Overcoming Security issues using OpenCV and Machine learning

Aditi[1], Vivek Kumar Jha[2], Saikat Maity[3]

[1] Student, Departmentof Computer Science and Engineering Technology, Dr.B.C.ROY Engineering College Durgapur, West Bengal, India

[2]Student, Departmentof Computer Science and Engineering Technology, Dr.B.C.ROY Engineering College Durgapur, West Bengal, India

[3] Associate Professor, Departmentof Information Technology, Dr.B.C.ROY Engineering College Durgapur, West Bengal, India

Abstract

As computer vision has become a part of our daily life. Using smartphones features like face recognition is all a part of computer vision. In this paper, we will talk aboutCV or Computer Vision which is preferably used to read and display a video stream in real time through which one can access the web camera and using machine learning one can let their system learn through data sent by user or through datasets which is easily available on the internet and then system trains itself. After training part it is ready to solve the real life problems. But using both computer vision and machine learning at the same time is always a challenging task asone has to capture and another has to train the system at the same timeso that the system could be able to recognize the things to which they are trained is the most innovative work in this paper as we have to keep in mind that both the things should be done at the same time.

Keywords — *Computer Vision, machine learning, smartphone.*

I. INTRODUCTION

In this era, where camera has become an important part of our life. This project mainly describes the use of computer vision for providing surveillance, security in very convenient way. Most of the people cannot afford the high cost of model which will give security to their important documents or anything which is important. This project aims to create a model through computer vision which will provide more security than the old ones. In simple words, computer vision is to provide vision i.e., eye to a computer. Using computer vision we are able to detect faces, eyes, nose and different body parts through haarcascade classifiers and hog. But I prefer to use haarcascade. But do you think about using more about machine learning and computer vision. Machine learning where a computer or a machine learns from the past data and used to predict or forecast future data. However machine learning is generally used to predict the future result which is known as dependent variable. And the past data which is known as an independent variable. The name dependent and independent is used because independent means which is not dependent. As we get past dataset which is easily available on the internet or one can make its own dataset and the predicted result depend on the past dataset that's why it is known as dependent variable. The machine learning is subdivided into three categories i.e., supervised, unsupervised and reinforcement learning. Using different algorithms of machine learning which will get best accuracy can be used as per the dataset or as per the model and users. The dataset which is generally used to train the model is easily available in UCI, Kaggle and many more. Actually, I prefer to use only two sites which I mentioned earlier. Machine learning is generally used to train 80% of data which is present in dataset. And test by using remaining 20% of data present in dataset.

However generally machine learning is used to train the data to predict stock prediction or disease prediction or something which can be predicted through past data. But using machine learning algorithm to predict the faces of humans by training with their images is really a big deal. This project aims to do so and later we will try to use third party to deliver some messages or giving missed call if possible. Currently we are working on this.

Our main concern is to train our system in such a way that model first opts to click a pic of yours for training purpose, so generally it will be better if we try to click a picture using different facial expressions so that it could address you more easily. And after that, it will train itself with a pic of yours and recognize you every time when you open your web cam. However this is easy anyone can do this. The big deal is to recognize others as intruders or unknown persons from the known ones i.e., from those whose images were added in the dataset, unless you introduce others to the web cam. This is the first motto of our paper the rest will be revealed later.

In this paperwork we will see how machine learning and computer vision works and how it works when implemented both one at a time.
How we are able to capture picture through web cam using computer vision and all these things we will discuss in detail. Surely you will love this project we can assure you this thing. Currently we are working with the web cam which is generally in laptops.

This document is a template. An electronic copy can be downloaded from the conference website. For questions on paper guidelines, please contact the conference publications committee as indicated on the conference website. Information about final paper submission is available from the conference website.

II. BACKGROUND

Here, we will discuss about all the things which we will use in our project in detail. Starting with the computer vision part first then we come into machine learning and later which algorithm best fits in this scenario, then it's applications and at last the conclusion part.

A. COMPUTER VISION

Computer vision has become a part of today's era as we generally use our face as a phone lock designed by some companies. But the basic thing is how to use all this [4]:

Like my friend did. The below figure Fig.1 shows detection of human face and eye

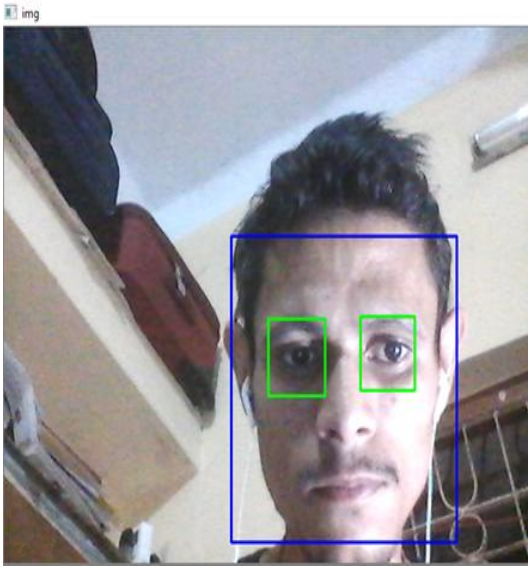


Fig.1: Face and eye detection

The above figure use simple haarcascade classifier to train our system so that it can easily detect face and eyes.

Computer vision[1] is used to read an image and display an image with open CV.
Things one should install in python before using any sort of detection:

- Numpy
- Cv2

To install Numpy in python use this command in cmd (command prompt): pip install Numpy
To install cv2 in python use this command in cmd (command prompt): pip install cv2
Numpy is used to support the open CV functionality. Generally Numpy is used for numerical calculations. And open cv-python i.e., cv2 is a library of python used to give vision to computer. Mainly there are two types of classifiers first one is hog (Histogram of oriented gradients) and haarcascade classifier. But we generally prefer to use haarcascade classifier. Haarcascade is generally used to detect a particular object or body parts for which it is trained. There are different classifiers of haarcascade like frontal face which is trained to detect front face of a human body, left eye which is trained to detect left eye of a human's body, right eye which is trained to detect right eye of a human's eye and there are many classifiers like this which is trained to detect some specific things.
First we have to import Numpy as it helps to represent images in a multidimensional array then we will import cv2 which is a library of python [10]. We use while (True) because we want to run the program infinitely so that it could capture video stream in real time without any disturbance and start taking images until waitkey is used. Actually this is used to capture real time video without stopping. Then to use face detection we preferred to convert BGR to gray to reduce noise and to convert 3dimensional array to 2 dimensional arrays. Then to read an image, we use imread which means image read which is generally stored in n dimensional array. For example:
Image = cv2.imread ('File path')
As we does not use semicolon in python like c. Then to display an image we read before, we use imshow function:
cv2.imshow ('Image name', image)
Generally, we use waitkey function to wait for a few milliseconds this is used so that user's input can be taken from keyboard. Then we have to release the camera so that camera is able to read and display a real time video stream. At last we have to destroy all windows and the only active thing in laptop is the running real time video.
If one has to do eye detection then he/ she must have to useROI [5] also known as region of interest. So that system will detect only those things which are a region of interest like eyes.
This is the basic of computer vision.

But do you think how good it will be if we use both computer vision and machine learning.
One helps to display a realtime image while other is busy in training and testing of an image. If both used one at a time then we will be on another level and might give the best security model.

B. MACHINE LEARNING

Machine learning[2] is generally used to train our machines. Through past data machines learn all by itself. The past data are known as an independent variable while new data or the data which is to be forecasted is known as dependent variable. Machine learning generally learns from the past data and trains itself from the past data and predicts the result for the future one. Prediction is generally used so that we will be able to know the future values in present. For example let's say you have to do a weather forecast on 5th of April then in past when machine learning came into being existence, we collected all the data on 5th April for at least 5-10 years and then predict it. But now everything changed, we do not need to take or include the data of 5-10 years, one can easily predict the weather by taking dataset of 5-10 days. Generally, machine learning algorithm, if applies to a dataset, then it will train 80% of its dataset and test with the 20% i.e., remaining of the dataset.

Machine learning is broadly classified into three parts:

a. Supervised learning
b. Unsupervised learning
c. Reinforcement learning

It is divided on the basis of pattern of their learning. Generally, people focus on machine learning because we do not have to write full code what to do, how to do, why to do, when to do computers learn all by itself. Now come on each types of machine learning.

a. SUPERVISED LEARNING

The meaning of word supervised means first observe or watch then do the work, or it simply means first learn how everything was going then predict what will be all by your own. Supervised learning means that humans have to feed all the data to the computers and when computer understand what's going on then it will try to solve it for you. Just like teacher- student relationship. Under supervised learning or supervision learning, there are many types of algorithms which first observe each and every data in the dataset then process it. The mostly used are two—

a) Regression
b) Classification

1. Regression

Regression[3] is a method from statistics which is used to train from the input data, input means the data which we will give to the system and predict in the form of output. In this, when user which acts as a teacher who gives all the past data which is known as independent variable and computer or a machine with whom user is working first observe all the past data or independent variable and after then it is ready to predict some output with the help of past data. The most important thing while using any algorithm is to check it is accuracy. Accuracy tells how much your algorithm is accurate to predict some valuable results. When your model's accuracy is more than 95% then your model is somehow worthy.
Under regression, there are many types like linear regression, logistic regression, and quadratic regression and so on.

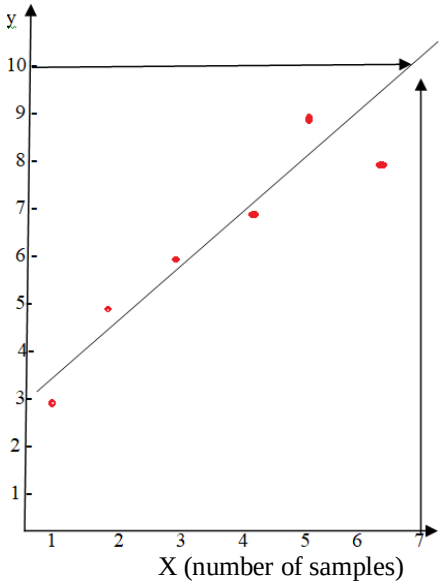


Fig.2: Plotting of table1

The coordinates of points is given in Table 1 where x represents the number of samples while y represents the cost per sample. But, here in sample number 7, cost per sample is not known and if someone wants to predict the cost per sample of sample number 7, the simple linear regression is best.

TABLE 1

X(number of samples)	Y(cost per sample)
1	3
2	5
3	6
4	7
5	9
6	8
7	?

Here some values are given in Table 1 through which we have to find the value of Y coordinate when x coordinate is 7. Through regression, we draw a line so that all the points fall near to the line. Based on the value we can predict the value of Y coordinate when x=7.
So through prediction by plotting graph as shown in Fig.2, we find that approximate value of y=10 when x=7. This is known as simple regression.

The above figure Fig.2 illustrates the linear regression model. The red line is the predicted value. The linear regression model is based on formulae:

$$Y = a * x + b \qquad (1)$$

Here y is a dependent variable as it depends on the value of x. x is an independent variable. b is the slope of the line and a is the y intercept. But if one wants to do hand calculations then there are certain formulae to find the value of a and b.

Actually, there is one more feature i.e., matplotlib which is used to do plotting in 2-D or 3-D.
If one wants to use matplotlib then he/she has to install matplotlib through cmd (command prompt): pip install matplotlib.

If one wants to use matplotlib to plot graph then he/she has to import matplotlib by writing this command: import matplotlib.pyplot as plt.
If one wants to show the graph then he/she has to show the graph by using this command: plt.show()
One can either add coordinates i.e., X or Y or use scatter points to plot only those coordinates not the whole line.

The result is generally close to the data given by user from where the machine learns to produce result with higher accuracy.
In linear regression model, linear graph is made while in case of quadratic regression model, the graph would be like:

TABLE 2

X(number of samples)	Y(cost per samples)
1	8
2	7
3	5
4	2
5	4
6	7
7	?

Here x represents the number of samples while y represents the cost per sample as shown in Table 2. The graph is drawn in Fig.3.. But, here in sample number 7, cost per sample is not known and if someone wants to predict the cost per sample of sample number 7

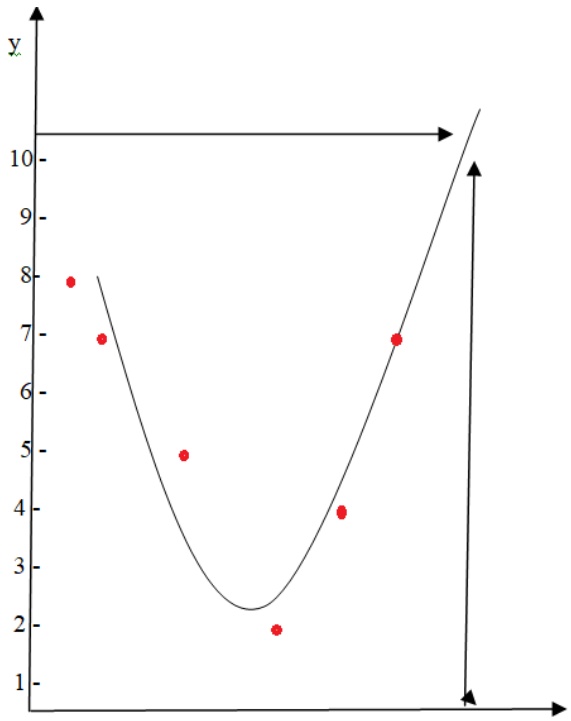


Fig.3: Plotting of table 2

Here, the above table 2shows some points but the value of y coordinates which represents cost per sample is missing or which we have to predict. We can find this by quadratic regression if simple regression or linear regression fails. Here, in the graph we draw in Fig.3 quadratic curve where the value of previous i.e., from x=1 to 6 or y=1 to 6, we have to predict the value of y when x=7. From this graph, we find that y=10. This is known as a quadratic regression.

The above graph in Fig.3 shows quadratic regression.
One generally use regression to predict or forecast the results which is directly dependent on input data through which computer understands something i.e., what's going. Just like prediction of estates, stock prediction and so on.

2. Classification

The word classification means to classify objects or data under a certain group. In classification, user's first gives the dataset or the input data through which computer trains itself. It is also a part of

supervised learning. The output generally produce classifies itself through input data sent by users or label the output that under which category it falls. There are several types of algorithms which come under classification like svm (support vector machine),knn(k nearest neighbor) etc...

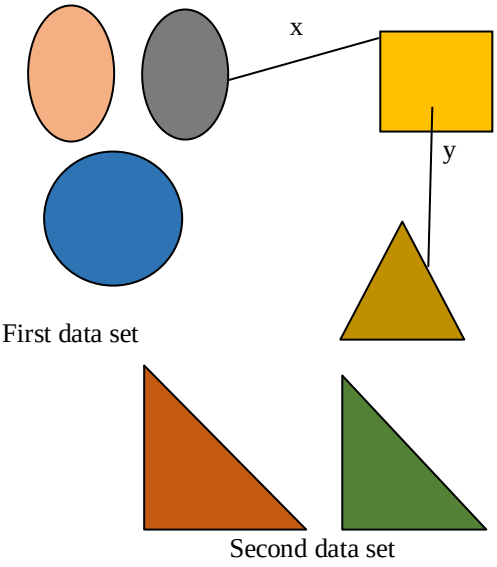


Fig.4: K-NearestNeighbor Example

In the above figure Fig.4, there are two different objects first one is oval in shape and second one is triangular in shape having different size, we have to predict the rectangle object belongs to which object category one which is oval or one which is triangular. In knn(K-nearest neighbor), we measure the distance of rectangle from one of the object whose shape is oval say the distance is x and the distance of rectangle from one of the object whose shape is triangular say the distance is y.
If x>y then rectangle object belongs to the triangular object category otherwise it belongs to oval object category. This is how k nearest neighbor works, whose distance is nearest to the data whose testing part is going, will fall under that category.
The distance which is minimum from the value which we will test to the input data, under that category the testing value belongs.

However, support vector machine is interested to find hyper planes.

This is all about supervised learning where we first have to let them learn what's going on, how going on and then computer decides the rest all by itself there is no need to mention or give the instructions to the computer like how to do it. Computer does all by itself. This is the great thing about machine learning that's why it has a different craze. While supervised learning comes under labels or we can say come under a category. The problems solved with the help of supervised learning are the problem under supervision of the user not on the computer. But unsupervised learning is the opposite of supervised learning.

b. Unsupervised Learning

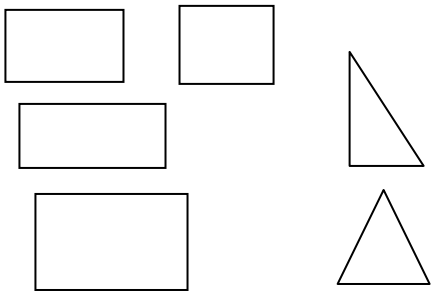
The word unsupervised means which is not under supervision. Like supervision, it does not have output variable. Or we can say it does not have teacher. Computer has to decide all by itself. There is no need to train computer. It is the opposite of supervised learning. The computer decides by itself under which category the testing value will fall. The user gives unlabeled data generally to understand more about data. Generally, unsupervised learning is used to those models in which user himself/herself do not have enough idea or information about the datasets which is used by the computer to decide. Computer decides all by itself under some category by looking into their color, shape and all this. It is broadly divided into two types:

a) Clustering
b) Association

1. Clustering

In Clustering [6], different data input is divided in the form of clusters.It is unsupervised learning, where users do not have enough knowledge about the data set. System learns everything from itself about the data and put the data in the order for example put the same data in the one category and form a cluster which is denoted by k. If you want the whole data is arranged in two parts then use k = 2, if you want in 3 parts then k=3 and so on.

Let's say we have 6 different data here consider data as objects i.e.,4 rectangle and 2 triangle and you want whole data to be divided into 2 parts as shown in Fig.5 then use k = 2



First data set **Second data set**

Fig.5: 2 cluster

Here all the data are arranged according to the shape and as per user choice it is divided into two parts.

2. Association

The word association[7] means forming some groups on the basis of something which maybe anything. Association in machine learning finds the relation between independent variable. Many datasets given by the user, now the work of machine is to understand the datasets and then predict the actions of testing value. Generally used in some sort of recommendation algorithm.
For example, say that A orders three types of books based on genre say 'horror', 'comedy', 'thriller' and B orders three types of books say 'horror', 'comedy', 'romance'. Now if C comes to order some books like 'horror', then machine will predict and recommend her to buy comedy.
This is a part of unsupervised learning, where no labeled input data are feed, no needed to train or give supervision of computer. The computer understands all by its own and then guide or tells the users.

C. Reinforcement Learning

This learning [7] is different from supervised learning and however, there is no need to train or something supervision like that. Computer finds itself what to do and learns everything from the past experience. In this the computer or machine under which this algorithm is going to use decides how to reach to an output through experience and reward itself when it reaches to the output. Machine or computer chooses the best path from itself. There is a chance that computer chooses a wrong path or complex path which takes more time and learns from its own mistakes. This learning is not like the old one i.e., supervised learning where first computer o machine observes and train itself from the past data or input data then predict the output. Here, there is no need to do anything just give a sample of input data then computer decides all by itself which path it should chose to reach to the reward i.e., output in less amount of time.

This is all the types of machine learning and through looking at the problems, we estimate that under which real life problem, we should choose which types of algorithms to produce best results with highest accuracy. The decision of which algorithms to use depends upon the user not on the machines or computers.

Algorithm

The biggest dilemma is to decide which algorithms should we use to train the pictures and let our computers or machines recognize the people all by its own. All of these come under the classification algorithm in supervised learning where we create labels according to the same types of data input so that computer can easily recognize peoples or anything for which it is train.

The first basic thing we should do is to create an empty list of labels to gives labels to each and every image of different objects so that we could easily store them. But how does it start from the beginning. If one selects to add a different set of input then how anyone can do it? Actually, here we are giving the inputs of data which is in the form of images. For taking images from the web camera of the laptop, the very basic thing one should do is to import cv2 which helps to read and display a video stream. And after reading the image from the web camera, we have to give labels to the image with which we are going to train our machine or our computer so that they recognize it more easily. After that, web camera captures the image in different angles so that in any angle it can recognize the image and tells the name as given in the label. Here, we have trained our computer to take up to 50 pictures of image in different angles. The user can add as many images he/she wants. Adding many images will increase the accuracy of this project. While capturing images, we have to destroy All Windows. After adding many images or a single image, user can now train the system. Adding images to train our computer or machine is an easy task. We just have to use cv2 library which helps one to access the web camera so that one could read the image. Adding labels to an image is like adding name to an image. Like we do in supervised learning, we give labels to the input where first computer observes all the input by taking some time and then learns from the input and then able to predict the real time things. Just like we do it to train with only 50 pictures per image. Anyone can use more than 50 pictures or less than 50 pictures to train a model. It all depends on the user choice who wants to train a model. After taking pictures the next thing is to train the model by one of the classification algorithm in machine learning. This is the biggest deal any user will face as the best part is to select which algorithm one should choose while using in a model, which will give best result or perfect result which accuracy will be the best, which result will match the result in the future if one is solving real time problem like stock prediction. All things should keep in one's mind before using the algorithm.

Algorithm used in this project:

Step 1. Make our own dataset by capturing images in real time where we have put a limit on the number of images to be captured by the system is 50.

Step 2. Give label to the dataset here we consider label as name.

Step3.Using knn (k nearest neighbor) of supervised learning where computer is trained to learn from the dataset, so after adding label and capturing images i.e., making our own dataset we have to train our model.

Step 4. After training the data through supervised learning, 80% of the data which is present in the dataset is used to train while remaining 20% of the data is used to test

.

Step 5. After training we will use it for testing whether it is able to recognize the persons.

Step 6. When used for testing purpose, it is used to recognize only those persons whose label i.e., name is added while for rest of the persons, system claims them as an intruder.

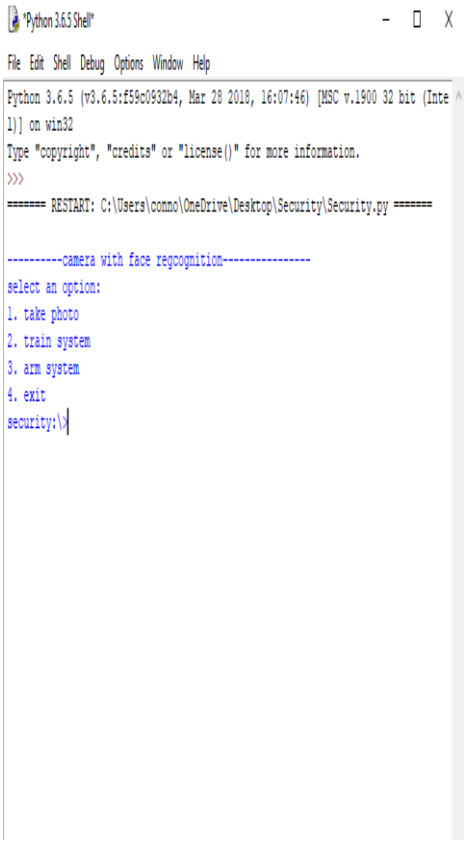


Fig.6: System asks for a few things

This model always asks the user whether he/she have to take a pic or not, whether he/she have to train the image or whether he/she wants to check whether a computer can recognize or not as shown in Fig.6. So capturing images will look like Fig. 7

This system is trained to capture a picture of 50 images as shown in Fig.7. One can increase the no of pictures taken by the system. It all depends on the user's choice. We are interested only in taking 50 pictures because this is a prototype.

After taking or capturing images, the next thing one should do is to train the image. Just like supervised learning do after observing the inputs, it trains itself in the same way after capturing the images computer or machine will train itself from the past data. Here, past data is the images. Computer will train itself to each and every image of a label.

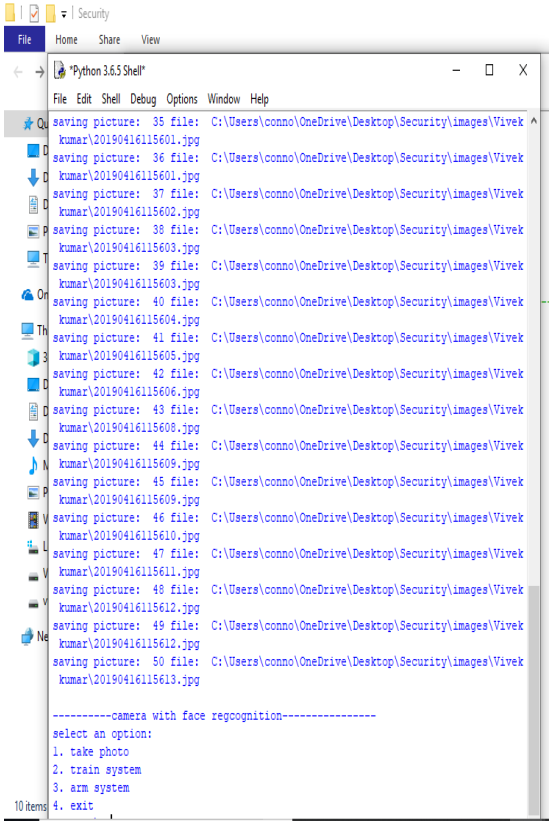


Fig.7: System is capturing the images

Machine learning generally trains data either on 80-20 or 70-30. This means that up to 80% of full data is used for training purpose and the rest 20% is used for testing purpose or 70% of the full data is used for training purpose and rest 30% is used for testing.

After training as shown in Fig.8 the next thing is to check how much a computer trains itself, through observing and training. Here, we will check

whether our system or computer is able to recognize itself or not. If it is able to recognize the path, then our model is successful. In this we have to set a path to the trained model, so that computer able to recognize each and every object which is trained. And this project is made to recognize a certain group of people, if another person comes in contact with the known person by the computer then computer name that person as an intruder.

After capturing the images and training the images, the next thing a user wants to see whether a computer is able to recognize a person or not. Well, we have tried to do the recognition part with only one person who is my team member in this project.

```
training the system to recognize faces
found images for  ['Vivek']
adding image  1  -  Vivek 1
adding image  1  -  Vivek 2
adding image  1  -  Vivek 3
adding image  1  -  Vivek 4
adding image  1  -  Vivek 5
adding image  1  -  Vivek 6
adding image  1  -  Vivek 7
adding image  1  -  Vivek 8
adding image  1  -  Vivek 9
adding image  1  -  Vivek 10
adding image  1  -  Vivek 11
adding image  1  -  Vivek 12
adding image  1  -  Vivek 13
adding image  1  -  Vivek 14
adding image  1  -  Vivek 15
adding image  1  -  Vivek 16
adding image  1  -  Vivek 17
adding image  1  -  Vivek 18
adding image  1  -  Vivek 19
adding image  1  -  Vivek 20
adding image  1  -  Vivek 21
adding image  1  -  Vivek 22
adding image  1  -  Vivek 23
adding image  1  -  Vivek 24
adding image  1  -  Vivek 25
adding image  1  -  Vivek 26
adding image  1  -  Vivek 27
adding image  1  -  Vivek 28
adding image  1  -  Vivek 29
adding image  1  -  Vivek 30
adding image  1  -  Vivek 31
adding image  1  -  Vivek 32
adding image  1  -  Vivek 33
adding image  1  -  Vivek 34
adding image  1  -  Vivek 35
adding image  1  -  Vivek 36
```

Fig.8: Training our system with the data taken in real time

In Fig.9 below system is trained or let's says in other words, Vivek is a known person for the system. Now let's see what system will call those who are unknown to the system.

As we have added label Vivek name of my project friend. So whenever system sees him, it will recognize him via name.

FONT_HERSHEY_SIMPLEX [9] is used to write something which means some text can be added in the image. We are training our system so that it will be able to recognize person via name.

Neither my (Aditi) images were trained nor captured in the database. The computer recognizes him by his name in the same way as computer should do with the known persons and me by the name with which computer is supposed to know the name of the unknown persons. In this model, Vivek is the known person for the system and I am the unknown one or intruder as the name suggests by the computer for me.

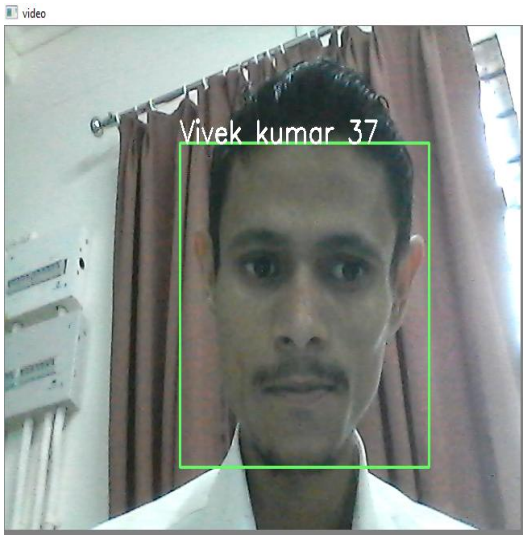
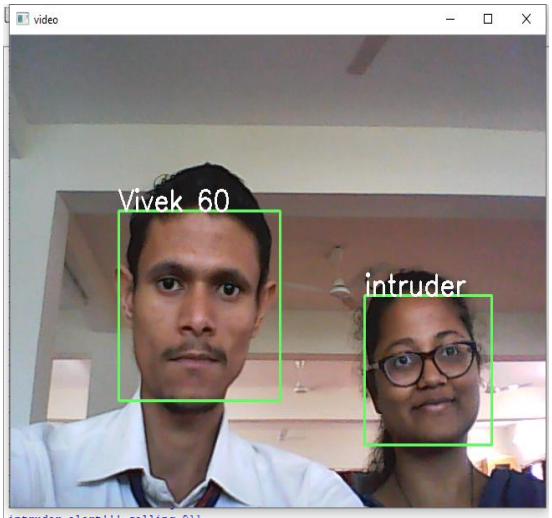


Fig.9: System is recognizing Vivek as system is trained to recognize by seeing vivek's face.




```
intruder alert!!! calling 911
intruder alert!!! calling 911
intruder alert!!! calling 911
intruder alert!!! calling 911
intruder alert!!! calling 911
intruder alert!!! calling 911
intruder alert!!! calling 911
intruder alert!!! calling 911
intruder alert!!! calling 911
intruder alert!!! calling 911
intruder alert!!! calling 911
intruder alert!!! calling 911
```

Fig.10: System is recognizing known and unknown persons.

The next thing to test is to whether he can recognize the unknown persons or not. We have trained "intruder" word to those whom computer or system becomes unable to recognize. In this model, computer is trained only to recognize Vivek (team member) and not me as shown above in Fig.10.
This model is trained to recognize and detection of face only, one can use it to recognize face and eye both. To detect eye, one has to use ROI(Region of interest) [1].

Applications

The biggest use of this project is to use this model in high confidential areas, where a few are allowed to access. Instead of using humans, we can use this model we can also save electricity or its power by using a sensor which can detect the human presence. Then camera will turn on when the sensor is able to detect the human and recognize whether the system know this person or not. If yes, then he/she is allowed to enter to that confidential area or if not then the system will give a call or give message to a security member whose number is saved on the system as the user want to use the system, the rest depends on user.

One can also use to train the system through a weapon dataset through system can easily recognize the weapons and give a call to the third party. We are currently working on the training part with weapons dataset. This project is highly used for security reasons. Recognizing the image through one of the best algorithms i.e., classification under supervised learning.

CONCLUSION

However this is a prototype for this system, but if we able to build the real model using night vision camera,human presence sensor to reduce power consumption, this will be the best thing for the security of our county's confidential things.

ACKNOWLEDGMENT

We, Aditi and Vivek Kumar Jha like to say thank you to Saikat Maity to give us the opportunity to make this project. Without his guidance, we are unable to complete this project. Without his motivation, that we can do this boost our confidence and helps us.

REFERENCES

[1]. https://www.pythonprogramming.net/loading-images-python-opencv-tutorial/
[2]. https://www.simplilearn.com/big-data-and-analytics/machine-learning-certification-training-course
[3]. https://www.geeksforgeeks.org/machine-learning/
[4]. https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html
[5]. https://realpython.com/face-detection-in-python-using-a-webcam/
[6]. https://www.expertsystem.com/machine-learning-definition/
[7]. https://towardsdatascience.com/machine-learning/home
[8]. https://www.edureka.co/masters-program/machine-learning-engineer-training
[9]. https://www.pytorials.com/face-recognition-using-opencv-part-3/
[10]. https://www.superdatascience.com/blogs/opencv-face-recognition
[11]. Computer Vision :A Modern Approach by Forsyth and Ponce
[12]. Introduction to Machine Learning using Python by Andreas C. Muller and Sarah Guido.
[13]. S. Maity, J. Sil, "Color Image Segmentation using type-2 fuzzy sets" – International Journal of Computer and Electrical Engineering, Vol. 1, No. 3, August 2009 1793-8163 (2009) pp 376 – 383.
[14]. S. Maity, J. Sil, "Denoising of Images using fuzzy Rule base and Clustering Approach"- CSIBIG 2014, DOI: 10.1109/CSIBIG.2014.7056994IEEE Explore, pp 1 – 7.
[15]. S. Maity, J. Sil, "CMYK model Color Image Segmentation using Type 2 Fuzzy Sets"-6th International IEEE Conference on Next Generation Web Services Practices (NWeSP 2010). IEEE Explore pp. 347-352