# 3. MATERIAL AND METHOD

## *3.1 Data Collection*

In this study, we collected a total of 1805 images that belong to four classes: no mask, fabric/cloth mask, surgical mask and filtering facepiece (FFP) mask.

|  |  |
| --- | --- |
| TABLE 1. NUMBER OF IMAGES PER CLASS IN DATASET | |
| **Class** | **Number of Images** |
| No Mask | 455 |
| Fabric/Cloth Mask | 404 |
| Surgical Mask | 523 |
| FFP Mask | 423 |

## *3.2 Data Annotation*

Collected dataset are then manually annotated for image localization. This is to determine the location a single object in a given image. The image localization implements a bounding box, a rectangle that encloses an object in an image. The bounding box are in YOLO format. Sample images with bounding box are shown in Figure 1.

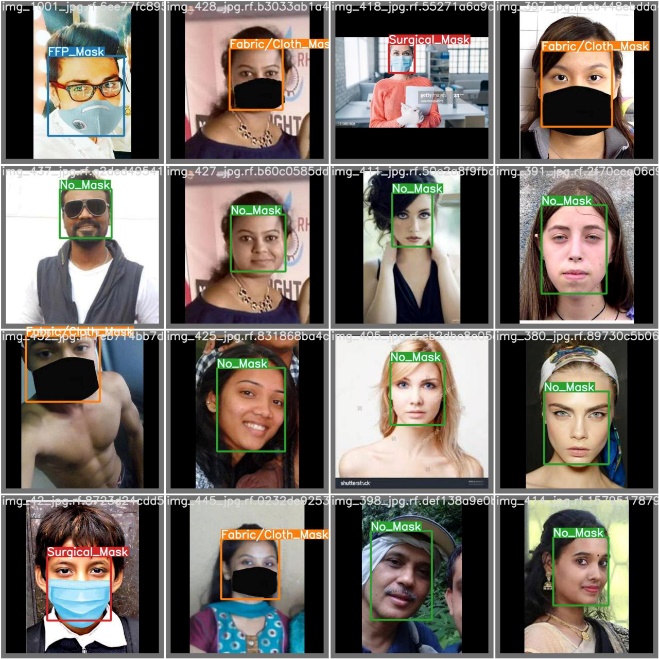


Fig. 1. Sample images with bounding box.

## *3.3 Data Pre-Processing*

Transformations of the dataset’s attributes are applied to ensure uniform learning and inference prior to training, validation and testing. Collected images are resized to 416×416 and to maintain the source image aspect ratio, black edges are fitted in.

## *3.4 Data Augmentation*

For better outcomes, data augmentation is required. The model accuracy will improve as we have a larger variety of data. The goal of data augmentation is to increase the variability of the input photos, resulting in a more robust object detection model [1]. Augmentation in the algorithm YOLOv4 is directly performed using various methods called CutMix, Mosaic Data Augmentation, Class Label Smoothing and Self-Adversarial Training (SAT) [2].

### *3.4.1 CutMix*

### *3.4.2 Mosaic Data Augmentation*

### *3.4.3 Class Label Smoothing*

### *3.4.4 Self-Adversarial Training (SAT)*

## *3.5 Darknet Framework*

## *3.6 YOLOv4-Tiny Algorithm*

### *3.6.1 Backbone Network – Feature Formation*

### *3.6.2 Neck – Feature Aggression*

### *3.6.1 Head – Detection Step*

## *3.7 Model Training Setup*

In this study, all of the dataset’s training in this project were carried out in Google's Jupyter notebook Colab using the provided free GPU. The GPU type is Tesla T4, NVIDIA-SMI 495.46, Driver Version: 460.32.03 and CUDA Version: 11.2.

The datasets are divided into three sets. For training, 70% of the dataset were used (1263 images) in training the dataset using YOLOv4-tiny algorithm. Another 20% of the dataset were used for validation (361 images). The validation set is used to assess the training using new data. A final set using 10% of the dataset were used for testing the accuracy of the trained model in detecting facemasks (181 images).

YOLOv4-tiny hyper-parameters were set as follows: Due to variances in the size of the images in the dataset, the image input size was set to 416×416 pixels to make detection easier. The cosine annealing decay method is used in learning rate. The initial learning rate was set at 0.00261 and the final rate was set at 0.1. At 6400 steps and 7200 steps, respectively, a factor of 0.1 was multiplied. The training employed a single GPU with batch size 64 to perform multi-scale training. As proposed by the creators of YOLOv4-tiny, the study also used the default momentum of 0.949, IoU threshold of 0.213, and loss normalizer of 0.07.

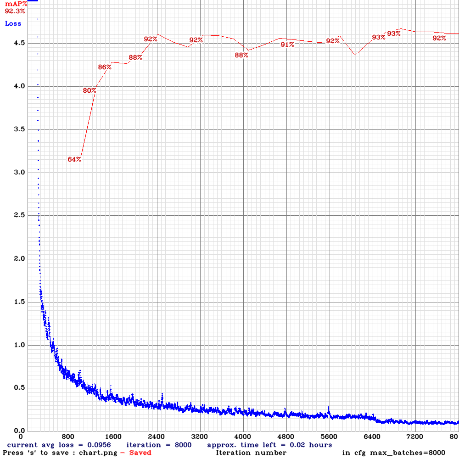
|  |  |
| --- | --- |
| TABLE II. PARAMETERS FOR MODEL TRAINING | |
| **Parameters** | **Value** |
| Input Size | 416×416 |
| Initial Learning Rate | 0.00261 |
| Batch Size | 64 |
| Iterations | 8000 |

## *3.8 Trained Model Implementation*

## *3.9 Evaluation Indicators*

# 4. EXPERIMENTAL RESULTS

The performance of the model was evaluated using the 181 images from the test dataset to determine the performance of the trained model in detecting facemask types.



Results are presented in Table III.

|  |  |  |
| --- | --- | --- |
| |  | | --- | | TABLE III. RESULTS | | |
| **Calculation** | **Score** |
| Precision | 89% |
| Recall | 90% |
| F-1 Score | 90% |
| Average IoU | 70.27% |
| mAP | 92.28% |

As seen in table III, gwapa ko,

Another table for the ap

|  |  |  |  |
| --- | --- | --- | --- |
| TABLE IV. AVERAGE PRECISION VALUES | | | |
| **Class** | **TP** | **FP** | **AP%** |
| No Mask | 105 | 8 | 94.91 |
| Fabric/Cloth Mask | 67 | 11 | 87.37 |
| Surgical Mask | 106 | 20 | 88.99 |
| FFP Mask | 62 | 2 | 97.84 |

TP means teleport

# 5. CONCLUSION