AN APPROACH OF FACE DETECTION WITH CAP AND MASK

BY

NAME: MD. SOLIMAN ALI
ID: 171-15-8572
NAME: MOBASSERA ASMA SADIA
ID: 172-15-1599
AND
NAME: TABASSUM TANHA
ID: 171-15-1282

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By
MS. TANIA KHATUN
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised By
MR. MOHAMMAD JAHANGIR ALAM
Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

APRIL 2021

APPROVAL

This Project titled "**AN APPROACH OF FACE DETECTION WITH CAP AND MASK**", submitted by Md. Soliman Ali, Mobassera Asma Sadia and Tabassum Tanha to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on April.

BOARD OF EXAMINERS

____

(Name) **Chairman**
Designation
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

______________________ **Internal Examiner**

(Name)
Designation
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

External Examiner

____

(Name)
Designation
Department of -------
Name of university

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Ms. Tania Khatun, Senior Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

____

Ms. Tania Khatun
Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

____

Mr. Mohammad Jahangir Alam
Lecturer
Department of CSE
Daffodil International University

Submitted by:

____

Md. Soliman Ali
ID: 171-15-8572
Department of CSE
Daffodil International University

____

Mobassera Asma Sadia
ID: 172-15-1599
Department of CSE
Daffodil International University

Tabassum Tanha
ID: 171-15-1282
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project successfully.

We really grateful and wish our profound our indebtedness to **Ms. Tania Khatun**, **Senior Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Deep Learning*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft, and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to Head**,** Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Face detection is a technology that is broadly implemented and studied on deep learning with hugely improved face detection system indicators. Extensive research has been done to detect faces under illumination changes, and the resulting degenerative image problems are largely ignored. Maximum existing face detection methods focus on overcoming the problem of facial obstruction due to sunglasses. As far we know, no research has been done on obstruction due to cap and mask. Yet this problem should emphasize because it is known that bank criminals adopt this process to hide their faces. Our system offers purposes for solving this problem. Face detection with cap and mask applying Deep

Learning is a highly challenging job. Using the TensorFlow Object Detection select for dataset training in the deep learning model. Specifically, we are going to use a pre-trained model called SSD-MobileNet. This method applies a camera, which captures images. First, Prepare the image labels using Labelimg. Then, Transfer learning using SSD MobileNet. Finally, Detection using a webcam and OpenCV and display the text and percentage. The implementation of the process is by using OpenCVPython. The technique applies to various libraries.

Keywords— Face detection, Object detection, Tensorflow's Object Detection API, TensorFlow, Dataset, Deep learning, SSD MobileNet, Security management.

TABLE OF CONTENTS

CONTENTS	PAGE
Board of examiners	i
Declaration	ii-iii
Acknowledgements	iv

Abstract	v
CHAPTER	**1-26**
CHAPTER 1: INTRODUCTION	**1-3**
1.1 Introduction	1-2
1.2 Motivation	2
1.3 Objective	2
1.4 Expected Outcome	2
1.5 Report Layout	3
CHAPTER 2: BACKGROUND	**4-7**
2.1 Introduction	4
2.2 Related works	4-7
2.3 Research summary	7
2.4 Challenges	7
CHAPTER 3: RESEARCH METHODOLOGY	**8-11**
3.1 Introduction	8-10
3.2 Proposed Model	11
CHAPTER 4: MATERIALS, EXPERIMENT RESULT AND DISCUSSION	**12-21**
4.1 Data Set	12
4.1.1 Data labeling	12-13
4.1.2 Dataset Structure	13
4.2 Experiment Results and Analysis	14-18
4.2.1 Result Analysis	19-20
4.3 Discussion	21

CHAPTER 5: EFFECT ON SOCIETY, ENVIRONMENT AND WORKABILITY	**22-23**
5.1 Effect on Society	22
5.2 Effect on Environment	22
5.3 Ethical Views	23
5.4 Workability Plan	23
CHAPTER 6: OVERVIEW, CONCLUSION AND FURTHER RESEARCH	**24**
6.1 Overview of the Study	24
6.2 Conclusions	24
6.3 Further Research	24
REFERENCES	**25-26**

LIST OF FIGURES

FIGURES	PAGE NO
Figure 3.1: Proposed Model	11
Figure 4.1: Labeling image	13
Figure 4.2: XML file of labeling image	13
Figure 4.3: Localization Loss	17

Figure 4.4: Classification Loss	17
Figure 4.5: Normalized total Loss	18
Figure 4.6: Learning Rate	18

LIST OF TABLES

TABLES	PAGE NO
Table 2.1. Literature Review	4-6
Table 4.1. The number of images per dataset	12

CHAPTER 1
INTRODUCTION

1.1 Introduction

Nowadays, object detection is one of the most significant leading research areas in deep learning. Here this technology predominantly identifies objects relating to particular labels inside an image or even a video stream. It does not confuse image labels, a different computer vision branch that can easily predict what an image contains. We could technically apply this technology for face detection with cap and mask. This technology has various real-world applications like video surveillance or image retrieval, with facial detection being one of the most used fields. We can also use it in Sign Language Detection (SLD) for Communicating with deaf and dumb people in our community.

In real-time, the Single Shot Multibox Detector (SSD) is a single convolution neural network mainly designed for object detection. SSD is a famous algorithm in object detection. It is generally faster than Faster RCNN. It learns to predict bounding box locations surrounding the thing. It labels the thing in just a one-shot instead of R-CNNs that apply a region proposal network to create the bounding boxes that are later used to classify objects. Therefore, this model is often trained to finish and consists of MobileNet architecture, followed by many other convolution layers.

Face-related occlusions, which might also happen for harmless behaviors, are often related to various severe security issues. In the Global pandemic, COVID-19 circumstances appeared dangerous influenza in the whole world. Center this epidemic situation many crimes have increased. Many of us who commit crimes wear hats and masks so that nobody can identify them. For example, ATM criminals tend to wear caps, sunglasses, and masks to prevent their faces from being detected. Due to the viewing angle problem, a close-distance face captured by a CCTV camera is often impeded by the visor, whereas a far-distance look is not hampered, but it has a low resolution. The occlusion generated by a cap with a brim is therefore varying along with people moving. The trade-off between occlusion and picture quality collected face pictures is usually undefined from the recorded videos. Thus, it is imperative to equip automatic cap detection systems in entrance surveillance, detect suspicious persons and provide prior knowledge of occlusion to improve criminals' face detection. However, according to many recent police reports, shop thieves, bank robbers usually wear a cap and mask when entering the places where they commit crimes. If we follow our model, then we could lower the number of crimes related to ATM Booth.

This project focused on face detection with cap and mask. Using the TensorFlow Object Detection select for dataset training in the deep learning model. Specifically, we are going to use a pre-trained model called SSD-MobileNet.

1.2 Motivation

Face detection technologies have studied over the past decades. Deep learning procedures can leverage rich datasets of faces and learn expensive and miniature representations of faces, permitting advanced models, to begin with, to perform as well and afterward to defeat people's face detection abilities. The approach of Deep learning has developed so rapidly still sometimes to find any criminals. We want to build a model that can quickly detect face with cap and mask. For building a deep learning network that solves object detection problems, we use the TensorFlow object detection API framework. For that, we use an automated detection approach with the deep learning models.

1.3 Objective

- Detect face of a person
- Detect cap and without cap
- Detect mask and without mask
- Prevent crime
- Make society safe to live

1.4 Expected outcome

- Labelling Images for face detection with cap and mask
- We will train the deep learning model using TensorFlow Object detection API, and specifically we are going to use a pre-trained model called SSD-MobileNet
- Finally find the accuracy of the proposed algorithm
- At last, we will easily be able to detect the face with cap and mask through this application

1.5 Report Layout

This paper is organized as follows, In chapter one is the project title is elaborated and explained in the introduction part. This section includes the introduction of the project in 1.1, Motivation of the project in 1.2, Objective of the project in 1.3, Expected outcome of the project in 1.4, and report layout in 1.5. Chapter two is Background, where previous related works is deliberated. This section embodies, Terminologies in 2.1, preceding research activity or related works in 2.2, research analysis or short summary in 2.3 and at the last of the section, challenges in 2.4. Chapter three justifies Research Methodology. This chapter has described the project techniques in 3.1 and at the last of the section, A proposed processing model in 3.2. The most important part is Materials, Experiment Result and Discussion part found in Chapter four. Here Data set utilized developing the project in 4.1, Experiment Results and Analysis in 4.2 and at the last of the section, Discussion in 4.3. Chapter five reports Effect on Society, Environment, Ethical Views and Workability Plan in 5.1, 5.2, 5.3 and 5.4. At the last section, in Chapter six which includes Overview, Conclusion and Further Research for Future Work in 6.1, 6.2 and 6.3 followed by the relevant references.

CHAPTER 2
BACKGROUND

2.1 Introduction

The goal of face detection is to discover if there are any faces in the image or video. If multiple fonts are present, each face is enclosed by a bounding box, and thus we all know the faces location. Human faces are difficult to model as many variables can change, such as facial expression, orientation, lighting conditions, and partial occlusions such as caps, sunglasses, scarf, masks, etc. The detection result gives the face location parameters. In our case, we used images to identify the face with caps and masks. No research has been done on this. Some research is related to it and used several algorithms, which does not give an accurate result. Here, the technique we applied in this paper successfully performs to determine the face with cap and mask in real-time. In this way, we will get a more accurate result.

2.2 Related works

Table 2.1. Literature Review

Author (year) [ref]	Title	Description	Achievement Claimed
Shaukat Hayat, She Kun, Zuo Tengtao, Yue Yu, Tianyi Tu, Yantong Du (2018) [8]	A Deep Learning Framework Using Convolutional Neural Network for Multi-class Object Recognition.	They proposed model was further tuned, and the recognition performance was improved. They use nine different object classes taken from wide varied image dataset caltech101 and deploy five layers Convolutional Neural Network model. They compared the proposed model's performance with different classical bag-of-words (BOW) approaches trained on a novel	This algorithm achieved an accuracy level of 90.12% is much better than different classical BOW approaches.

		dataset consisting of 5 other wild animals class objects.	
Kanimozhi S, Gayathri G and Mala T (2019) [13]	Multiple Real-time object identification using Single shot Multi-Box detection.	In this paper, they have tried to recognize the object shown in front of a web camera. They can detect and track the object in the sports field to make the computer learn Deeply, which is none other than the application of Deep Learning.	The proposed method increases the accuracy level in identifying real-time household objects.
Hanen J abnoun, Faouzi Benzarti and Hamid Amiri (2015) [12]	Object Detection and Identification for Blind People in Video Scene.	This paper presents a visual substitution system for blind people based on object recognition in the video scene. This system uses SIFTS keypoints extraction and features matching for object identification.	The proposed system showed good accuracy for detecting objects.
Qinfeng Li (2019) [7]	An Improved Face Detection Method Based on Face Recognition Application.	This paper suggests an advanced diagonal detection method, using a K parallel bottleneck connection structure, spacing parameters in each bottleneck connection structure, and using parameter partition sharing to decrease overfitting.	The proposed method has good detection accuracy and robustness.

			5
Aniruddha Srinivas Joshi, Shreyas Srinivas Joshi, Goutham Kanahasabai, Rudraksh Kapil and Savyasachi Gupta (2020) [10]	Deep Learning Framework to Detect Face Masks from Video Footage.	In this paper, a new approach for detecting face masks from videos is proposed. A highly effective face detection model is applied for obtaining facial images and cues. A distinct facial classifier is built using deep learning to determine the presence of a face mask in the facial images detected.	The proposed method has shown its good effectiveness in identifying facial masks by achieving high precision, recall, and accuracy.
Wenxuan Han, Zitong Huang, Alifu.kuerban, Meng Yan and Haitang Fu (2020) [1]	A Mask Detection Method for Shoppers Under the Threat of COVID-19 Coronavirus.	This paper proposed a modified SSD method to detect whether shoppers are wearing masks in the supermarket. To identify whether shoppers are wearing masks, we created the COVID-19-Mask dataset, which can provide future study data.	The proposed method results show the high detection precision and real-time performance of the proposed algorithm.
Rui Min and Jean-Luc Dugelay (2011) [6]	Cap Detection for Moving People in Entrance Surveillance.	This paper introduces a solution to a recently identified occlusion problem time-variant occlusion due to cap in entrance surveillance.	The proposed algorithm evaluates several surveillance videos and yields

		The proposed system can be applied to a wide range of security management applications nowadays in video surveillance.	good detection rates.

"Table 2.1" shows a literature review, which is that the other research works done using different techniques and algorithms.

2.3 Research summary

In this paper, we showed the way to face detection with cap and mask. We have taken some data from the google site, but we create maximum data. We build a model that can quickly detect the face with cap and mask. For building a deep learning network that solves object detection problems, we use the TensorFlow object detection API framework. For that, we use an automated detection approach with the deep learning models. We run the algorithm in the data set. For our selected algorithm, we got the result. We determine if our algorithm works properly, and it can give us accurate results, Whether or not to identify the cap and the masked face.

2.4 Challenges

- Not getting enough data to work
- Less amount of training data
- Non-representative Training data
- Poor quality of data

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

Here are the techniques how we complete our project is given:

1) Object Detection
2) SSD (Single Shot Detector) Architecture
3) Tensorflow API for object detection

Object Detection

Object detection is one of the most significant leading research areas in deep learning, computer vision, and image processing that recognizes specific class semantic objects (such as people or animals) in computerized pictures and video recordings. Pedestrian detection and face detection are among the well-researched domains of object detection. Object detection becomes a vital application inside various deep learning fields, including counting picture recovery and video checking.

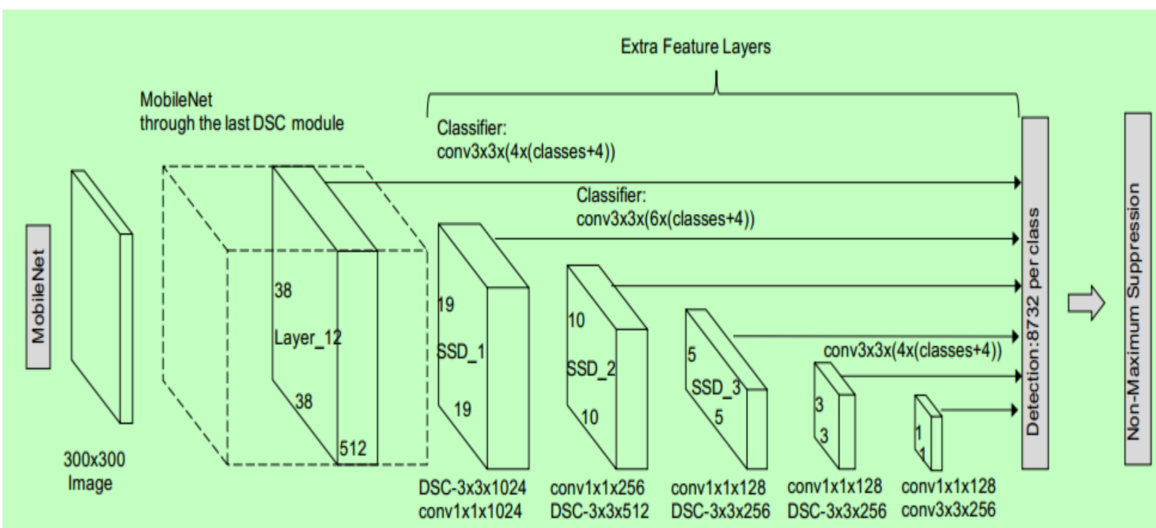
It uses a multi-label classifier that can distinguish both objects from one another but still doesn't know which ones. This is where Image Localization comes in as it identifies the object's location in the image and returns a bounding box identifying the object in the picture.

Object Detection = Image Localization + Multi-Label Image Classifier

This technology has various real-world applications like video surveillance or image retrieval, with facial detection being one of the most used fields. It can also be used at traffic lights for walker detection to better direct traffic or even help the blind walk and Sign Language Detection (SLD) for Communicating with deaf and dumb people in our community.

SSD Architecture

The complete form of SSD is Single Shot Detector. The Single Shot Detector (SSD) architecture could be an individual convolution network that classified these locations in one pass and prognosticates bounding box locations. Therefore, SSD is often trained end-to-end. The Single Shot Detector (SSD) network organize of base architecture design accompanied by various convolution layers:



SSD Mobilenet Layered Architecture

Using the SSD, we had to be compelled to take one single shot to identify numerous objects inside the picture. Simultaneously, regional proposal network-based methods such as the R-CNN group require two shots to generate region proposals, one for recognizing each proposal's object. Hence, SSD will be much speedier compared with two-shot regional proposal network-based methods.

Tensorflow API for object detection

For building a deep learning network that solves object detection problems, we use the TensorFlow object detection API framework. They provide a set of detection models pre-trained on the COCO dataset. These models can help for out-of-the-box inference if we're curious about categories already in those datasets.

They are also helpful for initializing any models when training on novel datasets. The varied architectures utilized in the pre-trained model are described in this table.

Model name	Speed (ms)	COCO mAP	Outputs
CenterNet HourGlass104 512x512	70	41.9	Boxes
CenterNet HourGlass104 Keypoints 512x512	76	40.0/61.4	Boxes/Keypoints
CenterNet HourGlass104 1024x1024	197	44.5	Boxes
CenterNet HourGlass104 Keypoints 1024x1024	211	42.8/64.5	Boxes/Keypoints
CenterNet Resnet50 V1 FPN 512x512	27	31.2	Boxes
CenterNet Resnet50 V1 FPN Keypoints 512x512	30	29.3/50.7	Boxes/Keypoints
CenterNet Resnet101 V1 FPN 512x512	34	34.2	Boxes
CenterNet Resnet50 V2 512x512	27	29.5	Boxes
CenterNet Resnet50 V2 Keypoints 512x512	30	27.6/48.2	Boxes/Keypoints
CenterNet MobileNetV2 FPN 512x512	6	23.4	Boxes
CenterNet MobileNetV2 FPN Keypoints 512x512	6	41.7	Keypoints
EfficientDet D0 512x512	39	33.6	Boxes
EfficientDet D1 640x640	54	38.4	Boxes
EfficientDet D2 768x768	67	41.8	Boxes
EfficientDet D3 896x896	95	45.4	Boxes
EfficientDet D4 1024x1024	133	48.5	Boxes
EfficientDet D5 1280x1280	222	49.7	Boxes
EfficientDet D6 1280x1280	268	50.5	Boxes
EfficientDet D7 1536x1536	325	51.2	Boxes
SSD MobileNet v2 320x320	19	20.2	Boxes
SSD MobileNet V1 FPN 640x640	48	29.1	Boxes
SSD MobileNet V2 FPNLite 320x320	22	22.2	Boxes
SSD MobileNet V2 FPNLite 640x640	39	28.2	Boxes
SSD ResNet50 V1 FPN 640x640 (RetinaNet50)	46	34.3	Boxes
SSD ResNet50 V1 FPN 1024x1024 (RetinaNet50)	87	38.3	Boxes
SSD ResNet101 V1 FPN 640x640 (RetinaNet101)	57	35.6	Boxes
SSD ResNet101 V1 FPN 1024x1024 (RetinaNet101)	104	39.5	Boxes
SSD ResNet152 V1 FPN 640x640 (RetinaNet152)	80	35.4	Boxes

There are a variety of models.

3.2 Proposed Model

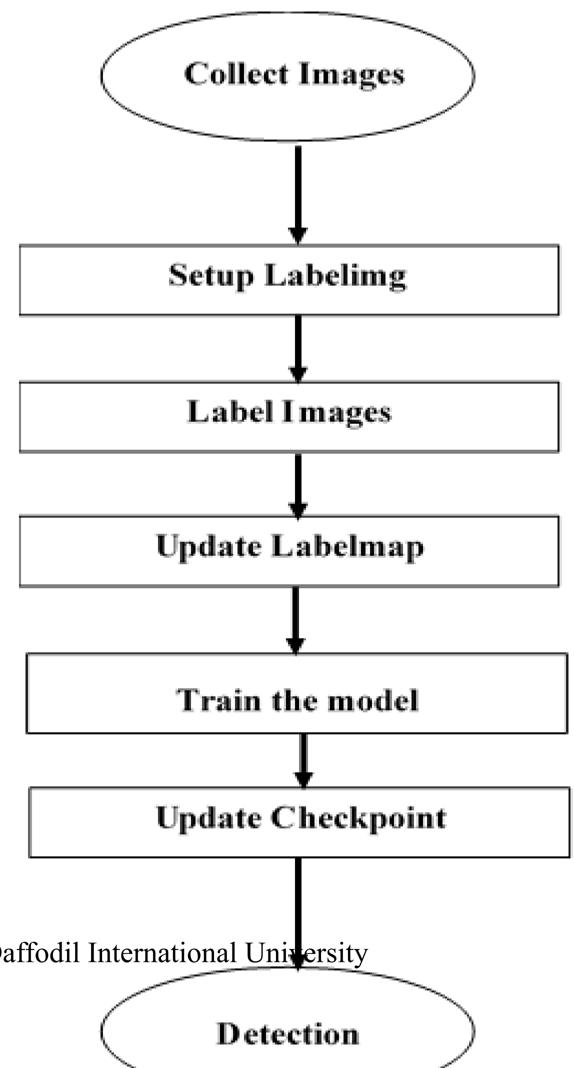


Fig 3.1: Proposed Model

CHAPTER 4

MATERIALS, EXPERIMENT RESULT AND DISCUSSION

4.1 Dataset

For this project, we are collecting 1000 real-time images. There are 750 images for training purposes with 750 label images XML files, and there are also 250 images for test purposes with 250 label images XML files. We use 75% of the images for training purposes, and for test purposes, we use 25% of images.

Table 4.1. The number of images per dataset

DATASET	ORGINAL IMAGES	LABEL IMAGES (XML)
Training Set	750	750

Test Set	250	250

4.1.1 Data labeling

Make a dataset; the dataset must be labelled perfectly.

The labelling image is an attachment below:

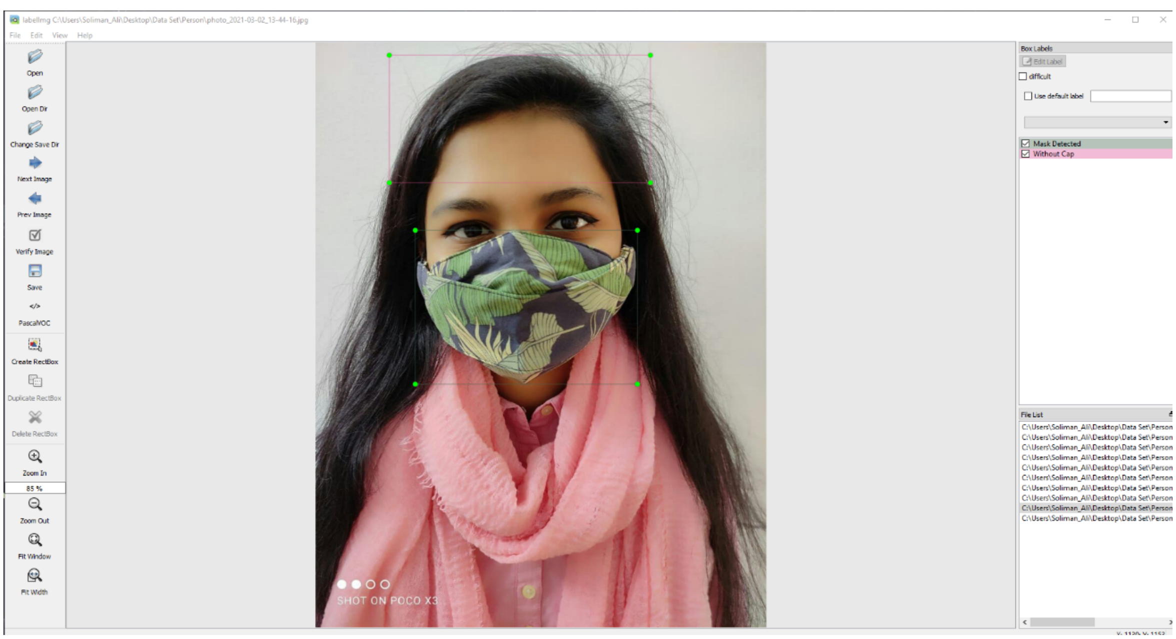


Fig 4.1: Labeling image

4.1.2 Dataset Structure

After the labeling process is finished, an XML file is generated for each sheet image. The sample XML file is given below:


```
  <?xml version="1.0"?>
- <annotation>
    <folder>collectedimages</folder>
    <filename>photo_2021-03-02_13-44-18.jpg</filename>
    <path>C:\Users\Soliman_Ali\Desktop\Face Mask Cap Detection\Tensorflow\workspace\images\collectedimages\photo_2021-03-02_13-44-18.jpg</path>
  - <source>
      <database>Unknown</database>
    </source>
  - <size>
      <width>868</width>
      <height>1156</height>
      <depth>3</depth>
    </size>
    <segmented>0</segmented>
  - <object>
      <name>Without Cap</name>
      <pose>Unspecified</pose>
      <truncated>0</truncated>
      <difficult>0</difficult>
    - <bndbox>
        <xmin>141</xmin>
        <ymin>27</ymin>
        <xmax>677</xmax>
        <ymax>270</ymax>
      </bndbox>
    </object>
  - <object>
      <name>Mask Detected</name>
      <pose>Unspecified</pose>
      <truncated>0</truncated>
      <difficult>0</difficult>
    - <bndbox>
        <xmin>186</xmin>
        <ymin>373</ymin>
        <xmax>619</xmax>
        <ymax>678</ymax>
      </bndbox>
    </object>
  </annotation>
```


Fig 4.2: XML file of labeling image

4.2 Experiment Results and Analysis

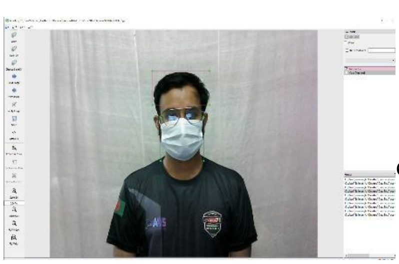
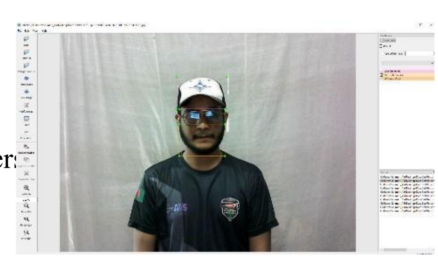
Ascertaining the section of various things on a particular picture and composing a conventional machine learning model that identifies these things is yet one of the complicated matters in deep learning. The TensorFlow's Object Detection API implies an open-source application interface built on TensorFlow's by Google that promotes the invention, training, and deployment of object apprehension models.

This technology detects a person with a cap & the mask also sees the person without a cap and mask. This analysis will train the deep learning model using TensorFlow Object detection API; We can apply SSD _mobile net, which is a per trained model. It identified a person when the person stayed in front of the web camera.

Here, we given some training data set for trained it and label them. We also put some data set as a test purpose and provide the output when detecting a person.

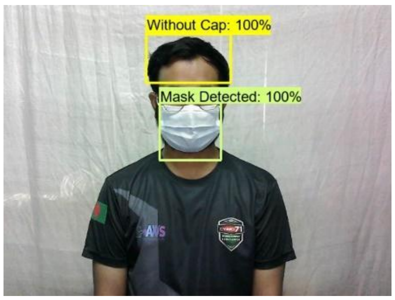


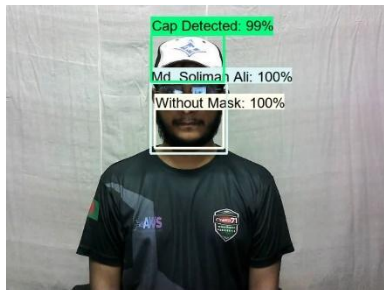
Original Images

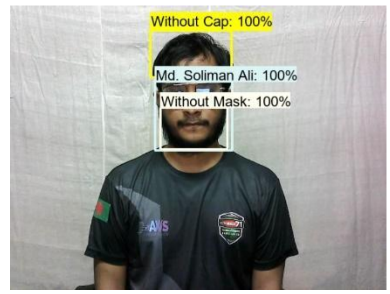




er

Labeling Images



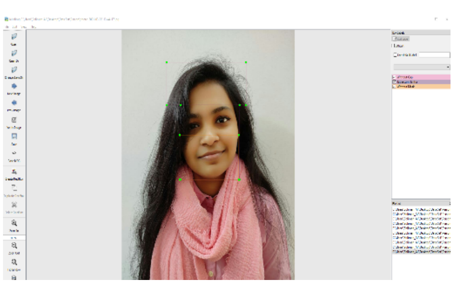
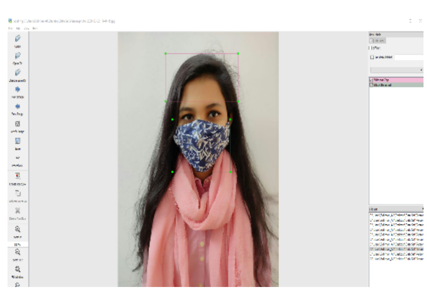
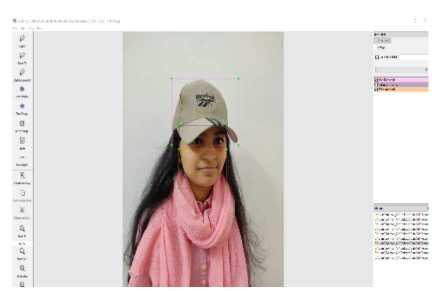




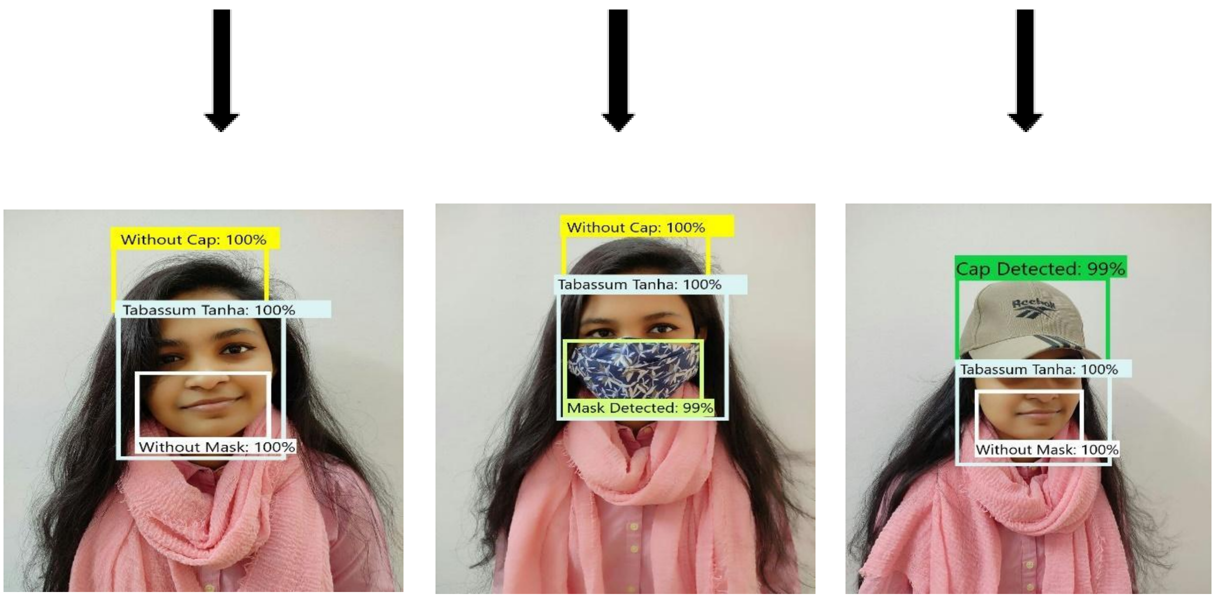
Output Images



Original Images

Labeling Images



Output Images

In the overview of all of the images, we give some original images, images with caps, and images with masks of a person. In the second step, it labels them, and the final stage when the person stands in front of the web camera, which also texts the person's name and either the person wears a mask or cap, also detects that or no mask or no cap it shows the text.

Different types of loss values diagram are given below:

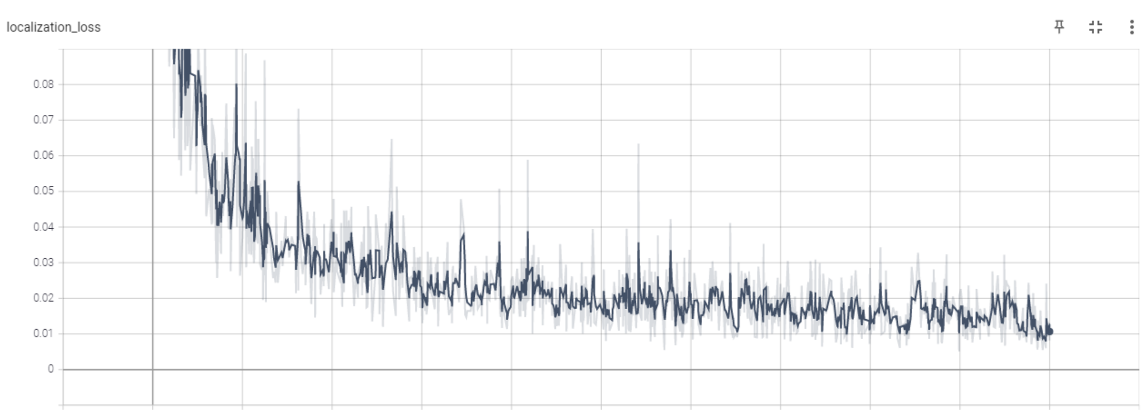


Fig 4.3: Localization Loss

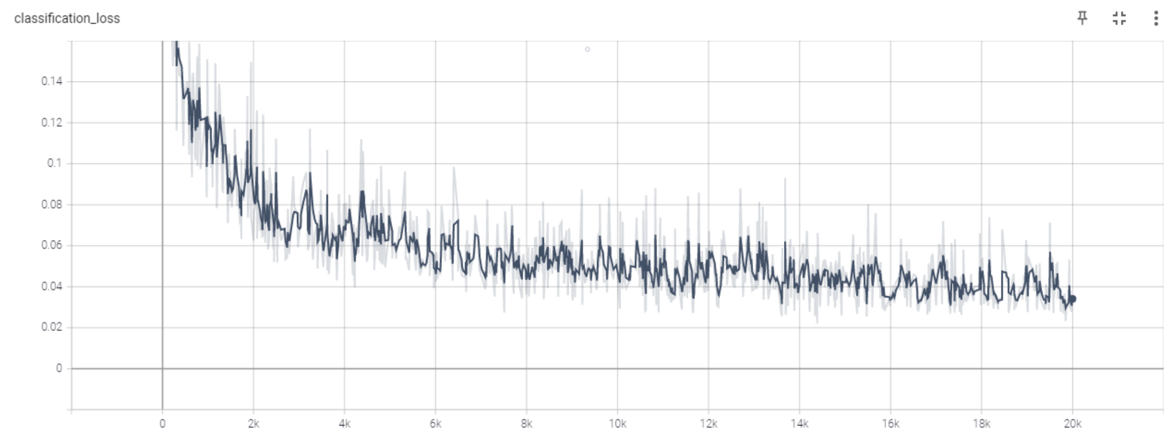


Fig 4.4: Classification Loss

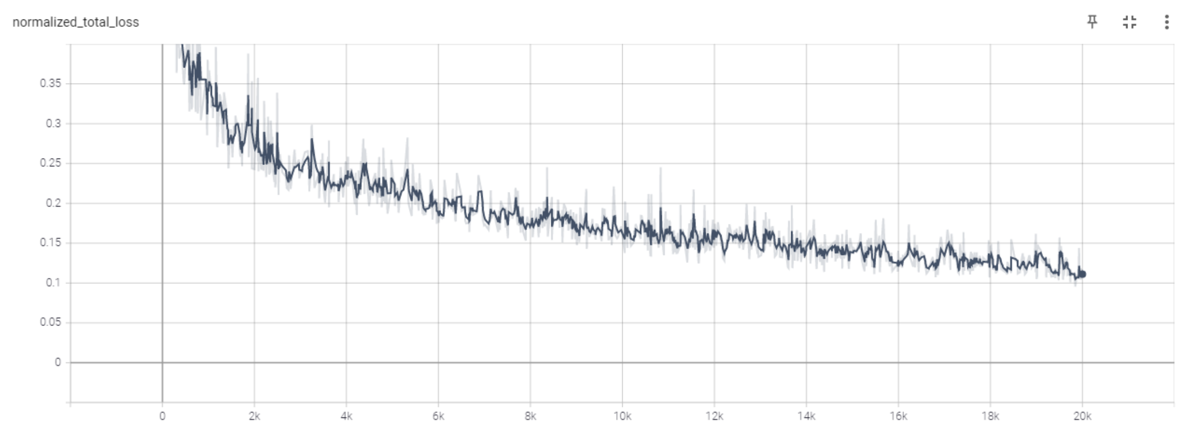


Fig 4.5: Normalized total Loss

And also, the learning rate values diagram are given below:

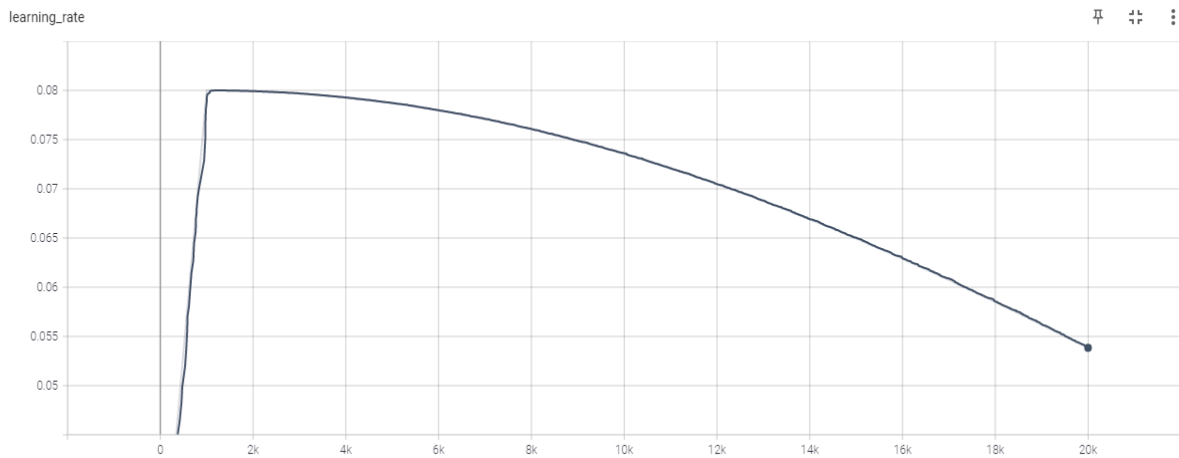
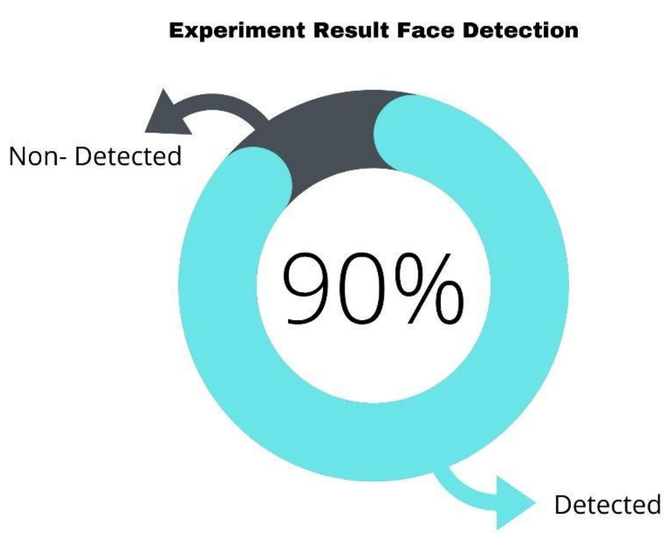


Fig 4.6: Learning Rate

4.2.1 Result Analysis

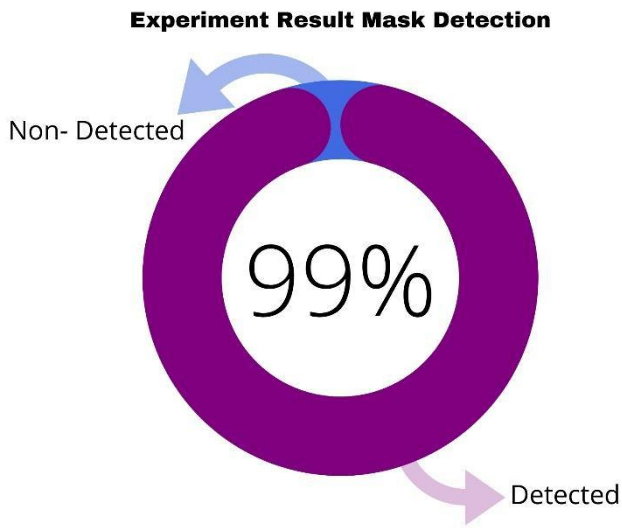
Face detection

Our project identifies a person perfectly, but when a person wears a cap and masks, it is challenging to detect the person. Still, it distinguishes the person almost 90% accurately.



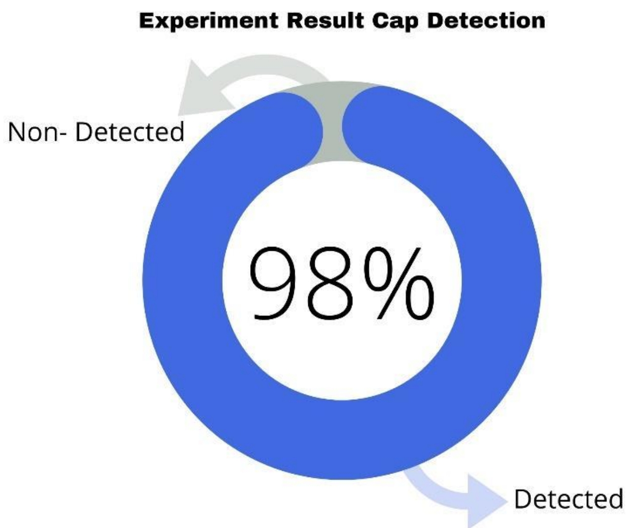
Mask detection

In our project, it distinguishes the person wears a mask or not. It detects that 99% accurately.



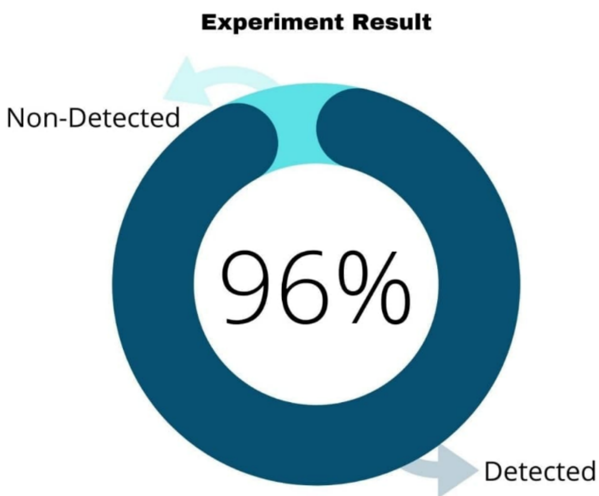
Cap detection

In our project, it identifies the person wears a cap or not. It detects that 98% accurately.



Final Analyses

In our project, the overall accuracy rate is 96%.



4.3 Discussion

In the experiment result and result analysis, we show how our project works step by step and the exactness of output. The project works 96% accurately applying this SSD mobile net model. On the other hand, some technology works not much, but our model frequently uses this model, which gives good efficiency.

CHAPTER 5

EFFECT ON SOCIETY, ENVIRONMENT AND WORKABILITY

5.1 Effect on Society

We implemented an algorithm to solve this matter, which elevates according to society and the environment. We anticipate that our strategy will produce an immeasurable impression on our community. Nowadays, crime has risen in our society, country a lot. While looking for those criminals search through the CCTV footage is a lengthy process to obtain the criminal but using our project when the person in front of the camera determined the person instant & easily. Thus, there are many good presents to society.

Point out the areas of society where this strategy contributes to:

- In protected sectors or criminal investigation branch.
- In industrial security sectors
- Find any missing person
- Covid-19-mask detection.

Those are the main significant sectors where face detection with cap & mask technique or algorithm assigns.

5.2 Effect on Environment

We believe that our technology proffers a great impression on the environment also. Implementing our technology, people will be available, which works in the defense department for catching criminals, which is helpful for general people. When there are no rebellious people around us, the environment feels free.

5.3 Ethical Views

This algorithm is entirely free of charge and open-sources, which helps usual people. This is the principal ethics we follow. Furthermore, this solution does not accord personal images or identities and also ensures safety and security.

5.4 Workability Plan

While this algorithm is entirely free of charge, a donation from the government, top-ranking bank, or investors to execute a different sector strategy is demanded. If this strategy is combined into the country's financial sector, a massive possibility for predicted the project.

CHAPTER 6

OVERVIEW, CONCLUSION AND FURTHER RESEARCH

6.1 Overview of the Study

Deep learning is an immense field for identifying anything. In this paper, a suitable algorithm is applied. We are associating with distinct methods. We raised the model, and it furnishes immeasurable feedback and apprises that the algorithm has a magnificent yield rate. There is seven-part to completing the project: collect Images, setup labelling, label images, update label map, train, update checkpoint, and detection. The algorithm is lattice enrichment to the research field.

6.2 Conclusions

Detecting from live capture is a very challenging task. Deep learning and SSD_mobile net is very promising machine learning strategies to deal with this difficulty. Our project is mainly to detect a person, cap, and mask. When the person in front of a web camera then caught those things. Our dataset is adequate in times of ascertaining the person. However, the model obtains person, cap, and mask types trained with sufficient samples with total efficiency. Experiments in the conventional data set and natural environment show that the recommended method has good detection accuracy.

6.3 Further Research

In this study, the algorithm which we applied has some limitation at present.

Limitation:

- If there are no data of the person in the data set, it may not detect them.
- When the data set increase, it may not give efficiency before it may be a slight bit decrease.
- It is cost-efficient because we need a webcam and a device.

Future scope:

- We will develop our project with a better efficient algorithm.
- We will try to convert it into a small device to reduce the cost.
- We set the alarm if the person wears a mask or cap into the ATM booth.

References

[1] Wenxuan Han, Zitong Huang, Alifu Kuerban, Meng Yan and Haitang Fu, "A Mask Detection Method for Shoppers Under the Threat of COVID-19 Coronavirus," *International Conference on Computer Vision Image and Deep Learning,* pp. 442-447, 2020.

[2] Chenghao Zheng, Menglong Yang and Chengpeng Wang, "A real-time face detector based on an end-to-end CNN," *International Symposium on Computational Intelligence and Design,* pp. 393-397, 2017.

[3] Edy Winarno, Imam Husni Al Amin, Herny Februariyanti, Prajanto Wahyu Adi, Wiwien Hadikurniawati and Muchamad Taufiq Anwar, "Attendance System Based on Face Recognition System Using CNN-PCA Method and Real-time Camera," *International Seminar of Research of Information Technology and Intelligent System,* pp. 301-304, 2019.

[4] Stanislav Smirnov and Alma Eguizabal, "Deep learning for object detection in fine-art paintings," *the Institute of Electrical and Electronics Engineers,* pp. 45-49, 2018.

[5] Nishajith.A, Nivedha.J, Shilpa.S.Nair and Prof.Mohammed Shaffi.J, "SMART CAP – WEARABLE VISUAL GUIDANCE SYSTEM FOR BLIND," *International Conference on Inventive Research in Computing Applications,* pp. 275-278, 2018.

[6] Rui Min and Jean-Luc Dugelay, "Cap Detection for Moving People in Entrance Surveillance," *the Institute of Electrical and Electronics Engineers,* pp. 01-04, 2011.

[7] Qinfeng Li, "An Improved Face Detection Method Based on Face Recognition Application," *4th Asia-Pacific Conference on Intelligent Robot System,* pp. 260-264, 2019.

[8] Shaukat Hayat, She Kun, Zuo Tengtao, Yue Yu, Tianyi Tu and Yantong Du, "A Deep Learning Framework Using Convolutional Neural Network for Multi-class Object Recognition," *the Institute of Electrical and Electronics Engineers,* pp. 194-198, 2018.

[9] Aniruddha Srinivas Joshi, Shreyas Srinivas Joshi, Goutham Kanahasabai, Rudraksh Kapil and Savyasachi Gupta, "Deep Learning Framework to Detect Face Masks from Video Footage," *12th International Conference on Computational Intelligence and Communication Networks,* pp. 435-440, 2020.

[10] Astha Singh, "Face Detection using Deep Recurrent Learning and SMQT Technique," *International Conference on Electronics and Sustainable Communication Systems,* pp. 17-22, 2020.

[11] Hanen J abnoun, F aouzi Benzarti and Hamid Amiri, "Object Detection and Identification for Blind People in Video Scene," *the Institute of Electrical and Electronics Engineers,* pp. 363-367, 2015.

[12] Kanimozhi S, Gayathri G and Mala T, "Multiple Real-time object identification using Single shot Multi-Box detection," *International Conference on Computational Intelligence in Data Science,* pp. 01-05, 2019.

[13] Qi Guo, Zhihui Wang, Caixia Wang and Dong Cui, "Multi-face dection algorithm suitable for video surveillance," *International Conference on Computer Vision, Image and Deep Learning,* pp. 27-33, 2020.

[14] Xiujie Qu, Tianbo Wei, Cheng Peng and Peng Du, "A Fast Face Recognition System Based On Deep Learning," *11th International Symposium on Computational Intelligence and Design,* pp. 289-292, 2018.

[15] Ya Wang, Tianlong Bao, Chunhui Ding and Ming Zhu, "Face Recognition in Real-world Surveillance Videos with Deep Learning Method," *2nd International Conference on Image, Vision and Computing,* pp. 239-243, 2017.

[16] Md. Rafiuzzaman Bhuiyan, Sharun Akter Khushbu and Md. Sanzidul Islam, "A Deep Learning Based Assistive System to Classify COVID-19 Face Mask for Human Safety," *the Institute of Electrical and Electronics Engineers,* pp. 01-05, 2020.

[17] Gabriel T.S. Draughon, Peng Sun, PhD and Jerome P.Lynch, PhD, "Implementation of a Computer Vision Framework for Tracking and Visualizing Face mask Usage in Urban Environments," *the Institute of Electrical and Electronics Engineers,* pp. 01-08, 2020.

[18] Suleman Khan, Ehtashan Ahmed, M. Hammad Javed, Syed A A Shah and Syed Umaid Ali, "Transfer Learning of a Neural Network Using Deep Learning to Perform Face Recognition," *International Conference on Electrical Communication and Computer Engineering,* pp. 01-05, 2019.

[19] Mohammad Marufur Rahman, Md. Motaleb Hossen Manik, Md. Milon Islam, Saifuddin Mahmud and Jong-Hoon Kim, "An Automated System to Limit COVID-19 Using Facial Mask Detection in Smart City Network," *the Institute of Electrical and Electronics Engineers,* pp. 01-05, 2020.