Real-Time **Face Mask Recognition** in Low Resolution **Surveillance Videos**

1 Introduction

Mask mandates have been recently introduced throughout the world to combat the virus pandemic. These mandate must be enforced by authorities hence stretching their resources further with many other restrictions that they have to enforce.

The project aims to:

- **Encourage people to wear mask** by having surveillance systems with face mask recognition act as a deterrent.
- **Relieve pressure** for authorities that are enforcing mask mandates.
- **Provide reliable statistical data** of mask compliance in an area to be used in conducting further investigation into face mask effectiveness.

2 Background

Deep convolutional neural networks imitate the human brain's visual cortex. It works by extracting general features of images that are then narrowed down to class-specific. CNN is dominating the object recognition field, with some CNN based models have achieved human-like performance on large benchmark datasets.

3 Technology

Python a high level programming language widely used in machine learning for rapid prototyping due to its simplicity.

TensorFlow is a machine learning framework used for training and analysis of deep learning models.

OpenPose is a multi-person pose tracker, it is the first real-time capable solution for human pose estimation.

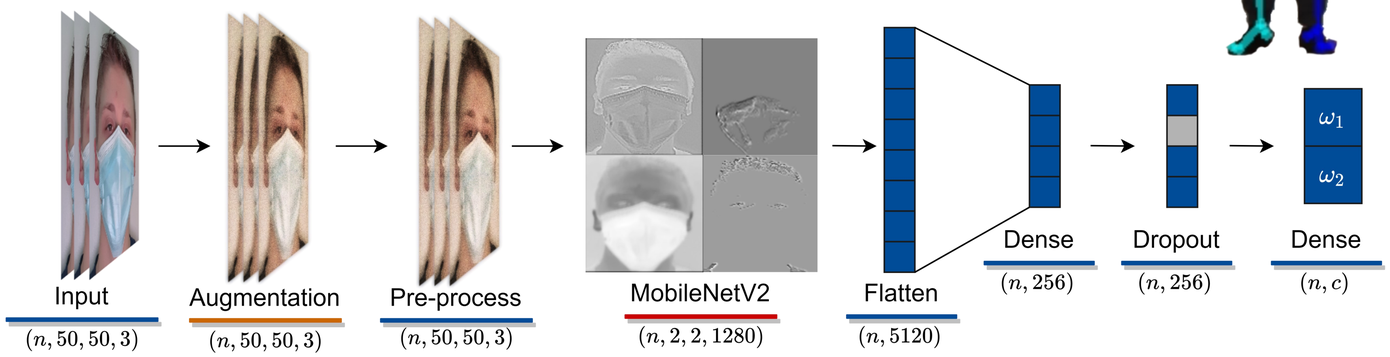
4 Methodology

The proposed systems consist of using human pose estimation to track and extract visible faces from pedestrians and classify them using a deep convolutionary neural network model.

Face mask classifier is trained on a dataset consisting of **19360** images. The model's architecture uses a MobileNetV2 for feature extraction that was pre-trained on a large dataset. It leverages previously learned weights on a new problem for better generalisation.

Face detection approach uses a multi-person pose estimation solution developed by OpenPose. The detected keypoints are used to extract pedestrians' visible faces that are then fed through the face mask recognition model to determine if pedestrian is wearing a mask.

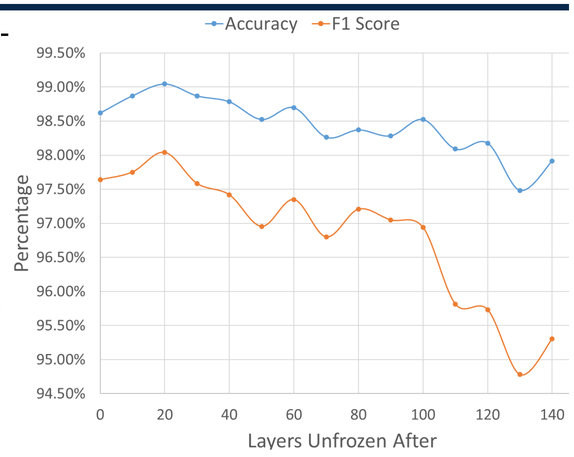
Pedestrians are identified and tracked throughout the frames using the **Kalman filter** to determine the pedestrian's trajectory and future position based on preliminary motion information gathered by extracting human body keypoints to create a motion model.



5 Findings

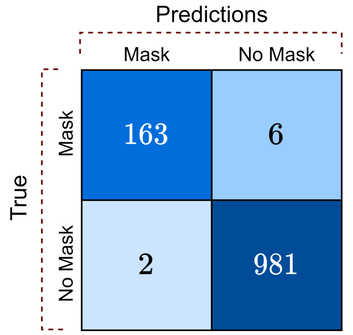
The use of **transfer learning and fine-tuning** for pre-trained MobileNetV2 has shown to provide **best performance** when layers are unfrozen after 20^{th} layer rather than training the base model from scratch for dataset of **19360** images.

Further investigation has shown that pre-trained model's initial layers produced general features, whereas deeper layers' features were more specific to the original classification task, however reusing some of the first layers for a similar classification task provided better performance than training model from scratch.



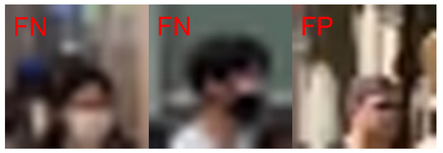
6 Results

The face mask model has achieved an F_1 score of **97.61%** and an **accuracy** of **99.31%**. The model can accurately classify low-quality images of faces that are of high variation.



Precision	98.79%
Recall	96.45%
F_1-score	97.61%
Accuracy	99.31%

The model has misclassified only **2** out of **983** faces without masks, and **6** out of **169** faces with masks.



7 Conclusion

A real-time capable system has been developed to detect, track, and classify multiple pedestrians across frames in low-resolution videos to determine mask compliance in an area.

Armandas Bespalovas

Supervisor Dr Franck P. Vidal

PRIFYSGOL BANGOR UNIVERSITY