Together for Tomorrow!
Enabling People
Education for Future Generations

AI Course

Team Project Action Plan

For students (instructor review required)

Samsung Innovation Campus

Real Time Face Mask Detection using YOLOv5

04/02/2022

GROUP-06

WAMAV Coders

Ahlam Irshad Shaikh
Azan Altaf
Varuni Gupta
Medha Mohan A P
Wania Zehra

Content

1. Introduction
    1.1. Background Information
    1.2. Motivation and Objective
    1.3. Members and Role Assignments
    1.4. Schedule and Milestones

2. Project Execution
    2.1. Data Acquisition
    2.2.Training Methodology
    2.3. Workflow

3. Results
    3.1. Data Preparation
    3.2. Modeling
    3.3. Final Product

4. Projected Impact
    4.1. Accomplishments and Benefits
    4.2. Future Improvements

5. Team Member Review and Comment

6. Instructor Review and Comment

1. Introduction

1.1. Background Information

The COVID-19 pandemic has caused a global health crisis. In response, the World Health Organization (WHO) has suggested wearing a face mask in public for effective protection. While much of the global population has adhered to these recommendations, some continue to wear the face mask improperly or refuse to wear the mask at all. It is essential that face masks are properly worn in public. To address this, we implemented computer vision, a recent advanced technology, to detect the status of face masks on individuals in crowded public places. Our research is intended to aid in minimizing the spread of coronavirus by developing technology for authorities to discern if face masks are being worn properly. The proposed mask status detection system can aid in reducing the spread of COVID-19 if deployed in a real-world scenario.

1.2 Motivation and Objective

Since the breakout of the deadly virus, COVID -19 has become the worst health problem that has affected the world population. The virus mutated and formed variants that are even more dangerous. Apart from being vaccinated, it is also necessary to stay away from this infectious virus. However, wearing a face mask that inhibits the spread of droplets in the atmosphere can still help combat this pandemic. Therefore, we aim to develop a real-time face mask detector using deep learning techniques that can be deployed for public use. The design uses the latest Yolov5 and PyTorch for the implementation. The work mainly detects three types of wearing face masks; Person wearing the mask, Person not wearing the mask, and Person not correctly wearing the mask. We believe that this work can reduce the spread of the virus by notifying the public about the proper usage of masks if installed at appropriate places.

1.3 Members and Role Assignments

1. Research on Latest Algorithms and Methods - Medha
2. Data Acquisition and Preprocessing - Azan
3. Custom Model Training by all Members
4. Real-Time Implementation - Varuni
5. Reports Documentation and Presentation - Wania and Ahlam

1.4 Schedule and Milestones

Task	End Date
Research	20th Jan 2022
DataAcquisition & Preparation	29th Jan 2022
Custom Training	30th Jan 2022
R-T Implementation	30th Jan 2022
Report & PowerPoint	3rd Feb 2022

2. Project Execution

2.1 Data Acquisition

The Data set used is a collection of images with people either wearing masks properly or improperly or not wearing at all. 60% of the data set was taken from kaggle.com which was already annotated in ‘xml’ format and was later transformed into ‘txt’ format to be used by the YOLOv5 Model. The remaining 40% Data set was only images also taken from Kaggle but were manually annotated and labeled using Roboflow.

2.2 Training Methodology

The Data set was divided into Train/Validate/Test in a ratio of 70%/20%/10%. The Training, Validation, Testing was done using the Yolov5 ‘train.py’ file pre included in the Yolov5 model. The batch size used was 16 and the number of epochs used were 200.

2.2.1 About YOLOv5

- YOLO – You Only Look Once
- YOLO (1-5)
- The algorithm is used in many applications because of its faster detection and accuracy.
- YOLO-v5 is an open-source project by Ultralytics that constitutes a group of several object detection models and techniques based on the YOLO model, pre-trained on the COCO dataset with 80 pre-labelled classes.
- Uses PyTorch

- They can be used for real-time object detection based on the data streams.
- They require very few computational resources.

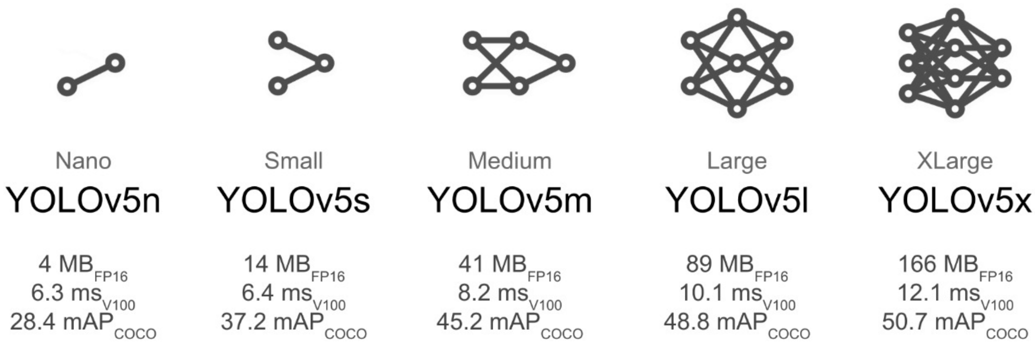


Figure 1: YOLOv5 Model Options

2.3 Workflow

The work flow included:
- Installing dependencies required by YOLOv5
- Downloading annotated data from Roboflow
- Reviewing the number of classes
- Training the data
- Custom loading the model
- Testing on a random image from google.com
- Implementing real time mask detection using webcam

3. Results

3.1. Data Preparation

3.1.1 Handling Imbalanced Data

The datasets used for training the real-time face mask detector were obtained from Kaggle. Using only one dataset resulted in an imbalanced model as the number of instances per class heavily varied from class to class:

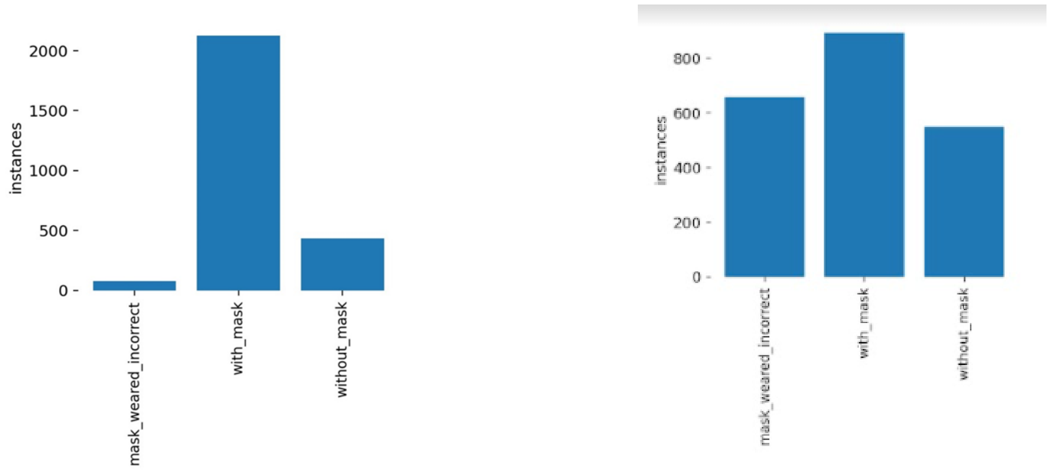


Figure 2: Data before and after balancing

Hence, the second additional dataset was used to feed in the balanced number of instances per class. This aims to provide reliable unbiased results.

One of the datasets already consisted of annotations in XML format. These needed to be converted to TXT, this was done using Roboflow. The other dataset consisted of unannotated images, which were annotated and exported in YOLOv5 txt format on Roboflow.

```
0 0.4963942307692308 0.4963942307692308 0.9927884615384616 0.9927884615384616
```

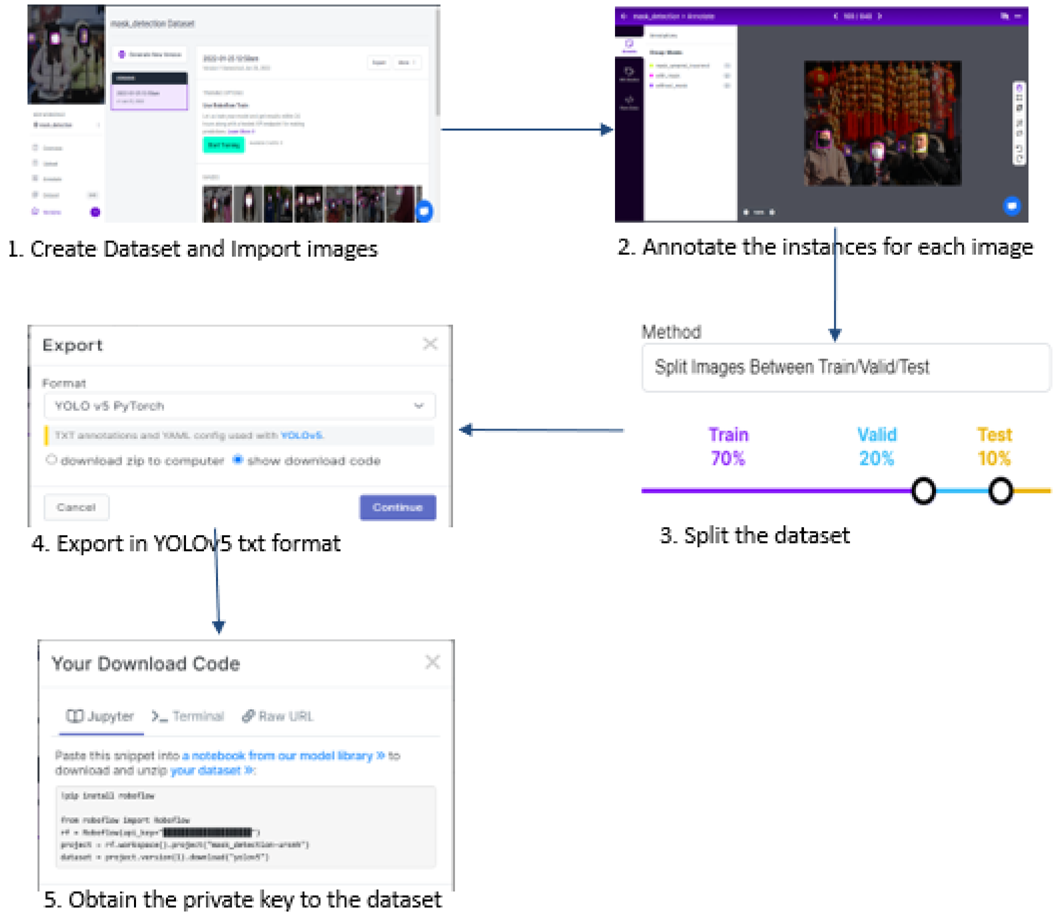
Figure 3: Snippet of label txt annotation

As seen in the figure above, the Yolov5 darknet txt format results in an annotation text file for each image. This text file consists of the annotations and a numeric representation of the label as well as a mapping of the label which is used to map the numeric IDs to strings that are readable by humans. The resulting numeric labels for the classes are as shown below:

Class Name	Numeric Label
mask_weared_incorrectly	0
with_mask	1
without_mask	2

3.1.2 2. Data annotation on Roboflow

The Roboflow online platform was used in order to annotate the images in YOLOv5 txt format. The steps and results of the annotation process are shown below:



The final dataset, after the preprocessing step, consists of approximately 600 instances for each class.

3.2 Modeling

Customized training on the YOLOv5 object detection algorithm was used in order to build the mask detection model. This model has generated several output visualizations that can be used to interpret the model training and performance. These outputs are defined below:

3.2.1 Results of the YOLOv5 Model Training & Validation through the epochs

The YOLOv5 model was trained for 200 epochs, for which the results are shown in the figure below. These graphs show the improvement in our model through the 200 epochs, which display different performance metrics for both the training and validation sets.

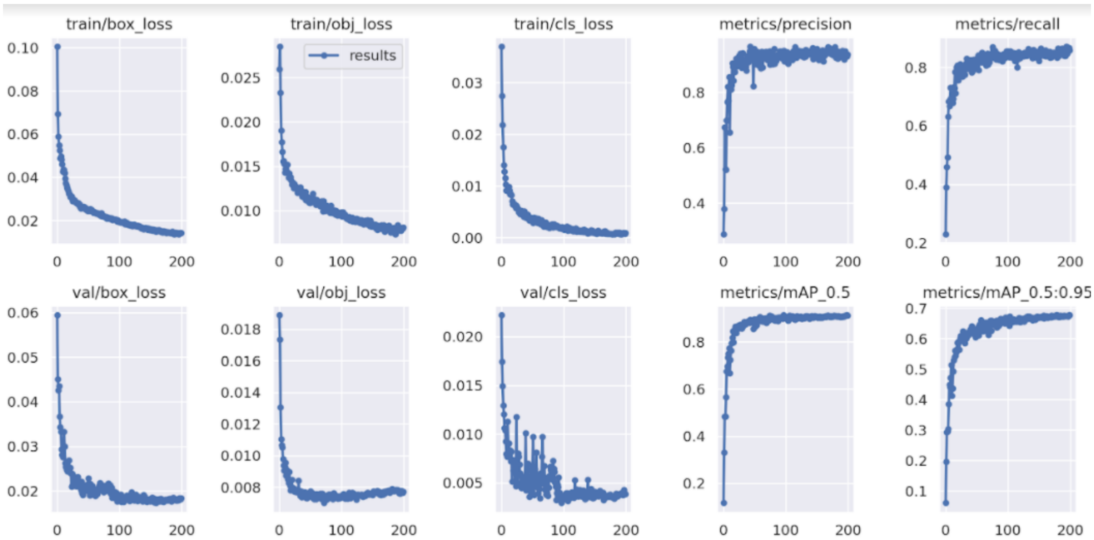


Figure 4: Model training and validation results

The box loss denotes the performance of the algorithm with reference to how well it can locate the center of an object and how well the predicted bounding box covers an object. The objectness loss represents the probability that an instance exists in a proposed region of interest. A high objectivity shows that the image is likely to contain an object. Lastly, the classification loss gives knowledge of the algorithm predicting the right class of a given instance. The box, objectness and classification losses of the validation data have shown a drastic drop in the first 50 epochs. Moreover, the precision, recall and mean average precision has drastically improved before plateauing (at almost 0.9) after around 50 epochs.

3.2.2 PR Curve

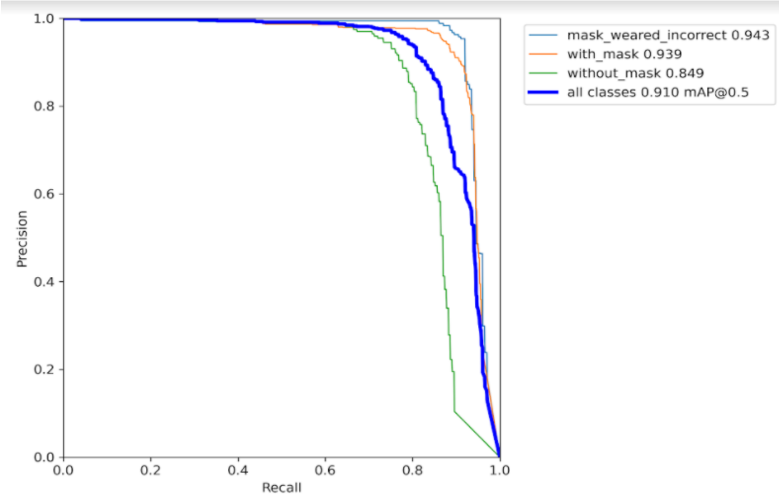


Figure 5: PR Curve

The precision-recall curve is used to obtain an ideal threshold for the precision and recall of the model classes, in which:

- The precision is the number of positive class predictions which are actually the positive class.

$$precision = PPV = \frac{TP}{TP+FP}$$

- The recall is the number of positive class predictions from all the positive class instances in the dataset.

$$recall = sensitivity = \frac{TP}{TP+FN}$$

The best PR curve threshold is one that results in the precision and recall being closest to 1. Therefore, from the key of the PR curve shown in the figure above, it can be seen that the precision-recall threshold is fairly high for all of the classes (mAP@0.5 0.91 for all the classes).

3.2.3 Confusion Matrix

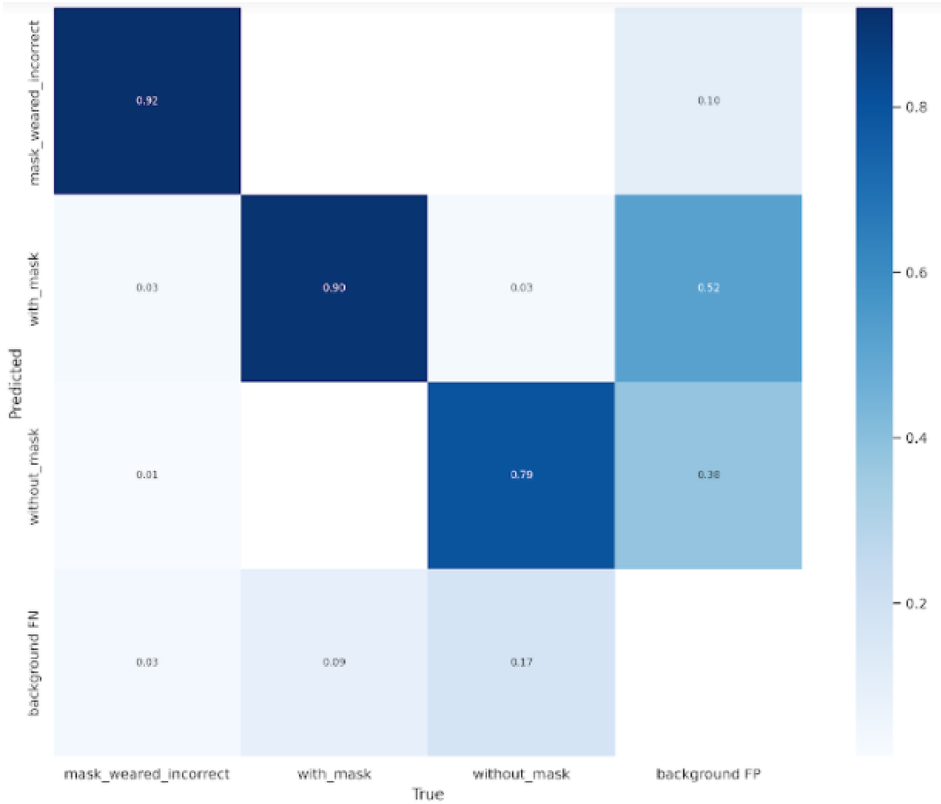


Figure 6: Confusion Matrix

A confusion matrix is used to evaluate the performance of a classification model. The figure above depicts that the true positive rate is high (0.92 for mask wear incorrect, 0.90 for with mask and 0.79 for without mask) for all three classes. This rate defines the proportion of the predicted classes that were actually those classes.

3.3 Final Product.

The final product of this project involves the real-time classification of whether a person is wearing a mask, not wearing a mask or incorrectly wearing a mask. Prior to the real-time video implementation, it is essential to test the classifier with images outside from the dataset that was used to train, test and validate the model. The image below shows that this classifier has correctly identified the target class for each instance with a fairly high confidence.

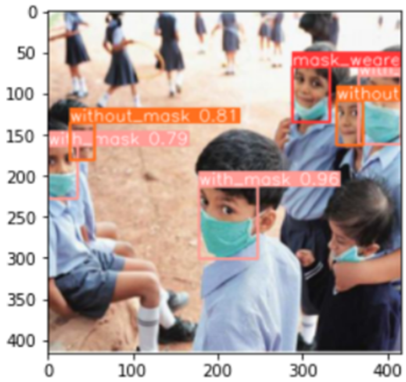
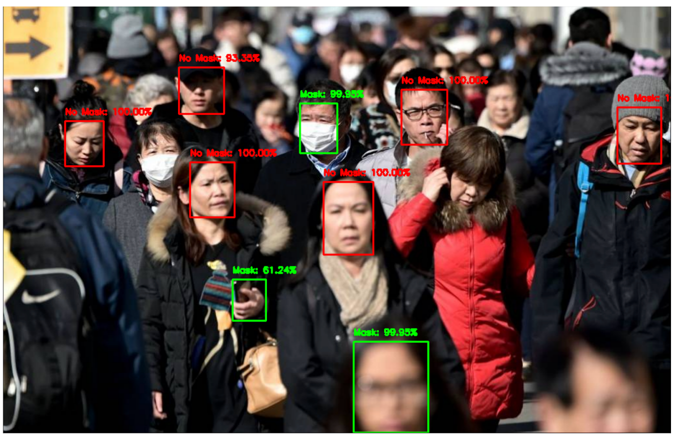


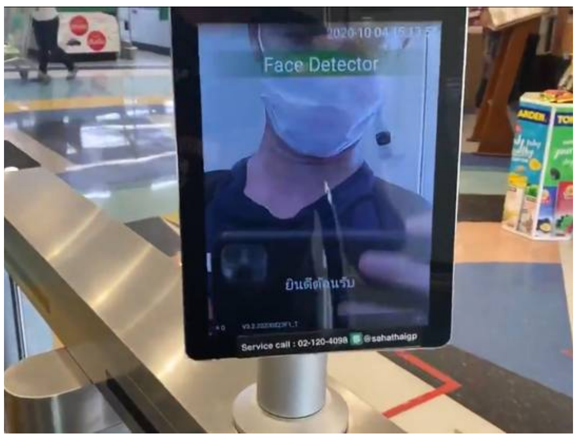
Figure 7: Final Output Classification Example

4. Projected Impact

4.1. Accomplishments and Benefits

The YOLOv5 algorithm helped attain a highly-accurate and fast face mask detection model, given its blazing fast inference speed. Through training and testing, our model achieved a success rate of about 85%. Most models detect either with a face mask or without a face mask but this model detects incorrectly worn masks as well. Since it is trained using a pre-existing algorithm, it is easy to use, cheaper to set up and faster to implement. By using this model at the entrance of stores, hospitals, schools, colleges and offices, we will be able to curb the spread of the coronavirus. It will help reduce exposure distances and implement effective control over the public.

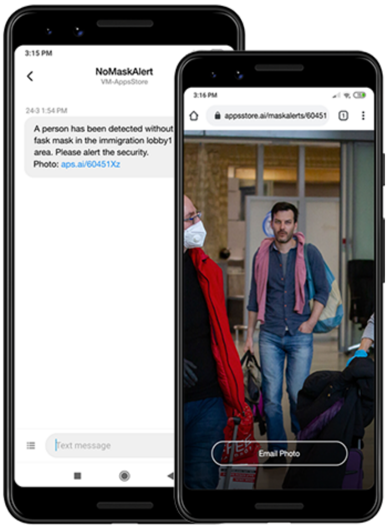




4.2 Future Improvements

Since the model was trained using bright images, an improvement that can be done would be to train the model with dark images as well so that it can work equally well in dim-lit places. Another improvement could be to enhance the model so that it can detect the presence of face masks on those faces that are not directly facing the camera. This Real-Time Face Mask Detector can be integrated into CCTVs at public places such as hospitals, malls, colleges, airports, public events, movie theaters, offices, and sidewalks. It can identify the people that are not wearing masks or are wearing it incorrectly and give them intimation warnings. This can be implemented by using the face of the person to obtain their name, mobile number and other details which are stored in that state/region's databases. After a few warnings, a fine can be issued to that person.

Moreover, it can be created into an application for the local governments and police stations. It will notify them about the number of people wearing face masks and not wearing face masks and will help them take the necessary steps/actions into controlling the spread of the coronavirus.



5. Team Member Review and Comment


Ahlam


Azan


Varuni


Medha


Wania

NAME	REVIEW and COMMENT
Ahlam Irshad Shaikh	Thankful for the opportunity of exploring new tools and techniques with this great team. Wonderful experience of cooperation and teamwork.
Azan Altaf	Working on this project with the whole team was a delightful experience. Everyone participated and was actively working on the project.
Varuni Gupta	With the help of an amazing team that is very supportive, this project was a fun learning experience. All members are extremely hard-workers and a pleasure to work with.
Medha Mohan AP	Extremely happy to work with this wonderful team. Considering myself fortunate to be a part of this program and to acquaint myself with these incredibly talented and hardworking team members.
Wania Zehra	Had a great experience working with my team. Everyone was very supportive and cooperative.

6. Instructor Review and Comment

CATEGORY	SCORE	REVIEW and COMMENT
IDEA	__/20	
CODING	__/20	

PROJECT MANAGEMENT	___/30	
PRESENTATION & REPORT	___/30	
TOTAL	___/100	