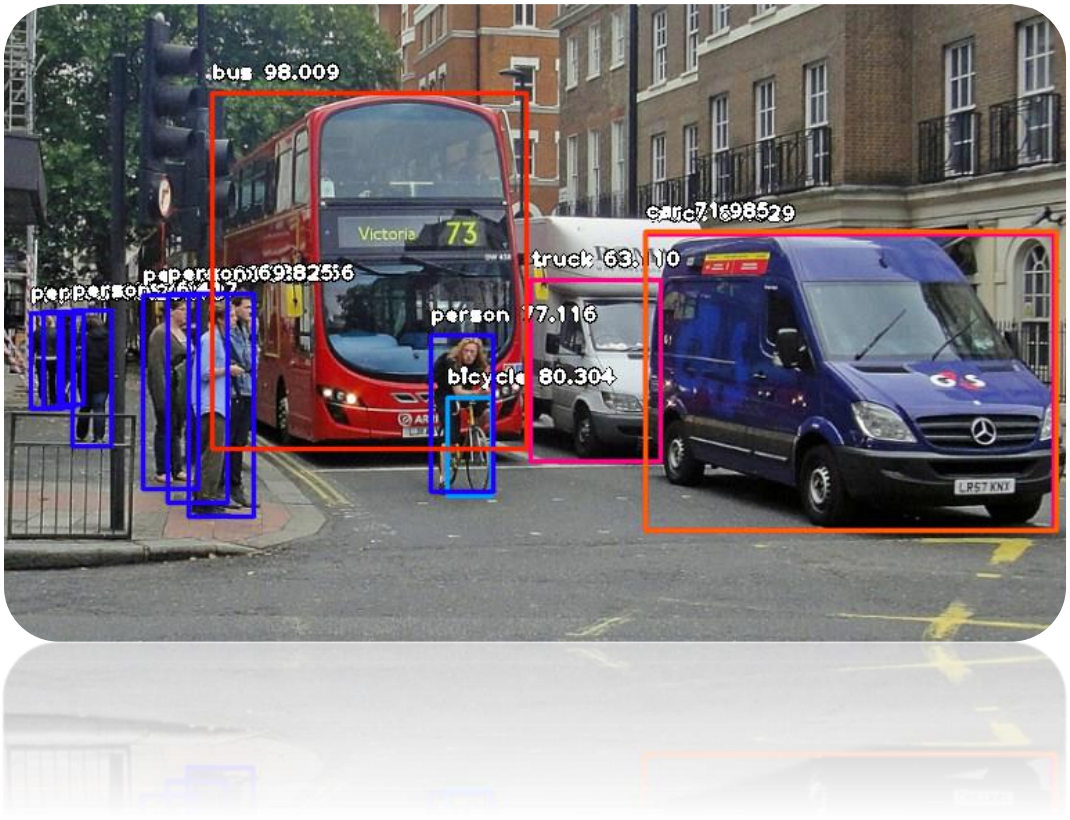
https://www.researchgate.net/publication/274569172_A_survey_on_Real_time_Object_Detection_and_Tracking_Algorithms



V. Object Detection Using ImageAI

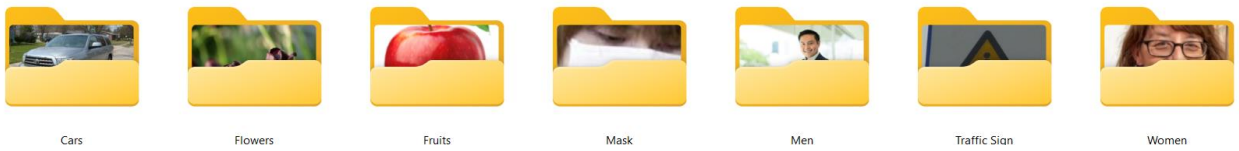
ImageAI provides very powerful yet easy to use classes and functions to perform Image Object Detection and Extraction.
ImageAI allows you to perform all of these with state of the art deep learning algorithms like Retina Net, YOLOv3 and TinyYOLOv3. With ImageAI you can run detection tasks and analyse images.
Find below the classes and their respective functions available for you to use. These classes can be integrated into any traditional python program you are developing, be it a website, Windows/Linux/MacOS application or a system that supports or part of a Local-Area-Network.

https://imageai.readthedocs.io/en/latest/detection/



5- The Dataset:

Our dataset consists of 7 categories



Category 1:

Name: cars

Number of images: 1000

Category 2:

Name: flowers

Number of images: 986

Category 3:

Name: fruits

Number of images:990

Category 4:

Name: men

Number of images: 990

Category 5:

Name: Traffic sign

Number of images: 963

Category 6:

Name: Women

Number of images: 990

Category 7:

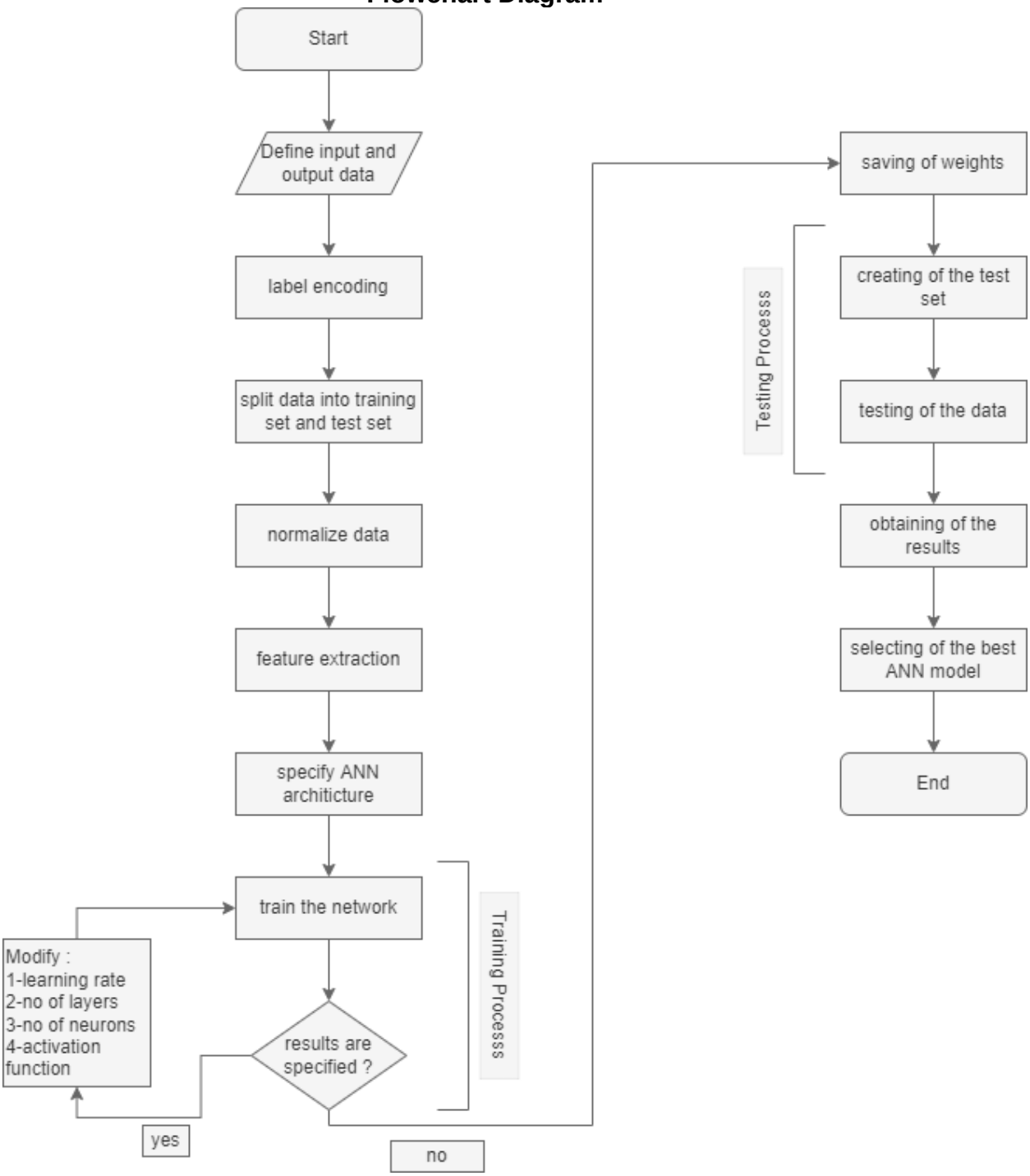
Name: Mask

Number of images: 989

https://drive.google.com/file/d/1kJSAEH2SGALlzTI0O9qmJPar64G9fInU/view

6- Details of algorithm:

Flowchart Diagram

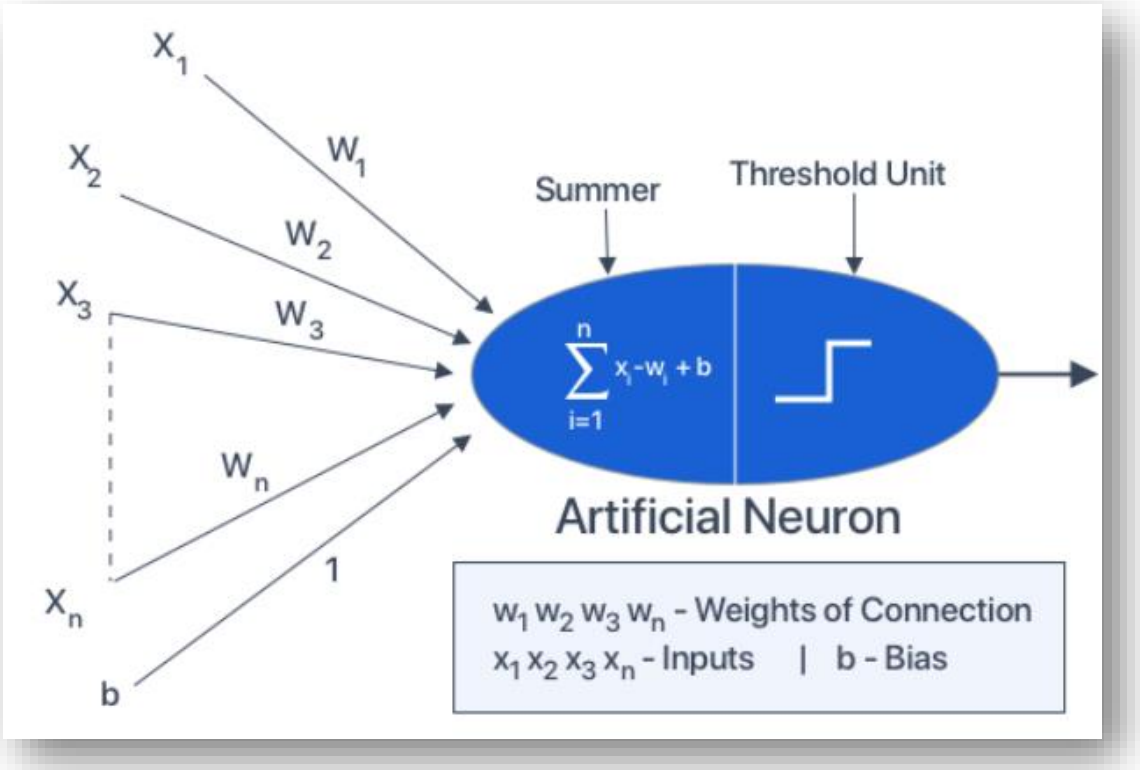


These are computational models and inspire by the human brain. Many of the recent advancements have been made in the field of Artificial Intelligence, including Voice Recognition, Image Recognition, Robotics using it. They are the biologically inspired simulations performed on the computer to perform certain specific tasks like -

- Clustering
- Classification
- Pattern Recognition

In general - It is a biologically inspired network of artificial neurons configured to perform specific tasks. These biological methods of computing are known as the next major advancement in the Computing Industry.

How does it work?



- It can be viewed as weighted directed graphs in which artificial neurons are nodes and directed edges with weights are connections between neuron outputs and neuron inputs.
- The Artificial Neural Network receives information from the external world in pattern and image in vector form. These inputs are designated by the notation x(n) for n number of inputs.
- Each input is multiplied by its corresponding weights. Weights are the information used by the neural network to solve a problem. Typically, weight represents the strength of the interconnection between neurons inside the Neural Network.
- The weighted inputs are all summed up inside the computing unit (artificial neuron). In case the weighted sum is zero, bias is added to make the output not- zero or to scale up the system response. Bias has the weight and input always equal to '1'.
- The sum corresponds to any numerical value ranging from 0 to infinity. To limit the response to arrive at the desired value, the threshold value is set up. For this, the sum is forward through an activation function.
- The activation function is set to the transfer function to get the desired output. There are linear as well as the nonlinear activation function.

What are the commonly used activation functions?

Some of the commonly used activation function is - binary, sigmoidal (linear) and tan hyperbolic sigmoidal functions(nonlinear).

- **Binary -** The output has only two values, either 0 or 1. For this, the threshold value is set up. If the net weighted input is greater than 1, the output is assumed as one otherwise zero.
- **Sigmoidal Hyperbolic -** This function has an 'S' shaped curve. Here the tan hyperbolic function is used to approximate output from net input.

Learning Techniques:

The neural network learns by adjusting its weights and bias (threshold) iteratively to yield the desired output. These are also called free parameters. For learning to take place, the Neural Network is trained first. The training is performed using a defined set of rules, also known as the learning algorithm.

Training Algorithms

- Gradient Descent Algorithm

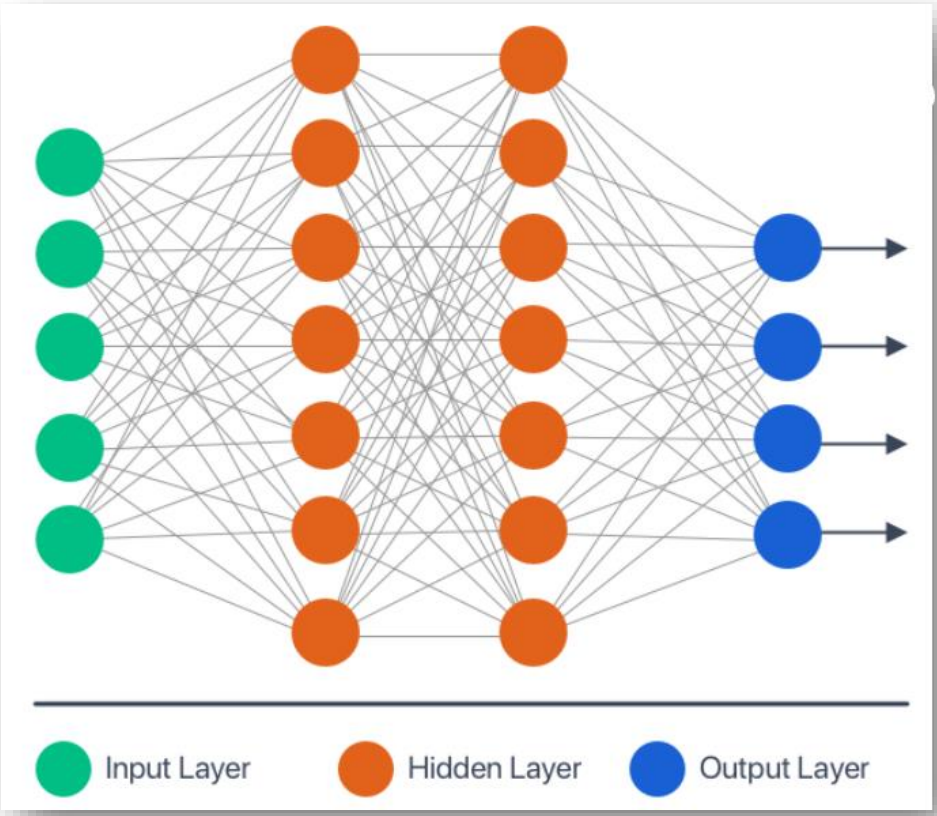
This is the simplest training algorithm used in the case of a supervised training model. In case the actual output is different from the target output, the difference or error is find out. The gradient descent algorithm changes the weights of the network in such a manner to minimize this mistake.

- Back Propagation Algorithm

It is an extension of the gradient-based delta learning rule. Here, after finding an error (the difference between desired and target), the error is propagated backward from the output layer to the input layer via the hidden layer. It is used in the case of Multi-layer Neural Network.

What is the architecture of it?

A typical Neural Network contains many artificial neurons called units arranged in a series of layers. In typical Artificial Neural Network comprises different layers -



- **Input layer** - It contains those units (Artificial Neurons) which receive input from the outside world on which the network will learn, recognize about, or otherwise process.
- **Output layer** - It contains units that respond to the information about how it learns any task.
- **Hidden layer** - These units are in between input and output layers. The hidden layer's job is to transform the input into something that the output unit can use somehow.

What are the Learning Techniques in Neural Networks?
Here is a list of Learning Techniques

- Supervised Learning:

In this learning, the training data is input to the network, and the desired output is known weights are adjusted until production yields desired value.

- Unsupervised Learning:

Use the input data to train the network whose output is known. The network classifies the input data and adjusts the weight by feature extraction in input data.

- Reinforcement Learning:

Here, the output value is unknown, but the network provides feedback on whether the output is right or wrong. It is Semi-Supervised Learning.

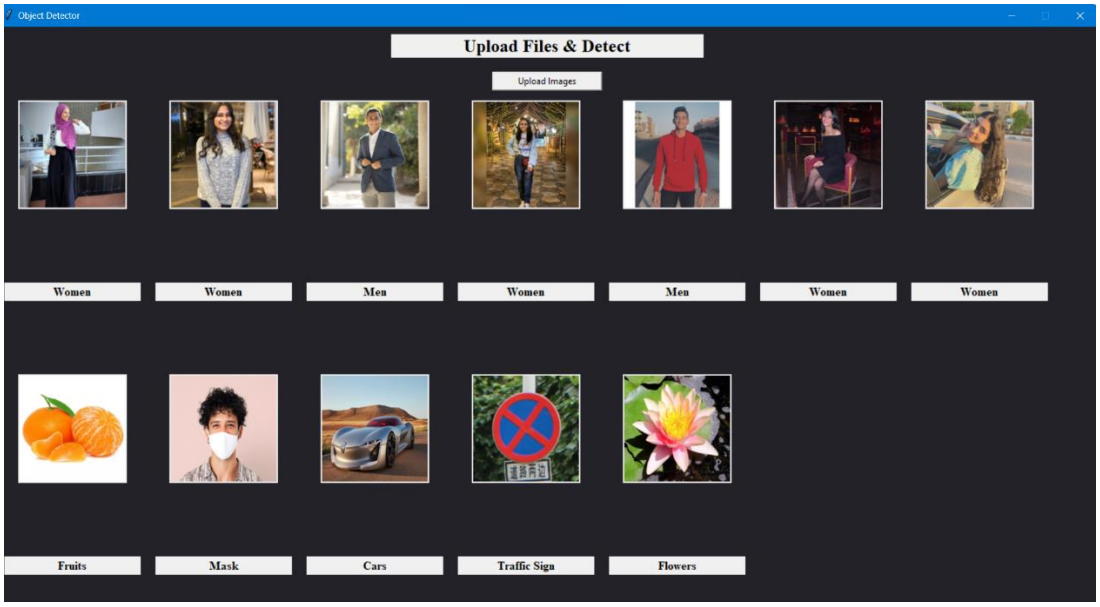
- Offline Learning:

The weight vector adjustment and threshold adjustment are made only after the training set is shown to the network. It is also called Batch Learning.

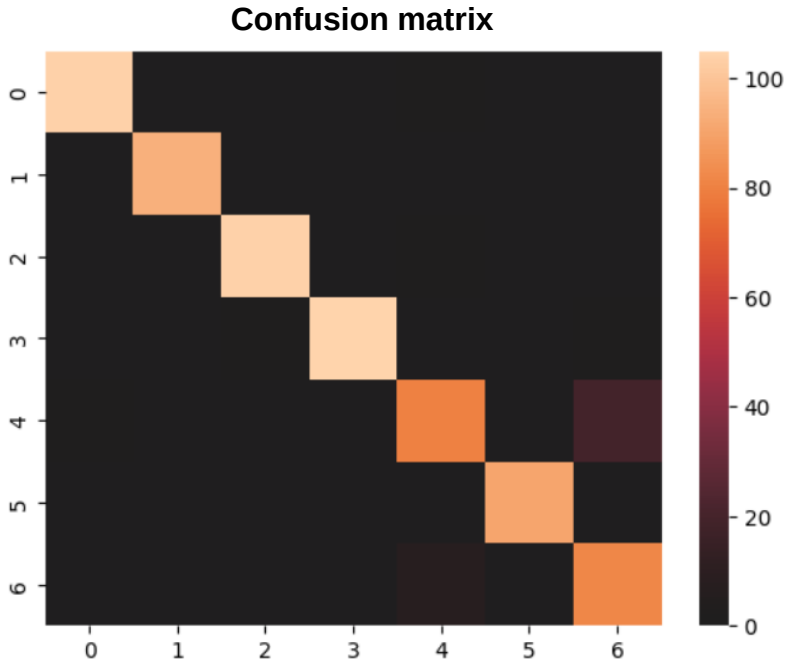
- Online Learning:

The adjustment of the weight and threshold is made after presenting each training sample to the network.

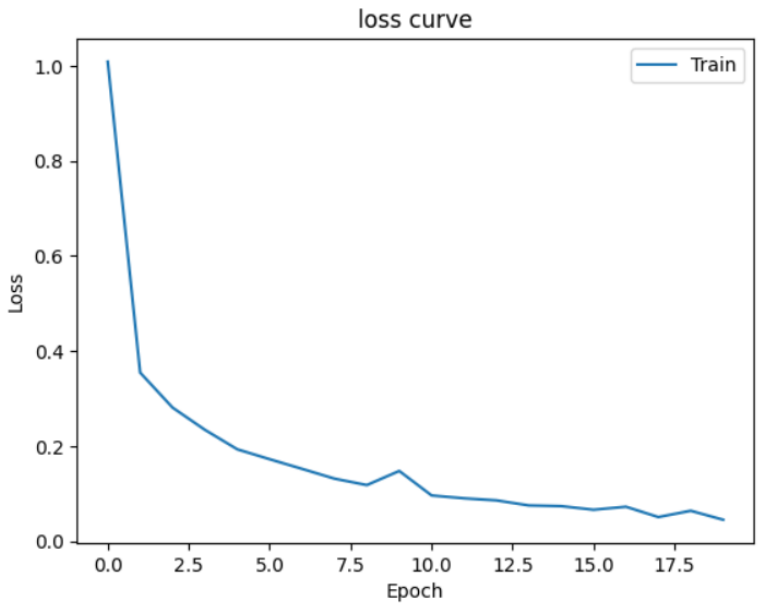
OUTPUT TEST FROM OUR PROJECT



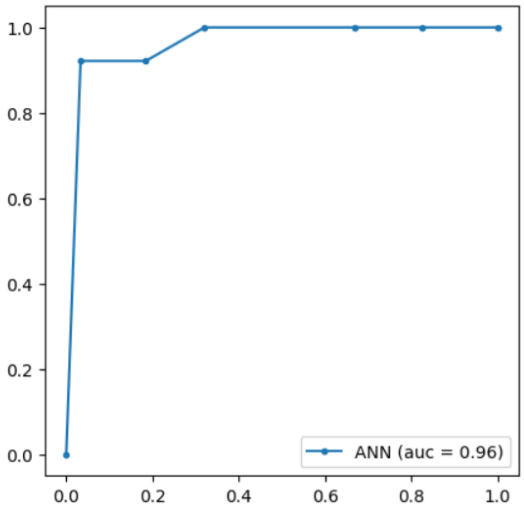
Some plots FROM OUR PROJECT



Loss curve



Roc curve



7- Development platform:

Tools:

Anaconda (Spyder, Jupyter)

Libraries:

OpenCV (cv2)
os
numpy
sklearn (sklearn.ensemble , sklearn.model_selection)
tensorflow (tensorflow.keras.utils)
keras (keras.application.vgg16)
tkinter (tkinter.filedialog)
PIL

Programming Language:

python