Mobile Application used to Detect the Correct Face Mask wearing based on Image Segmentation

By Rokas Budreika

A report submitted as part of the requirements for the degree
of BSc (Hons) in Computing: Computer Science
at the School of Computing
Robert Gordon University,
Aberdeen, Scotland

November 2021

Supervisor Carlos Moreno-Garcia

Abstract

COVID-19 is a disaster to human mankind. The whole world is now struggling to reduce the spread of COVID-19 virus. Wearing masks is a good practice that helps to mitigate the COVID-19 effectively. From the results of South Korea and China is clear that wearing face mask reduces the virus spread. However, ensuring all peoples wearing facemask is not an easy thing. This research attempts to develop a simple and effective model for real-time surveillance. The proposed model should successfully recognize if a person is wearing a face mask or not.

TABLE OF CONTENTS

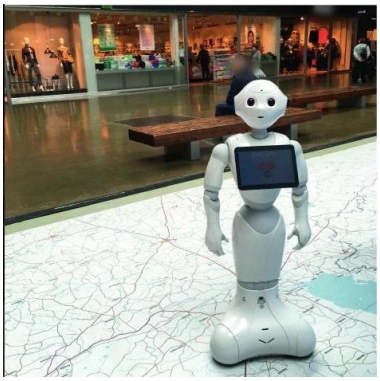
List of Figures

List of Tables

I. Introduction

Nowadays more and more people come across problems that could affect their health. Because of the particle pollution in the air, the risk of getting sick with chronic diseases are getting higher, especially in countries like China. Flu virus can also cause the pandemic and negatively affect public life. These days predominant Covid-19 pandemic causes to take drastic actions, which would stop spreading of virus, to save people's lives and to let to continue development of economy. Protective face masks are recommended or even necessary tool to stop spread of flu and Covid-19 viruses. Protective masks also help from particle pollution. However, not all of the people are concerned enough to wear masks when it is necessary or simply wear them incorrectly. Thinking about bigger picture and to protect the public health from upcoming health issues and possible death, especially pandemic time, we have to raise higher number of people wearing the masks.

Living this high-tech century, where AI (artificial intelligence) is becoming more popular, efficient and effective as a tool, we can use this tool to ensure face masks wearing. Using neural networks and machine learning technology, we can create AI method, which would monitor places where there are a lot of people. The main function of this method would be to decide if a person wearing a face mask and if he/she wearing it correctly. Using this information, we can use it as a tool to decrease number of people without face masks or wearing incorrectly. We can use this system within robots in shopping centres, which not only consult, navigate people but also would ask to wear face mask or adjust face mask to wear correctly.



Picture 1. Robot Pepper able to recognise people faces, communicate and etc. [13]

1.1. Significance of wearing face masks

From previous ages there was a lot of people that were wearing face masks to try preserving their health. XVII century plague doctors were wearing exceptional looking face masks in process to not get affected by "Bubonic plague". XIX century it was discovered that bacteria were on dust particles. 1878 it was suggested to wear face masks to reduce infections during the pandemic period. At 1897 start appearing first surgeons, which started using face masks during operations time. XX century view to wearing face masks drastically changed. Wearing face masks was normal even for public people its terms to stop 1918 prevailing flu pandemic. 1920 wearing a mask for surgeons was normal during operation times. However, even nowadays wearing a face masks it is not an exception.

Every year flu bacteria are spreading in the whole world. At some point there is new outbreak, which could cause world pandemic. 2009 flu H1N1, more popular known as “Swine Influenza” (“Pig flu”), during pandemic time national and international health agencies were recommending use of face masks. During pandemic a study was conducted, in which there was 9 volunteers which was sick from this flu. Volunteers had to cough 5 times on “Petri dish” with viral transmission medium. This experiment was also conducted with same volunteers that were wearing a medical face mask and also respirators N95. Results were shocking, without face masks all 9 volunteers were detected sick, while with any kind of face masks there were no detection of them being sick on “Petri dish” samples [23]. That proves significance of wearing face masks and that does indeed stop bacteria spreading.

December of 2019 China, Wuhan city had pneumonia outbreak of unknown cause, which caused big attention not just for China, but for the whole world [24]. This severe acute respiratory syndrome corona virus (“SARS CoV 2”), causes Covid-19 disease, also called corona virus [25]. Main symptoms of this disease are: fever, cough, fatigue, shortness of breath, sore throat, headache, a loss or change of smell or taste [26]. Corona virus infected people could feel no symptoms as well [27]. These without symptoms cases could spread disease without persons’ knowledge. It is really important to notice that this disease could complicate treatment for people that are at risk [28].

Air pollution – it is another important fact, for this reason it is important to wear face mask. In China air pollution causes around 1.6 million before age death and disturbs human life by making their health worse. Beijing solid particles (“PM2.5”) concentration often reaches minimal unhealthy levels at Winters time. [29]

Wearing a mask while going outdoors during pandemic has been a great helping hand in controlling the spread of coronavirus. It is a symbol of being a sensible citizen of a country. Several countries like China and Korea successfully controlled the Covid-19 in a short time due to the habit of using masks regularly. It was recommended to use masks, no matter what type of mask is available to you, just use it and you will be safe. The mask acts as a physical barrier to prevent the entry of the virus. Many individuals unknowingly infected many other people with the Covid-19. Mask is necessary because of two benefits. It does not let the virus enter your mouth or nose directly from the infected person’s sneeze or cough. The irresponsibility of a few people has to lead to the death of many others due to the spread of the virus. Secondly, if you touch a virus-contaminated surface and then your mouth or nose, the mask will stop the transmission of the virus. Many governments made it compulsory to use masks and people were compelled and monitor to act upon the order. The paper understudy will make deep learning about using masks to prevent the spread of deadly coronavirus. It will also learn the facemask detection by deep learning strategy. This is a useful technique of machine learning being used these days.

1.2. Motivation/Justification

This project is about preserving public health. This project would be important for all companies which have places that have public areas where people gather, from attractions to workplaces. This project is to make sure that customers, workers or in general people, would wear the face covering masks and make sure that they are wearing it correct way. For example, shops, cinemas, offices, schools and etc. This project would reduce the potential of diseases spreading.

1.3. Aims and Objectives

Aim: To create a functioning mobile face mask detection application that determines/classifies if a person is wearing face mask correctly.

Objective 1: To train and test a machine learning algorithm that detects if a person is wearing a mask.

Objective 2: Once the system has identified the mask, to detect if the person is wearing it correctly.

Objective 3: Deploy the model in a mobile application.

1.4. Contents

The following document is structured as follows. Chapter 2 is about researching existing models how to tackle face wearing problems. Benefits and downsides of them. Chapter 3 is about methodology that I have encountered researching previous related models. Chapter 4 is about proposed methodology (libraries and tools) and how I am planning to create mobile face mask detection application that determines/classifies if a person is a wearing face mask correctly. Finally, chapter 5 is about conclusions for the whole research that I have done.

II. Related Work

In this chapter, I will present how others have approached this problem. This includes the data that they have used plus the methods that they have implemented.

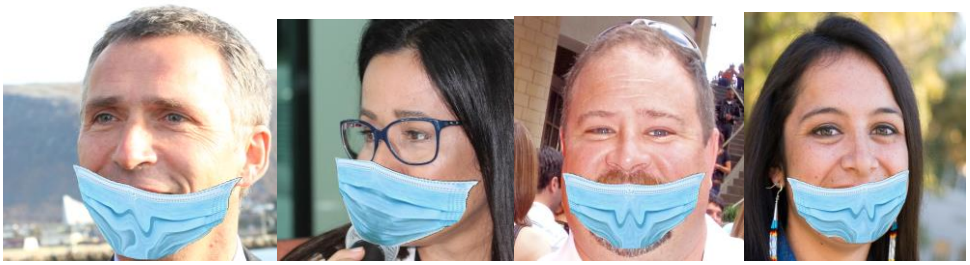
2.1. ResNet-50 convolutional neural network, SVM and ensemble method

This research was done using three datasets. First dataset is "Real-World Masked Face Dataset" ("RMFD") [9], which have 5’000 faces with masks and 90’000 faces without mask. However, for this research they have used only 10000 images, 5000 with facemasks and 5000 without. This dataset is used for learning, approval and testing stages [10].

Pictures 2. RMFD dataset pictures with masks examples [10]

Pictures 3. RMFD dataset pictures without masks examples [10]

Second dataset is about artificial (simulated) facemasks ("SMFD"), which is made of from 1570 pictures. This set is divided into 2 sets of 785 pictures with artificial facemasks and 785 without. "SMFD" is also used for learning, approval and testing stages [10].



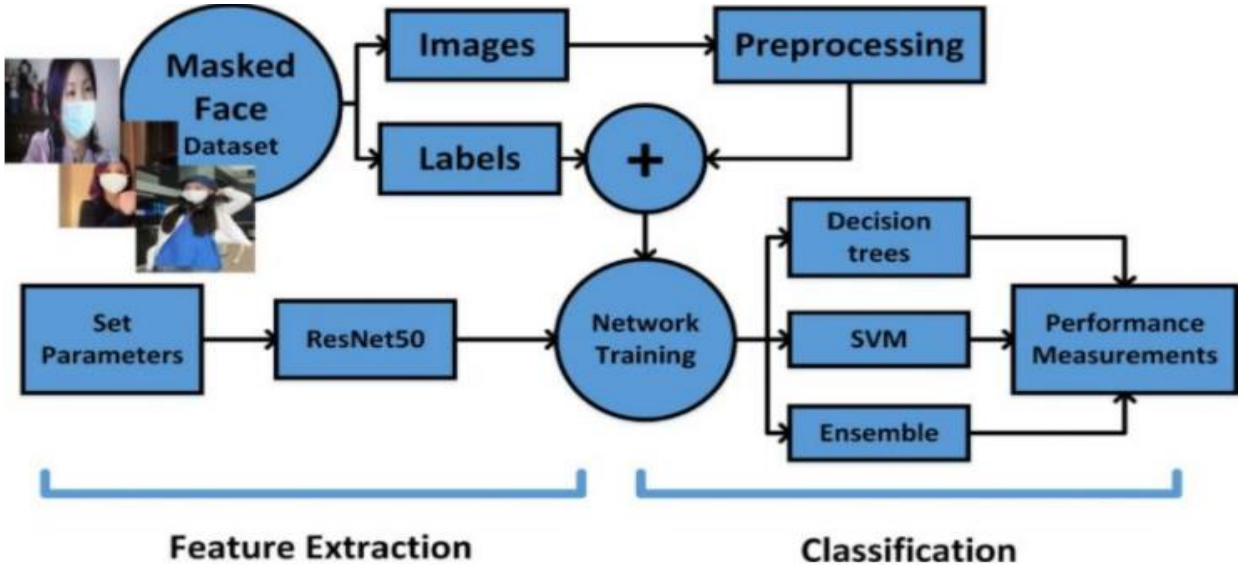
Pictures 4. SMFD pictures with masks examples [10]

Pictures 5. SMFD pictures without masks examples [10]

Third dataset is "Labelled Faces in the Wild" ("LFW"). This dataset contains celebrities faces with masks, which is made of 13000 pictures. This dataset was made only for testing, that is why this method never learnt using these images [10].

Pictures 6. LFW pictures with masks [10]

This method is made from 2 main components. First competent is "ResNet50" - convolutional neural network, which is used for extracting main face features. Second component is "SVM" - machine learning component, which is made from support vector machine ("SVM") and "ensemble". How it works that "ResNet50" extracts face features meanwhile machine learning method classify the results [10].



Pictures 7. Diagram of the proposed method [10]

Results for this method was [10]**:**

- "RMFD" got 99,64% accuracy.
- "SMFD" got 99,49% accuracy.
- "LWD" got 100% accuracy.

2.2. YOLO-v2 with ResNet-50

This research was made from 2 datasets.

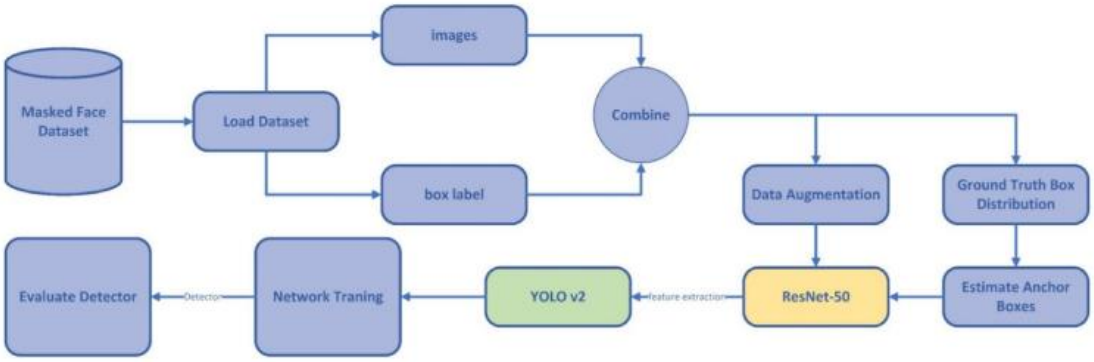
First dataset was "Medical Masks Dataset" ("MMD"). Which is made from 682 pictures and more than 3000 faces with masks [11].

Pictures 8. MMD dataset pictures examples [11]

Second dataset was faces with masks "Face Mask Dataset" ("FMD"), which is made from 853 pictures [11].

Pictures 9. FMD dataset pictures examples [11]

This method is made from 3 main components. First component is about face mapping, second about making more data, third about detecting. Detector used "YOLOv2" with "ResNet-50" for feature extraction and face masks detection learning, approval and testing stages [11].



Pictures 10. Detector method [11]

Average accuracy using this model was 81% [11].

2.3. MaskedFace-Net

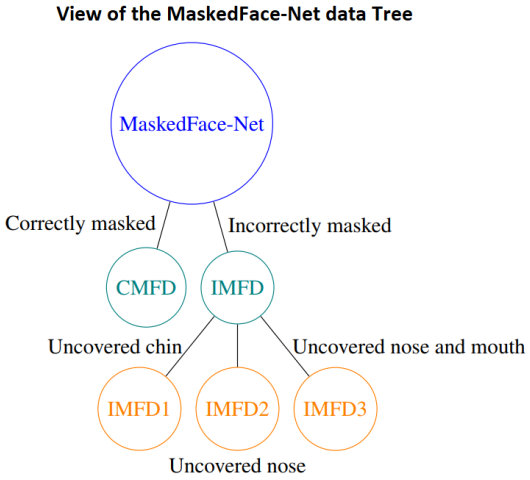
This research contains 2 datasets. First dataset, correctly masked face dataset ("CMFD") contains 67049 images [14] [15].

Pictures 11. CMFD dataset pictures examples [14] [15]

Second dataset, incorrectly masked face dataset ("IMFD") contains 66734 images [14] [15].

Pictures 12. IMFD dataset pictures examples [14] [15]

This method does not say which algorithms they have used to tackle the problem; however, it is good to include this to have big dataset of human pictures. This dataset highlights synthesis of masked face images, which we can use as benchmark other datasets for creating machine learning model. This system can then be used for enhancing vision-based monitoring systems towards several applications such as checking the respect of laws related to the mask wearing or generating crowd statistics. Moreover MaskedFace-Net has been generated for studying behaviours and contamination processes related to the Covid-19 virus.



Pictures 13. Proposed method data tree [14] [15]

2.4. Face Mask Detection

This research dataset contains total of 4095 images. Total amount images with masks 2165 [16].

Pictures 14. CMFD dataset pictures examples [16]

Total amount images without masks 1930 [16].

Pictures 15. IMFD dataset pictures examples [16]

This method uses “OpenCV”, “Caffe”, “Keras”, “TensorFlow” and “MobileNetV2” as its main components: “OpenCV” it’s an open source library for real-time computer vision; “Caffe” is used to make a deep learning framework which would be expressive, efficient and modular; uses “Keras” to minimize the number of user actions required for common use cases; also uses “TensorFlow” for easy way to develop and train Machine Learning model; followed by “MobileNetV2” which efficiently helps to run Machine Learning model on mobile systems.

2.5 Conclusions of the findings / summary

Table 1. Method Comparison [10] [11] [14] [15]

No.		ResNet-50 convolutional neural network, SVM and ensemble method	YOLO-v2 with ResNet-50	MaskedFace-Net	Face Mask Detection
1	Used Algorithms	“ResNet-50”, convolutional neural network, “SVM”, “ensemble”	“ResNet-50”, “YOLO-v2”, “SGDM”, “ADAM”	It is not publicly available to check, it is only a given a “gif” or working prototype	“OpenCV”, “Caffe-based face detector”, “Keras”, “TensorFlow”, “MobileNetV2”
2	Used Datasets	“RMFD”, “SMFD”, “LFW”	“MMD”, “FMD”	“CMFD”, “IMFD”	“CMFD”, “IMFD”
3	Capability to recognise face masks	+	+	+	+
4	Capability to recognise if a face mask is worn correctly	-	-	+	-

5	Average Accuracy	-	81%	95%	93%
6	Testing Accuracy	99.64%	-	98%	98%

III. Methodologies

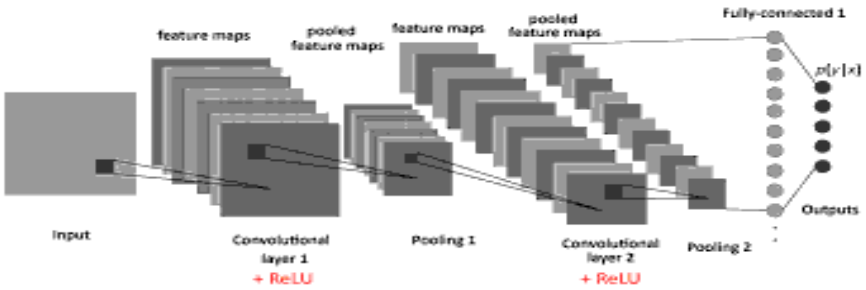
3.1. Neural Networks

"Neural Networks", also known as artificial neural networks (ANNs) or simulated neural networks (SNNs), are a subset of machine learning and are at the heart of deep learning algorithms. ANNs are comprised of a node layers, containing an input layer, one or more hidden layers, and an output layer. Each node, or artificial neuron, connects to another and has associated weight and threshold. Neural networks rely on training data to learn and improve their accuracy over time. However, once these learning algorithms are fine-tuned for accuracy, they are powerful tools in computer science and artificial intelligence, allowing us to classify and cluster data at a high velocity. Tasks in speech recognition or image recognition can take minutes versus hours when compared to the manual identification by human experts. One of the most well-known neural networks is Google's search algorithm.

3.2. Convolutional Neural Networks

"Convolutional Neural Network" (CNN) is one of the most significant tools in computer vision which is used in image recognition and processing that is specifically designed to process pixel data. Recognizing objects, face identification and etc., is just few areas, where we can use CNN [12].

CNN image-classifiers takes input image, process and classify according to specific layers. Computer sees input image as pixel array which depends on resolution. According to resolution, it will see h * w * d, where h – height, w – width, d – depth [12].



Pictures 16. CNN example [12]

3.3. Image Segmentation

“Image Segmentation” is the task of clustering parts of an image together that belong to the same object class. This process is also called pixel-level classification. In other words, it involves partitioning images (or video frames) into multiple segments or objects [22].

Picture 22. Annotated image for semantic image segmentation [22]

“Image Segmentation” plays a central role in a broad range of real-world computer vision applications, including road sign detection, biology, the evaluation of construction materials, or video surveillance [22].

“Image Segmentation” can be formulated as a classification problem of pixels with semantic labels (“Semantic Segmentation”) or partitioning of individual objects (“Instance Segmentation”). “Semantic Segmentation” performs pixel-level labelling with a set of object categories (for example, people, trees, sky, cars) for all image pixels. It is generally a more difficult undertaking than image classification, which predicts a single label for the entire image or frame. “Instance Segmentation” extends the scope of “Semantic Segmentation” further by detecting and delineating all the objects of interest in an image [22].

Due to” Deep Learning” models’ success in a wide range of vision applications, there has been a substantial amount of research aimed at developing image segmentation approaches using “Deep Learning”. At present, there are many general datasets related to image segmentation. The most popular image segmentation datasets are: “PASCAL VOC”, “MS COCO”, “Cityscapes”, “ADE20K”, “YouTube-Objects”, “KITTI” and other [22].

3.4. YOLO

YOLO (“You Only Look Once”) algorithm employs convolutional neural networks (“CNN”) to detects objects in real-time 45 frames per second (“FPS”) speed [17]. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects. This means that prediction in the entire image is done in a single algorithm run.

Yolo algorithm divides the input image into an **S x S** grid. IF the center of an object falls into a grid cell, that grid cell is responsible for detecting that object [17].

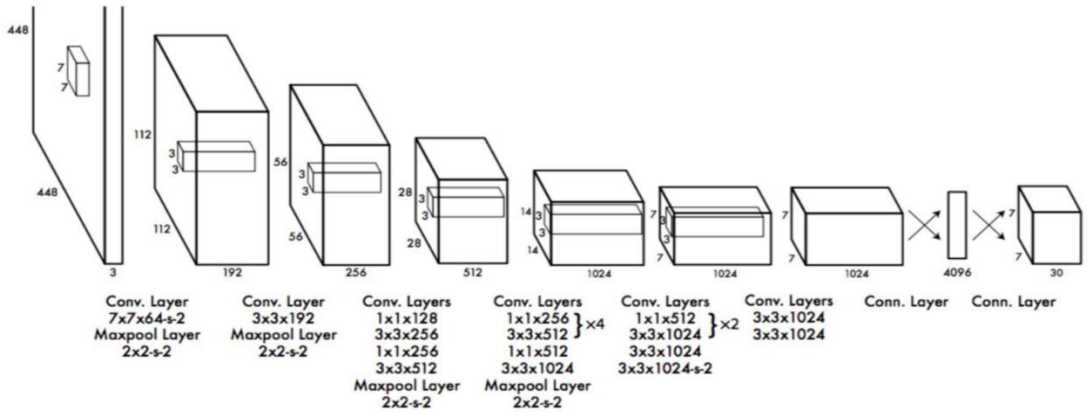
Each grid cell predicts B bounding boxes and confidence scores for those boxes. These confidence scores reflect how confident the model is that the box contains an object and also

how accurate it thinks the box is that it predicts. Formally we define confidence as **Pr (Object) * IOU**. If no object exists in that cell, the confidence scores should be zero. Otherwise we want the confidence score to equal the intersection over union (IOU) between the predicted box and the ground truth [17].

Each bounding box consists of 5 predictions, x, y, w, h, and confidence. The (x, y) coordinates represent the center of the box relative to the bounds of the grid cell. The width (w) and height (h) are predicted relative to the whole image. Finally, the confidence prediction represents the IOU between the predicted box and any ground truth box [17].

At test time we multiply the conditional class probabilities and the individual box confidence predictions, **Pr (Classi|Object) * Pr (Object) * IOU = Pr (Classi) * IOU**, which gives us class-specific confidence scores for each box. These scores encode both the probability of that class appearing in the box and how well the predicted box fits the object [17].

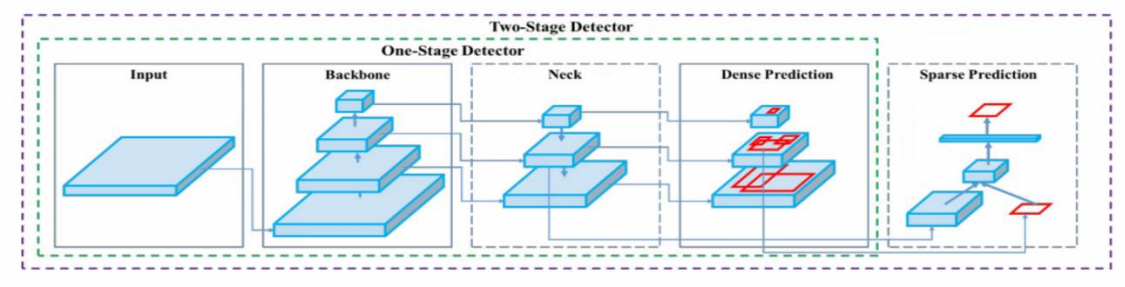
Yolo detection architecture has 24 convolutional layers followed by 2 fully connected layers. Alternating 1 x 1 convolutional layers reduce the features space from preceding layers. The convolutional layers on the ImageNet classification task at half the resolution (224 x 224 input image) and then double the resolution for detection [17].



Picture 17. Yolo algorithm architecture [17]

3.5. ResNet

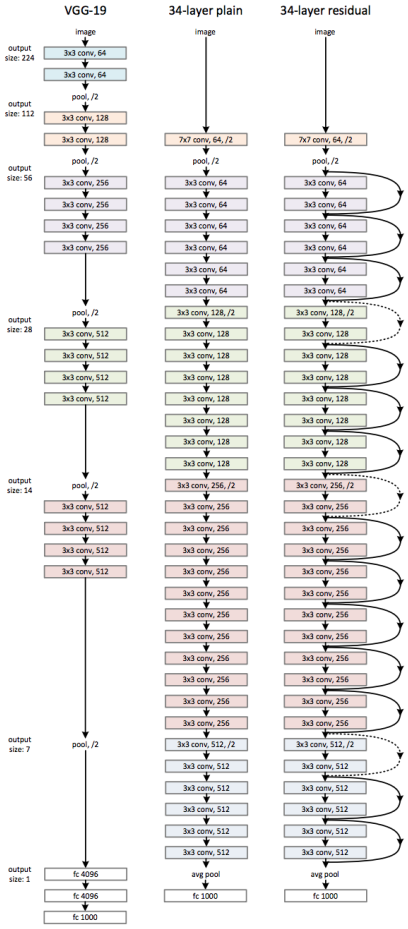
Nowadays object recognition methods are made from 2 parts: backbone, which is from get go trained using “ImageNet”, and head, which is used for object sensing. Normally backbone could be “VGG”, “ResNet” or “ResNeXt” if it used with GPU, if used with CPU, the backbone could be “SqueezeNet”, “MobileNet” or “ShuffleNet”. Usually head part is divided to 2 types: single or dual stages object recognition method. Most popular dual stages object recognition method is “R-CNN”, “fast R-CNN", “faster R-CNN" or etc. Single stage object recognition method would be “YOLO”, “SSD” and “RetinaNet”. Last few years created object recognition algorithms usually inserts few more layers between backbone and head, which is called neck. Neck is made of “from bottom to top” connections and several “from top to bottom” connections. Neck operation includes “Feature Pyramid Network” (“FPN”), “Path Aggregation Network” (“PAN”), “BiFPN” and “NAS-FPN” [18].



Picture 18. Methods architecture [18]

Residual Networks ("ResNet") is a classic neural network used as a backbone for many computer vision tasks. The fundamental breakthrough with ResNet was it allowed us to train extremely deep neural networks with 150+ layers successfully.

ResNet network uses a 34-layer plain network architecture inspired by VGG-19 in which then the shortcut connection is added. These shortcut connections then convert the architecture into the residual network as shown in the figure below:



Picture 19. ResNet network architecture [20]

3.6. SVM

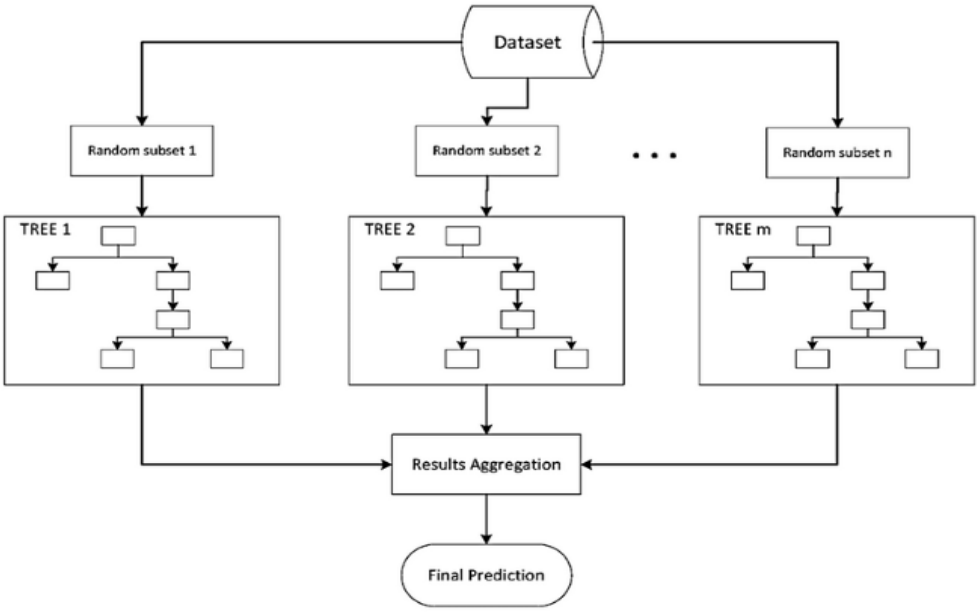
"Support Vector Machines" ("SVMs") are a set of supervised learning methods used for classification, regression and outlier's detection. It uses a technique called the kernel trick to transform your data and then based on these transformations it finds an optimal boundary between the possible outputs. When data is unlabelled, supervised learning is not possible, and an unsupervised learning approach is required. The support-vector clustering algorithm applies the statistics of support vectors, developed in the support vector machine algorithm, to categorize unlabelled data, and is one of the most widely used clustering algorithm in industrial applications. Simply put, it does some extremely complex data transformations, then figures out how to separate your data based on the labels or outputs defined [19].

3.7. Ensemble methods

Ensemble methods is a machine learning technique that combines several base models in order to produce one optimal predictive model. Ensemble methods allow us to take a sample of Decision Trees into account, calculate which features to use or questions to ask at each split, and make a final predictor based on the aggregated results of the sampled Decision Trees [21].

3.7.1 First model

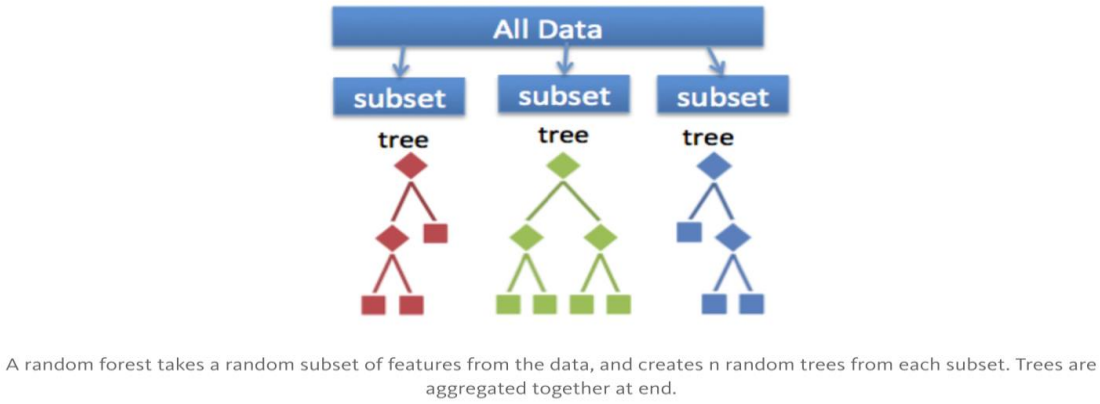
"BAGGing", or "Bootstrap AGGregating". "BAGGing" gets its name because it combines Bootstrapping and Aggregation to form one ensemble model. Given a sample of data, multiple bootstrapped subsamples are pulled. A Decision Tree is formed on each of the bootstrapped subsamples. After each subsample Decision Tree has been formed, an algorithm is used to aggregate over the Decision Trees to form the most efficient predictor [21].



Picture 20. BAGGing model [21]

3.7.2 Second model

"Random Forest" models can be thought of as "BAGGing", with a slight tweak. When deciding where to split and how to make decisions, "BAGGed" Decision Trees have the full disposal of features to choose from. Therefore, although the bootstrapped samples may be slightly different, the data is largely going to break off at the same features throughout each model. In contrary, "Random Forest" models decide where to split based on a random selection of features. Rather than splitting at similar features at each node throughout, "Random Forest" models implement a level of differentiation because each tree will split based on different features. This level of differentiation provides a greater ensemble to aggregate over, ergo producing a more accurate predictor [21].



Picture 21. Random Forest model [21]

IV Proposed Methodology

4.1. Libraries

"Kivy" **[4]** is an open source software library for the rapid development of applications equipped with novel user interfaces, such as multi-touch apps. "Kivy" is fast. This applies to both application development and application execution speeds. It is implemented using "C" level functionality to leverage the power of existing compilers. More importantly, this library uses intelligent algorithms to minimize costly operations, makes use of GPU instead CPU when it is efficient to do so. Another reason to use "Kivy" because of flexibility. "Kivy" can run on a variety of different platforms and devices: "Android", "Windows", "Linux", "OS X". Using this library user can write a simple application with a few lines of code. "Kivy" programs are created

using the "Python" programming language, which is incredibly versatile and powerful, yet easy to use. This language allows user to set up, connect and arrange your application elements quickly, in such a way user can focus on the essence of application rather than forcing to fiddle with compiler settings.

"**OpenCV**" **[3]** is an open source computer vision and machine learning software library. The library has more than 2500 optimized algorithms, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects, produce 3D point clouds from stereo cameras, stitch images together to produce a high resolution image of an entire scene, find similar images from an image database, remove red eyes from images taken using flash, follow eye movements, recognize scenery and establish markers to overlay it with augmented reality, etc. Library is extremely popular, and well-established in big companies such as Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda and Toyota. "OpenCV" has "C++", "Python" **[1]**, "Java" and "MATLAB" interfaces and supports "Windows", "Linux", "Android", and "Mac OS". "OpenCV" is written natively in "C++".

"**NumPy**" **[2]** is one of most popular "Python" libraries, it is mostly used for working with arrays, linear algebra, Fourier transform, and matrices. NumPy advantage is that it is really fast, up to 50x faster than traditional Python lists. The difference in speed comes from that NumPy most parts that requires fast computation are written in "C" language.

"**Android Studio**" is an open source integrated development environment (IDE) provided by Google for the development of apps for "Android" devices. "Android studio" provides several tools that lets user to develop and test applications. This includes an Android Virtual Device (AVD) that emulates various models of Android devices. Native Android apps can be developed in Java or Kotlin.

"**Pandas**" [5] is one of the "Python" libraries that lets user create fast, flexible and powerful tables, also called "Data frames"[6], as well as to create charts to display results.

4.2. Project Plan

- I will find datasets that are appropriate to my project, which is photos with people wearing masks, photos with people without masks and photos with drawn masks on top of the people faces (artificial databases).
- Research and analyse face masks wearing detection methods, compare similar models and identify their shortcomings.
- Designing and implementing a mobile application that real-time (live) detects if a person is wearing a mask and if person wearing it correctly.

- Identify the most efficient and accurate method to be used for face mask wearing detection.
- Identify average error rate of the mobile application.
- Comparison of data and summary of findings.

5. Conclusions

I have researched about the model which can surveil the area through the real-time camera, without any additional devices. The proposed system is a simple real-time video analyser. It has the capability to check whether the individuals wear masks or not. It can be installed in any stores and public gathering places. This helps us to tackle with the widespread of COVID-19 virus. Wearing masks reduces the community spread of COVID-19 virus. We can use this system to check and verify if all the customers are wearing face mask as well as the system thoroughly checks people who enter through the main gate. We can process the video records and find out whether the individual is wearing a face mask or not. If that specific individual wears his/her face mask, the door will open; otherwise it may say some error command similar to "please wear your face mask" or “please reposition your face mask correctly”.

After reviewing the subject area, similar apps and face recognition technologies, I have observed that not wearing face masks during pandemic, flu virus and air pollution can have a major impact affect human health. Many applications/models of a similar nature can only determine if a person is wearing face mask, which does not solve the problem. It has also been found that the YOLO method because of its speed and accuracy could perfectly cope with the recognition and classification of wearing face masks task. Which is why I will be trying to make model using YOLO method.

References

[1] Poole, D. L., & Mackworth, A. K. (2017). Python code for Artificial Intelligence: Foundations of Computational Agents. Version 0.7, 6.

[2] "NumPy". Link https://www.w3schools.com/python/numpy/numpy_intro.asp

[3] "OpenCV About". Link https://opencv.org/about/

[4] "Kivy Documentation". Link https://buildmedia.readthedocs.org/media/pdf/kivy/latest/kivy.pdf

[5] McKinney, W., & Team, P. D. (2015). Pandas-powerful Python data analysis toolkit. Pandas - Powerful Python Data Analysis Toolkit, 1625.

[6] "Intro to pandas". Link https://pandas.pydata.org/pandas-docs/stable/getting_started/index.html

[7] “Neural networks”. Link https://www.ibm.com/cloud/learn/neural-networks

[8] “Convolutional Neural Network”. Lubis, R. MACHINE LEARNING (CONVOLUTIONAL NEURAL NETWORKS) FOR FACE MASK DETECTION IN IMAGE AND VIDEO.

[9] "RMFD". Link https://github.com/X-zhangyang/Real-World-Masked-Face-Dataset

[10] Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2020). A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. Measurement, 167, 108288

[11] Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2020). Fighting against COVID19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. Sustainable Cities and Society, 102600

[12] Lubis, R. MACHINE LEARNING (CONVOLUTIONAL NEURAL NETWORKS) FOR FACE MASK DETECTION IN IMAGE AND VIDEO.

[13] Robot Pepper. Link Pepper the humanoid and programmable robot | SoftBank Robotics

[14] Adnane Cabani, Karim Hammoudi, Halim Benhabiles, and Mahmoud Melkemi, "MaskedFace-Net - A dataset of correctly/incorrectly masked face images in the context of COVID-19", Smart Health, ISSN 2352-6483, Elsevier, 2020, DOI:10.1016/j.smhl.2020.100144

[15] Karim Hammoudi, Adnane Cabani, Halim Benhabiles, and Mahmoud Melkemi,"Validating the correct wearing of protection mask by taking a selfie: design of a mobile application "CheckYourMask" to limit the spread of COVID-19", CMES-Computer Modeling in Engineering & Sciences, Vol.124, No.3, pp. 1049-1059, 2020, DOI:10.32604/cmes.2020.011663

[16] Face Mask Detection application. Link GitHub - chandrikadeb7/Face-Mask-Detection: Face Mask Detection system based on computer vision and deep learning using OpenCV and Tensorflow/Keras

[17] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

[18] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.

[19] “Support Vector Machines”. Link https://scikit-learn.org/stable/modules/svm.html

[20] “ResNet network”. Link What is Resnet or Residual Network | How Resnet Helps? (mygreatlearning.com)

[21] “BAGGing” and “Random Forest”. Link Ensemble Methods in Machine Learning: What are They and Why Use Them? | by Evan Lutins | Towards Data Science

[22] “Image Segmentation”. Link https://viso.ai/deep-learning/image-segmentation-using-deep-learning/

[23] COWLING, B. J., Zhou, Y. D. K. M., Ip, D. K. M., Leung, G. M., & Aiello, A. E. (2010). Face masks to prevent transmission of influenza virus: a systematic review. Epidemiology & infection, 138(4), 449-456.

[24] Wang, C., Horby, P. W., Hayden, F. G., & Gao, G. F. (2020). A novel coronavirus outbreak of global health concern. The Lancet, 395(10223), 470-473.

[25] Feng, S., Shen, C., Xia, N., Song, W., Fan, M., & Cowling, B. J. (2020). Rational use of face masks in the COVID-19 pandemic. The Lancet Respiratory Medicine, 8(5), 434-436

[26] Pascarella, G., Strumia, A., Piliego, C., Bruno, F., Del Buono, R., Costa, F., ... & Agrò, F. E. (2020). COVID-19 diagnosis and management: a comprehensive review. Journal of Internal Medicine.

[27] Wang, J., Pan, L., Tang, S., Ji, J. S., & Shi, X. (2020). Mask use during COVID-19: A risk adjusted strategy. Environmental Pollution, 115099.

[28] Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., ... & Cao, B. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The lancet

[29] Cherrie, J. W., Apsley, A., Cowie, H., Steinle, S., Mueller, W., Lin, C., ... & Loh, M. (2018). Effectiveness of face masks used to protect Beijing residents against particulate air pollution. Occupational and environmental medicine, 75(6), 446-452.