UNIVERSITÉ Concordia UNIVERSITY

Applied Artificial Intelligence

COMP 6721

AI Face Mask Detector

Jay Patel 40197003 - Data Specialist
Krupali Bhatt 40184668 - Training Specialist
Juhi Birju Patel 40190446 - Evaluation Specialist

Index

1. Dataset

The dataset used in this project is a combination of images taken from Kaggle[1] and other references as cited below[2].

The Kaggle dataset contained 4 folders: no mask, cloth mask, N95, and surgical mask. Based on the feedback received for the Build-1, We made the data set balanced, pre-augmentation. This was one of the major changes executed in Build 2 of the project.

Earlier We had 169 unique images of N95 with valve and we had performed image augmentation including resizing, flipping, cropping, brightness, saturation resulting into a total of 900 images. In the second phase of this project, for the fourth category which was N95 with a valve, we collected 328 unique images from a couple of kaggle[3][4] sources, and manually searching google images[5].

We performed image augmentation on this dataset by writing a python script, which included cropping, flipping and resizing images in order to improve the learning process for our model.

The dataset is balanced with 2054 RGB images in 5 folders named

- Cloth_mask
- ffp2 mask
- N95_with_valve
- surgical_mask
- and no_mask.

To perform an evaluation for a possible bias in the model, the attributes of age and gender have been used. To analyse gender bias of male and female classes have been considered and for the analysis of age bias three classes namely young, middle and old have been taken into account. The dataset for these two attributes has been taken from kaggle[6][7].

Given below are the statistics of our dataset:

Class	Number of Images
Cloth Mask	400
Surgical Mask	408
N95 Mask	400
N95 with a valve	414
No mask	432
Total	2054

Dataset Distribution for Gender Bias:

A) Gender Bias: MALE

Class	Number of Images
Cloth Mask	198
Surgical Mask	202
N95 Mask	220
N95 with a valve	150
No mask	230
Total	1000

B) Gender Bias: FEMALE

Class	**Number of Images**
Cloth Mask	202
Surgical Mask	206
N95 Mask	180
N95 with a valve	264
No mask	202
Total	1054

C) Overall

Class	**Number of Images**
Male	**1000**
Female	**1054**
Total	**2054**

Dataset Distribution for Age Bias:

D) Age Bias: CHILD

Class	Number of Images
Cloth Mask	50
Surgical Mask	192
N95 Mask	64
N95 with a valve	118
No mask	76
Total	500

E) Age Bias: YOUNG

Class	Number of Images
Cloth Mask	245
Surgical Mask	130
N95 Mask	148
N95 with a valve	160
No mask	210
Total	893

F) Age Bias: OLD

Class	Number of Images
Cloth Mask	105
Surgical Mask	86
N95 Mask	188
N95 with a valve	136
No mask	146
Total	661

G) Overall

Class	**Number of Images**
Young	**500**
Middle	**893**
Old	**661**
Total	**2054**

2. CNN Architecture

2.1 Model Design

- Input of [3*width*height] is feed into the first convolution layer, where 3 is the number of channels i.e. RGB.
- There are 5 convolution layers with a kernel size of (3,3), stride of (1,1) , and a padding of (1,1).
- The convolution layer is followed by a 2 Dimensional Batch Normalisation.
- Followed by the batch normalisation, is the ReLU Activation function.The rectified linear activation function (RELU) is a piecewise linear function that, if the input is positive say x, the output will be x. Otherwise, it outputs zero.
- There are two pooling layers , each after every 2 convolution layers. The pooling used is max pooling with a kernel of (2,2) and a stride of 2.
- We have used a drop out of 0.02 in order to prevent the model from overfitting.
- Finally, the output from the final pooling layer is flattened and fed as an input to the fully connected layer. Neurons in a fully connected layer are connected to all the activations in the previous layer. Therefore, it helps to classify the input into predicted classes.

2.2 Data Preparation for the Model

- The balanced dataset is fetched , shuffled and saved into .npy file for faster processing of the images.
- Images in the dataset are transformed to tensor objects, they are normalised and resize to 32 X 32, normalised and converted to tensor.
- We have used Stratified KFold for the train test split. K fold cross-validation divides the dataset into k folds however the splits lack proper distribution of classes and hence Stratified K fold is used to ensure that each fold of dataset has the same proportion of dat with a given class
- An equal proportion of all the five classes are taken for the purpose of testing to generate the training and testing confusion matrix.

2.3 Configuring the optimizer

- We have used the Adam optimizer, instead of the classical stochastic gradient descent procedure and the learning rate is set to 0.001, to update network weights iteratively.

2.4 Model Training

- We trained the model for various folds. We have used 10-folds here. Below given are the accuracy for each fold
- After each epoch, We print the training accuracy and the cross entropy training loss and we see that the accuracy increases and the loss decreases.
- The optimizer.zero_grad() function sets the gradients to zero.
- The backward() function is called on the loss to calculate back propagation.
- The optimizer.step() iterates over all the parameters and updates them using their internally stored grad values.

3. Bias Detection and Elimination:

1) Gender Bias :

After running the dataset of male masked images and female masked images individually on the dataset a disparity was observed in the accuracy of the model. The model was 35% accurate for the males and 27% accurate for the females indicating a selection bias in the attribute gender for the female class.

The model is now 36% accurate for the males and 33% accurate for the females.

2) Age Bias:

The results of running the dataset of child, young and old aged masked images on the previous model showed accuracies of 15%, 23%, and 18% respectively indicating a little lower representation of old aged masked images in the previous complete dataset.

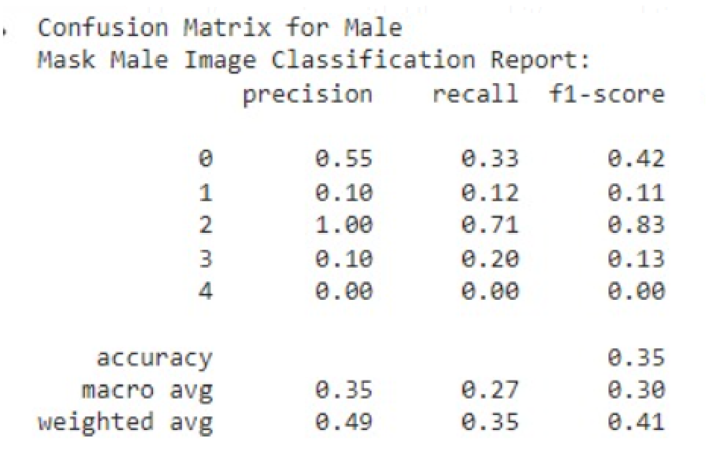
The results of running the dataset of child, young and old aged masked images on the current model shows accuracies of 23%, 38%, and 30% respectively.

Bias Elimination

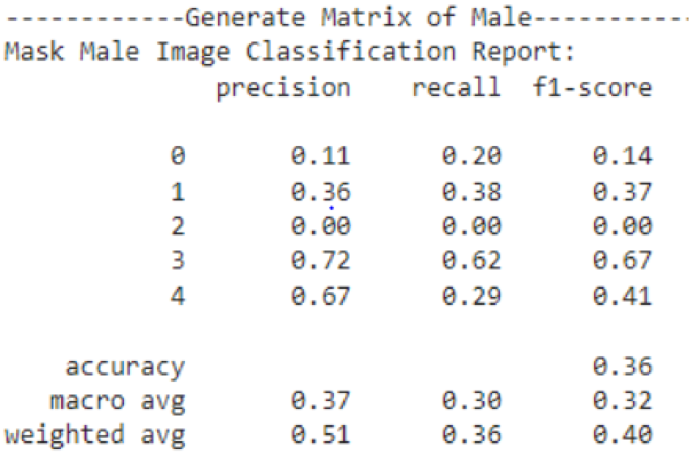
- ❖ To eliminate the bias, more images of male and female masked images have been added to create a balance between the two gender in the complete dataset.
- ❖ For the elimination of bias in age wise distribution , more images of middle and old aged people wearing masks have been added for a balanced dataset.
- ❖ Also to increase the accuracy, one more convolution layer has been added in the model. The model has been retrained using the new dataset and new model and shows significant improvement.

Comparing Precision Recall and F-1 Score for all the subclasses of old Model Vs New Model:

A. Comparing Gender Bias: Male



```
Confusion Matrix for Male
Mask Male Image Classification Report:
              precision    recall  f1-score

           0       0.55      0.33      0.42
           1       0.10      0.12      0.11
           2       1.00      0.71      0.83
           3       0.10      0.20      0.13
           4       0.00      0.00      0.00

    accuracy                           0.35
   macro avg       0.35      0.27      0.30
weighted avg       0.49      0.35      0.41
```

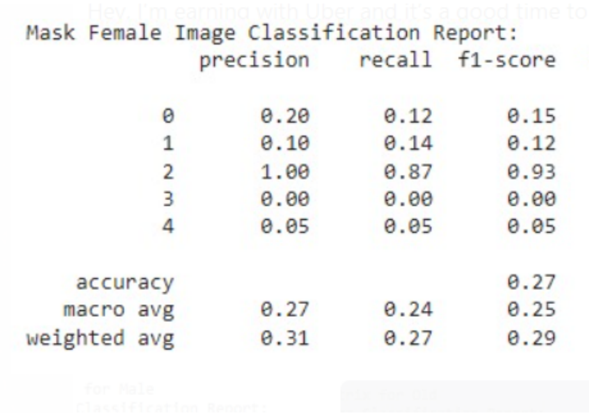


```
------------Generate Matrix of Male----------
Mask Male Image Classification Report:
              precision    recall  f1-score

           0       0.11      0.20      0.14
           1       0.36      0.38      0.37
           2       0.00      0.00      0.00
           3       0.72      0.62      0.67
           4       0.67      0.29      0.41

    accuracy                           0.36
   macro avg       0.37      0.30      0.32
weighted avg       0.51      0.36      0.40
```

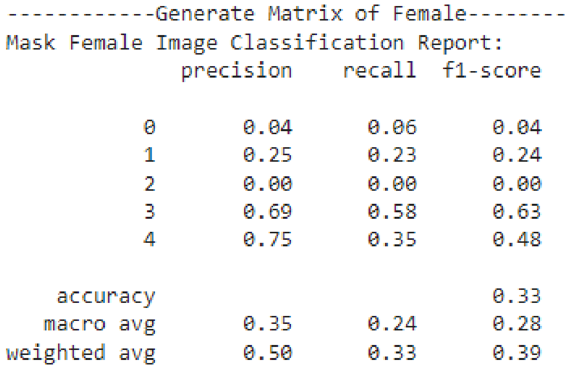
Confusion Matrix of old Model 35%

Confusion Matrix of newModel 36%

B. Comparing Gender Bias: Female



```
Mask Female Image Classification Report:
              precision    recall  f1-score

           0       0.20      0.12      0.15
           1       0.10      0.14      0.12
           2       1.00      0.87      0.93
           3       0.00      0.00      0.00
           4       0.05      0.05      0.05

    accuracy                           0.27
   macro avg       0.27      0.24      0.25
weighted avg       0.31      0.27      0.29
```

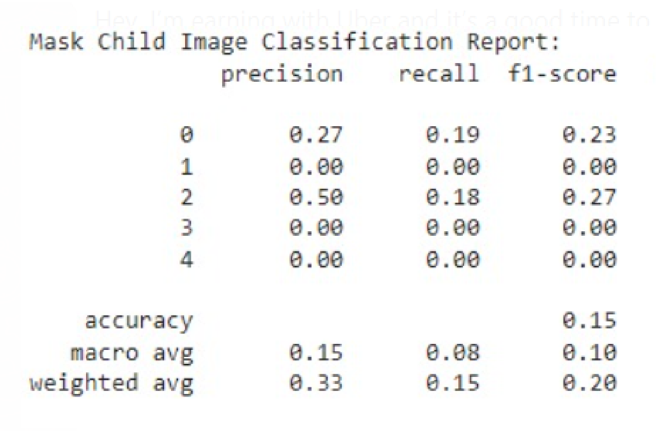


```
------------Generate Matrix of Female--------
Mask Female Image Classification Report:
              precision    recall  f1-score

           0       0.04      0.06      0.04
           1       0.25      0.23      0.24
           2       0.00      0.00      0.00
           3       0.69      0.58      0.63
           4       0.75      0.35      0.48

    accuracy                           0.33
   macro avg       0.35      0.24      0.28
weighted avg       0.50      0.33      0.39
```

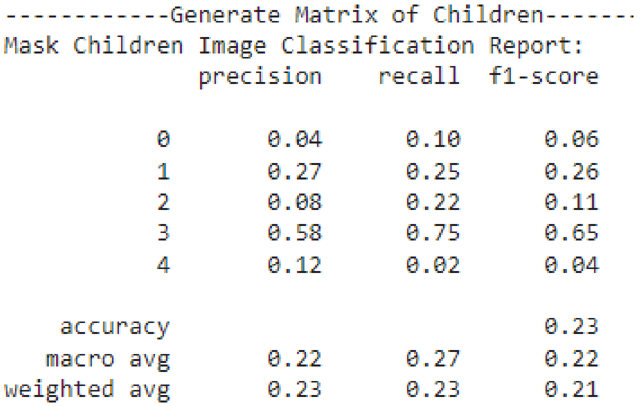
Confusion Matrix of old Model 27%

Confusion Matrix of newModel 33%

C. Comparing Age Bias: Child



```
Mask Child Image Classification Report:
              precision    recall  f1-score

           0       0.27      0.19      0.23
           1       0.00      0.00      0.00
           2       0.50      0.18      0.27
           3       0.00      0.00      0.00
           4       0.00      0.00      0.00

    accuracy                           0.15
   macro avg       0.15      0.08      0.10
weighted avg       0.33      0.15      0.20
```

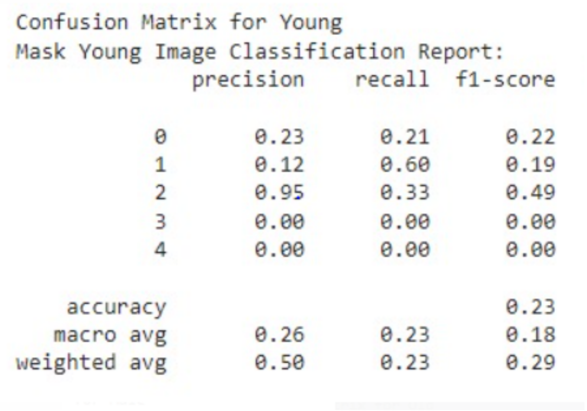


```
------------Generate Matrix of Children-------
Mask Children Image Classification Report:
              precision    recall  f1-score

           0       0.04      0.10      0.06
           1       0.27      0.25      0.26
           2       0.08      0.22      0.11
           3       0.58      0.75      0.65
           4       0.12      0.02      0.04

    accuracy                           0.23
   macro avg       0.22      0.27      0.22
weighted avg       0.23      0.23      0.21
```

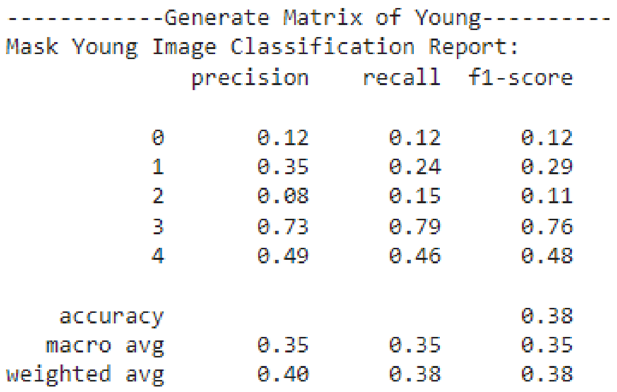
Confusion Matrix of old Model 15%

Confusion Matrix of newModel 23%

D. Comparing Age Bias: Young

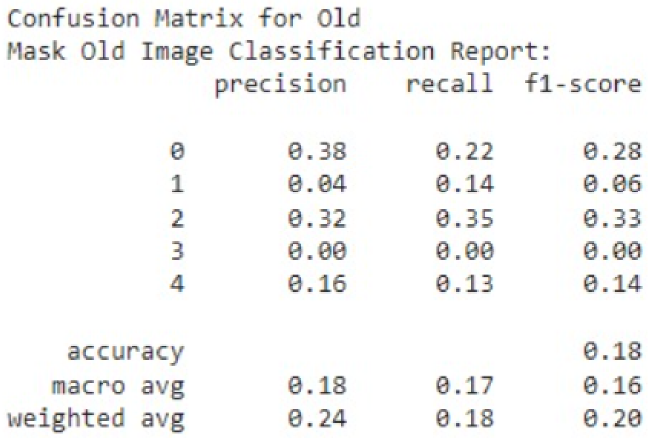


```
Confusion Matrix for Young
Mask Young Image Classification Report:
              precision    recall  f1-score

           0       0.23      0.21      0.22
           1       0.12      0.60      0.19
           2       0.95      0.33      0.49
           3       0.00      0.00      0.00
           4       0.00      0.00      0.00

    accuracy                           0.23
   macro avg       0.26      0.23      0.18
weighted avg       0.50      0.23      0.29
```

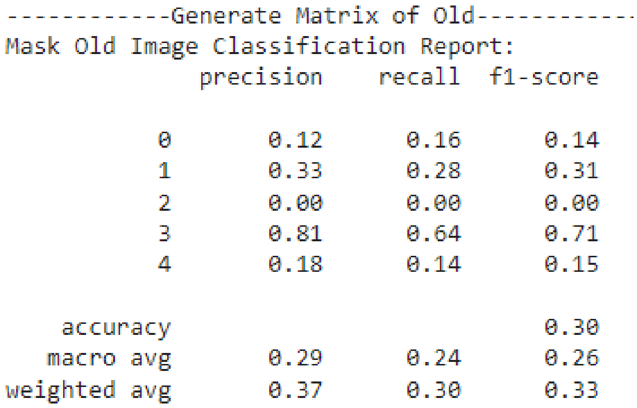


```
------------Generate Matrix of Young----------
Mask Young Image Classification Report:
              precision    recall  f1-score

           0       0.12      0.12      0.12
           1       0.35      0.24      0.29
           2       0.08      0.15      0.11
           3       0.73      0.79      0.76
           4       0.49      0.46      0.48

    accuracy                           0.38
   macro avg       0.35      0.35      0.35
weighted avg       0.40      0.38      0.38
```

E. Comparing Age Bias: Old



```
Confusion Matrix for Old
Mask Old Image Classification Report:
              precision    recall  f1-score

           0       0.38      0.22      0.28
           1       0.04      0.14      0.06
           2       0.32      0.35      0.33
           3       0.00      0.00      0.00
           4       0.16      0.13      0.14

    accuracy                           0.18
   macro avg       0.18      0.17      0.16
weighted avg       0.24      0.18      0.20
```

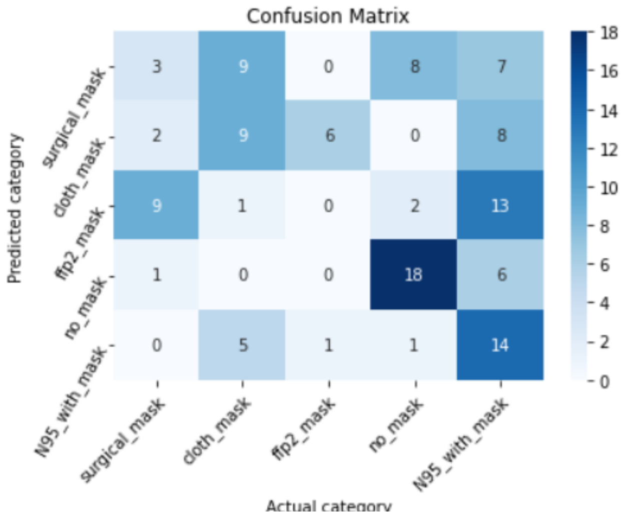


```
------------Generate Matrix of Old-----------
Mask Old Image Classification Report:
              precision    recall  f1-score

           0       0.12      0.16      0.14
           1       0.33      0.28      0.31
           2       0.00      0.00      0.00
           3       0.81      0.64      0.71
           4       0.18      0.14      0.15

    accuracy                           0.30
   macro avg       0.29      0.24      0.26
weighted avg       0.37      0.30      0.33
```

Confusion Matrix of old Model 18% Confusion Matrix of newModel 30%

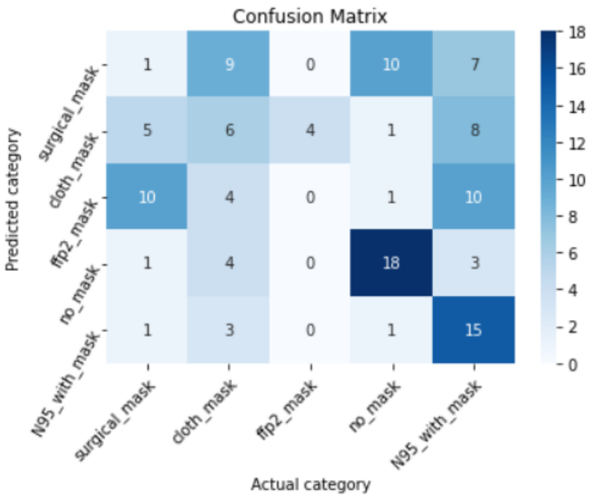
Confusion Matrix for the updated model:

Mask Male Image Confusion Matrix:



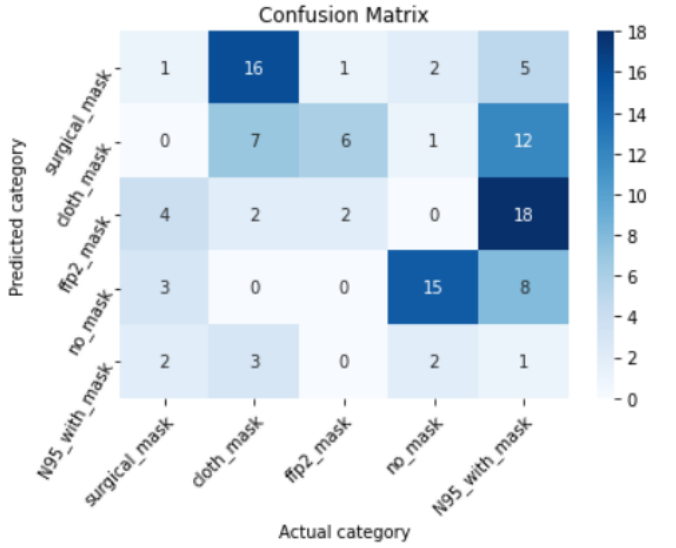
	surgical_mask	cloth_mask	ffp2_mask	no_mask	N95_with_mask
surgical_mask	3	9	0	8	7
cloth_mask	2	9	6	0	8
ffp2_mask	9	1	0	2	13
no_mask	1	0	0	18	6
N95_with_mask	0	5	1	1	14

Mask Female Image Confusion Matrix:

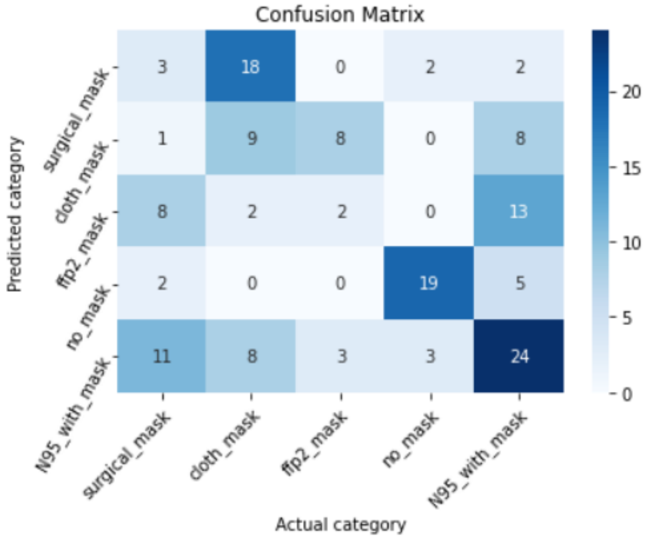


	surgical_mask	cloth_mask	ffp2_mask	no_mask	N95_with_mask
surgical_mask	1	9	0	10	7
cloth_mask	5	6	4	1	8
ffp2_mask	10	4	0	1	10
no_mask	1	4	0	18	3
N95_with_mask	1	3	0	1	15

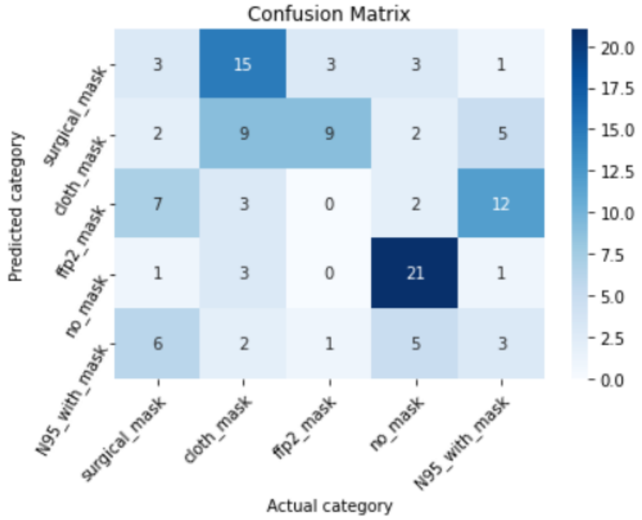
```
Mask Children Image Confusion Matrix:
```



```
Mask Young Image Confusion Matrix:
```



```
Mask Old Image Confusion Matrix:
```



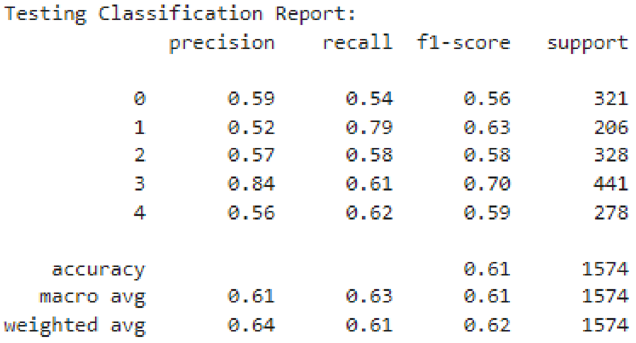
5. K- Fold Cross Validation

Model is evaluated using a stratified k fold strategy on training and testing set. The model is iterated over 10 folds with 10 epochs each and a batch size of 32. The average accuracy, precision, recall and f1-score of these folds are calculated at the end. The confusion matrix is visualised with plot_cm() function. The model is saved as a pickle file, then loaded using torch.load() and trained again. The model is made to run on a random set of images and a sample prediction is done which outputs the result as follows.

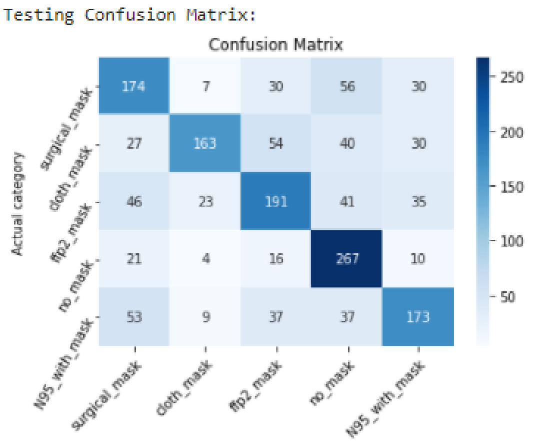
Performance Analysis:

Old Model:

- The dataset used in the old model was unbalanced before augmentation, and it was made balanced after performing extensive augmentation.
- The previous model had 4 convolution layers with a stride of (1,1) and a padding of (1,1). The model is evaluated using fixed training/test split of 70:30 under 10 epochs. After each epoch the accuracy and cross entropy training loss are printed and we observe that with increasing epochs the accuracy increases and the loss decreases.



```
Testing Classification Report:
              precision    recall  f1-score   support

           0       0.59      0.54      0.56       321
           1       0.52      0.79      0.63       206
           2       0.57      0.58      0.58       328
           3       0.84      0.61      0.70       441
           4       0.56      0.62      0.59       278

    accuracy                           0.61      1574
   macro avg       0.61      0.63      0.61      1574
weighted avg       0.64      0.61      0.62      1574
```



Testing Confusion Matrix:

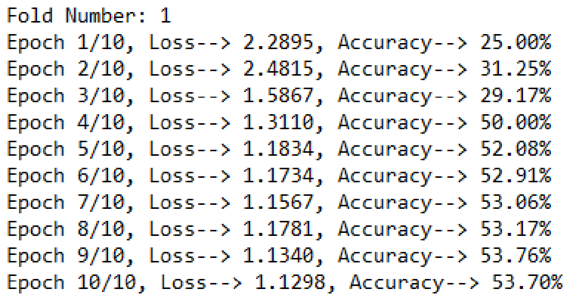
Confusion Matrix

Actual category	surgical_mask	cloth_mask	ffp2_mask	no_mask	N95_with_mask
surgical_mask	174	7	30	56	30
cloth_mask	27	163	54	40	30
ffp2_mask	46	23	191	41	35
no_mask	21	4	16	267	10
N95_with_mask	53	9	37	37	173

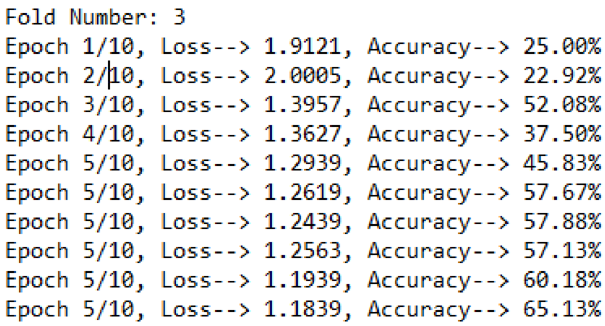
Testing Classification Result for Old Model

Testing Confusion Matrix for Old Model

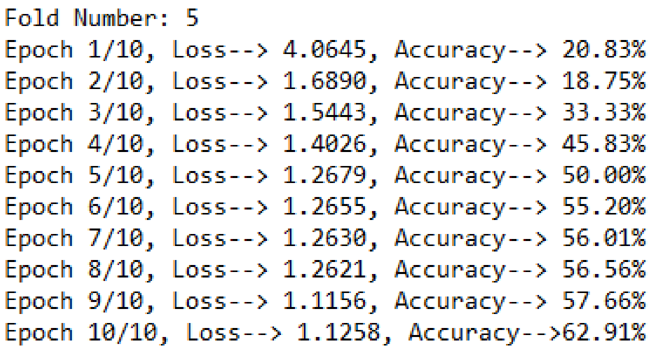
- On running the k-fold on the old method, following result was observed:



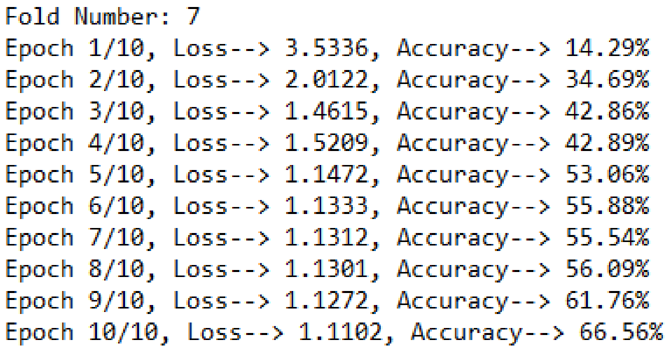
```
Fold Number: 1
Epoch 1/10, Loss--> 2.2895, Accuracy--> 25.00%
Epoch 2/10, Loss--> 2.4815, Accuracy--> 31.25%
Epoch 3/10, Loss--> 1.5867, Accuracy--> 29.17%
Epoch 4/10, Loss--> 1.3110, Accuracy--> 50.00%
Epoch 5/10, Loss--> 1.1834, Accuracy--> 52.08%
Epoch 6/10, Loss--> 1.1734, Accuracy--> 52.91%
Epoch 7/10, Loss--> 1.1567, Accuracy--> 53.06%
Epoch 8/10, Loss--> 1.1781, Accuracy--> 53.17%
Epoch 9/10, Loss--> 1.1340, Accuracy--> 53.76%
Epoch 10/10, Loss--> 1.1298, Accuracy--> 53.70%
```



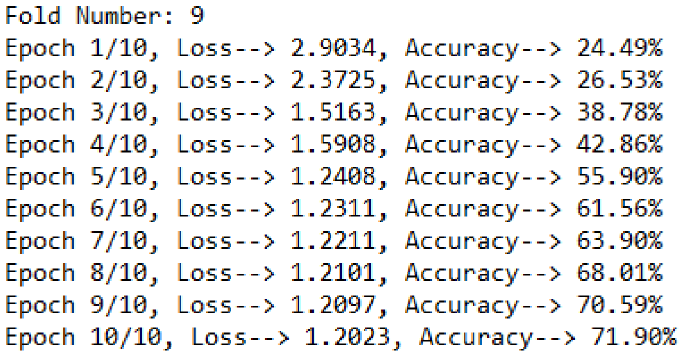
```
Fold Number: 3
Epoch 1/10, Loss--> 1.9121, Accuracy--> 25.00%
Epoch 2/10, Loss--> 2.0005, Accuracy--> 22.92%
Epoch 3/10, Loss--> 1.3957, Accuracy--> 52.08%
Epoch 4/10, Loss--> 1.3627, Accuracy--> 37.50%
Epoch 5/10, Loss--> 1.2939, Accuracy--> 45.83%
Epoch 5/10, Loss--> 1.2619, Accuracy--> 57.67%
Epoch 5/10, Loss--> 1.2439, Accuracy--> 57.88%
Epoch 5/10, Loss--> 1.2563, Accuracy--> 57.13%
Epoch 5/10, Loss--> 1.1939, Accuracy--> 60.18%
Epoch 5/10, Loss--> 1.1839, Accuracy--> 65.13%
```



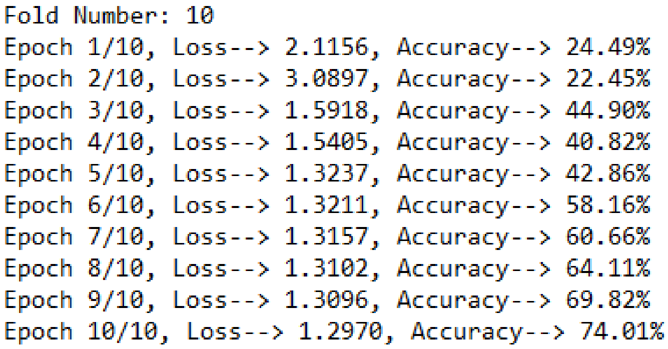
```
Fold Number: 5
Epoch 1/10, Loss--> 4.0645, Accuracy--> 20.83%
Epoch 2/10, Loss--> 1.6890, Accuracy--> 18.75%
Epoch 3/10, Loss--> 1.5443, Accuracy--> 33.33%
Epoch 4/10, Loss--> 1.4026, Accuracy--> 45.83%
Epoch 5/10, Loss--> 1.2679, Accuracy--> 50.00%
Epoch 6/10, Loss--> 1.2655, Accuracy--> 55.20%
Epoch 7/10, Loss--> 1.2630, Accuracy--> 56.01%
Epoch 8/10, Loss--> 1.2621, Accuracy--> 56.56%
Epoch 9/10, Loss--> 1.1156, Accuracy--> 57.66%
Epoch 10/10, Loss--> 1.1258, Accuracy-->62.91%
```



```
Fold Number: 7
Epoch 1/10, Loss--> 3.5336, Accuracy--> 14.29%
Epoch 2/10, Loss--> 2.0122, Accuracy--> 34.69%
Epoch 3/10, Loss--> 1.4615, Accuracy--> 42.86%
Epoch 4/10, Loss--> 1.5209, Accuracy--> 42.89%
Epoch 5/10, Loss--> 1.1472, Accuracy--> 53.06%
Epoch 6/10, Loss--> 1.1333, Accuracy--> 55.88%
Epoch 7/10, Loss--> 1.1312, Accuracy--> 55.54%
Epoch 8/10, Loss--> 1.1301, Accuracy--> 56.09%
Epoch 9/10, Loss--> 1.1272, Accuracy--> 61.76%
Epoch 10/10, Loss--> 1.1102, Accuracy--> 66.56%
```



```
Fold Number: 9
Epoch 1/10, Loss--> 2.9034, Accuracy--> 24.49%
Epoch 2/10, Loss--> 2.3725, Accuracy--> 26.53%
Epoch 3/10, Loss--> 1.5163, Accuracy--> 38.78%
Epoch 4/10, Loss--> 1.5908, Accuracy--> 42.86%
Epoch 5/10, Loss--> 1.2408, Accuracy--> 55.90%
Epoch 6/10, Loss--> 1.2311, Accuracy--> 61.56%
Epoch 7/10, Loss--> 1.2211, Accuracy--> 63.90%
Epoch 8/10, Loss--> 1.2101, Accuracy--> 68.01%
Epoch 9/10, Loss--> 1.2097, Accuracy--> 70.59%
Epoch 10/10, Loss--> 1.2023, Accuracy--> 71.90%
```

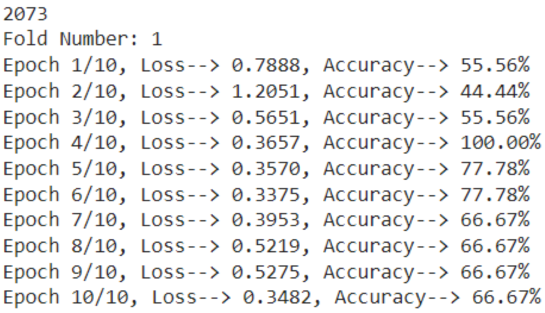


```
Fold Number: 10
Epoch 1/10, Loss--> 2.1156, Accuracy--> 24.49%
Epoch 2/10, Loss--> 3.0897, Accuracy--> 22.45%
Epoch 3/10, Loss--> 1.5918, Accuracy--> 44.90%
Epoch 4/10, Loss--> 1.5405, Accuracy--> 40.82%
Epoch 5/10, Loss--> 1.3237, Accuracy--> 42.86%
Epoch 6/10, Loss--> 1.3211, Accuracy--> 58.16%
Epoch 7/10, Loss--> 1.3157, Accuracy--> 60.66%
Epoch 8/10, Loss--> 1.3102, Accuracy--> 64.11%
Epoch 9/10, Loss--> 1.3096, Accuracy--> 69.82%
Epoch 10/10, Loss--> 1.2970, Accuracy--> 74.01%
```

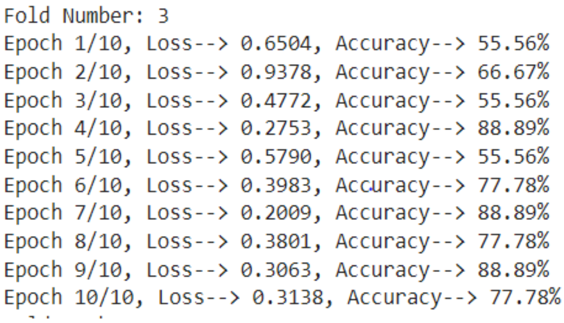
New Model:

- The dataset on which our new model has been trained, consists of filtered and balanced images in all the five classes, with each class containing an approximate of 400 images, and we have changed the resize parameter as compared to the build 1 for faster and better processing.
- In order to increase the accuracy of the model, one more convolution layers has been added each with a stride of (1,1) and a padding of (1,1)
- The model is evaluated using the stratified k-fold strategy with 10 folds and the shuffle parameter is set to true to ensure random shuffling. Each fold consists of 10 epochs and accuracy, recall, precision and f1-score are calculated for each fold as well.

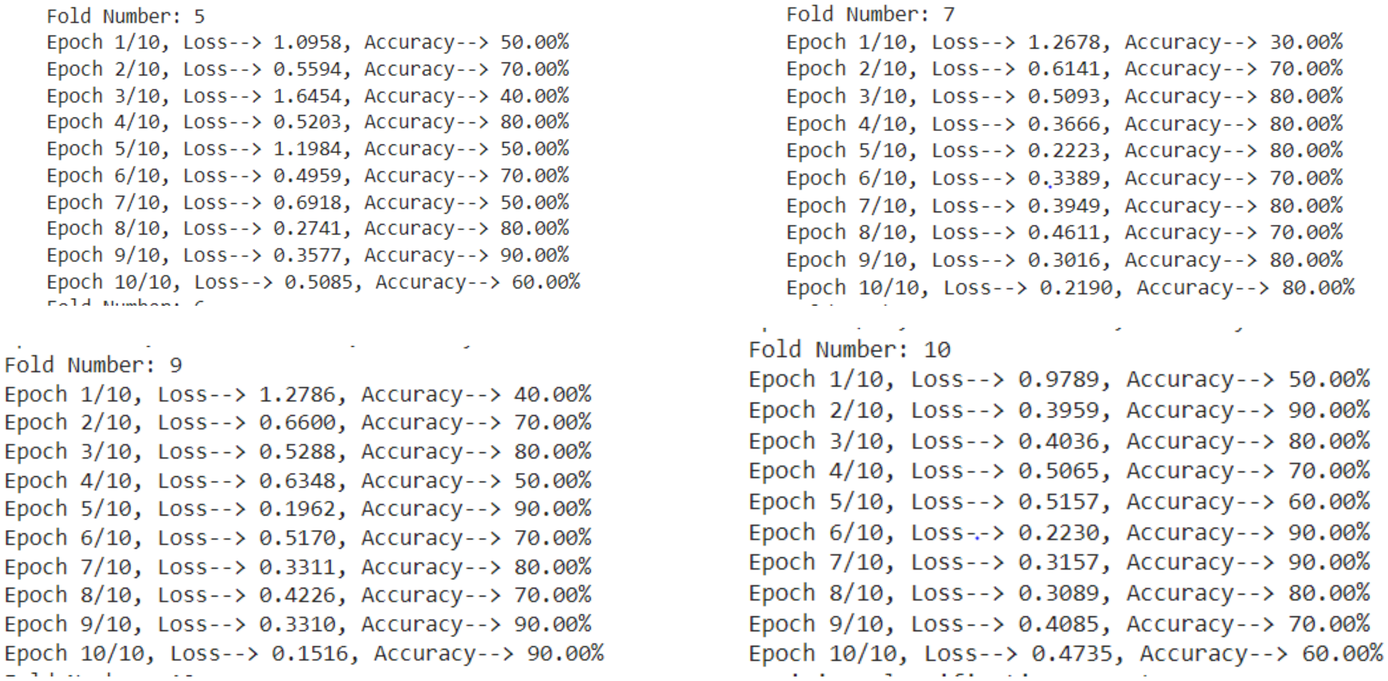
K Fold Results on New Model:



```
2073
Fold Number: 1
Epoch 1/10, Loss--> 0.7888, Accuracy--> 55.56%
Epoch 2/10, Loss--> 1.2051, Accuracy--> 44.44%
Epoch 3/10, Loss--> 0.5651, Accuracy--> 55.56%
Epoch 4/10, Loss--> 0.3657, Accuracy--> 100.00%
Epoch 5/10, Loss--> 0.3570, Accuracy--> 77.78%
Epoch 6/10, Loss--> 0.3375, Accuracy--> 77.78%
Epoch 7/10, Loss--> 0.3953, Accuracy--> 66.67%
Epoch 8/10, Loss--> 0.5219, Accuracy--> 66.67%
Epoch 9/10, Loss--> 0.5275, Accuracy--> 66.67%
Epoch 10/10, Loss--> 0.3482, Accuracy--> 66.67%
```



```
Fold Number: 3
Epoch 1/10, Loss--> 0.6504, Accuracy--> 55.56%
Epoch 2/10, Loss--> 0.9378, Accuracy--> 66.67%
Epoch 3/10, Loss--> 0.4772, Accuracy--> 55.56%
Epoch 4/10, Loss--> 0.2753, Accuracy--> 88.89%
Epoch 5/10, Loss--> 0.5790, Accuracy--> 55.56%
Epoch 6/10, Loss--> 0.3983, Accuracy--> 77.78%
Epoch 7/10, Loss--> 0.2009, Accuracy--> 88.89%
Epoch 8/10, Loss--> 0.3801, Accuracy--> 77.78%
Epoch 9/10, Loss--> 0.3063, Accuracy--> 88.89%
Epoch 10/10, Loss--> 0.3138, Accuracy--> 77.78%
```



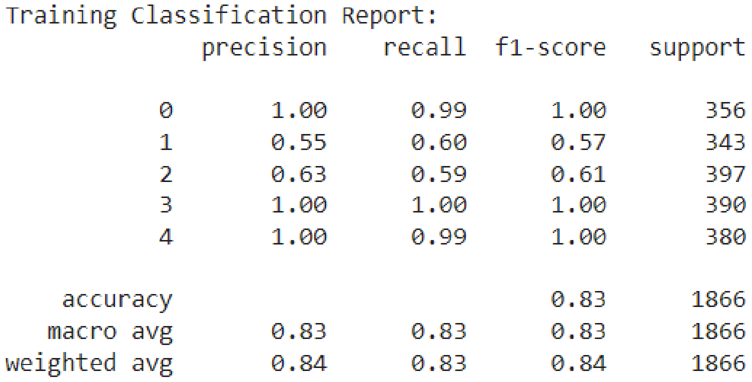
```
Fold Number: 5
Epoch 1/10, Loss--> 1.0958, Accuracy--> 50.00%
Epoch 2/10, Loss--> 0.5594, Accuracy--> 70.00%
Epoch 3/10, Loss--> 1.6454, Accuracy--> 40.00%
Epoch 4/10, Loss--> 0.5203, Accuracy--> 80.00%
Epoch 5/10, Loss--> 1.1984, Accuracy--> 50.00%
Epoch 6/10, Loss--> 0.4959, Accuracy--> 70.00%
Epoch 7/10, Loss--> 0.6918, Accuracy--> 50.00%
Epoch 8/10, Loss--> 0.2741, Accuracy--> 80.00%
Epoch 9/10, Loss--> 0.3577, Accuracy--> 90.00%
Epoch 10/10, Loss--> 0.5085, Accuracy--> 60.00%
```

```
Fold Number: 7
Epoch 1/10, Loss--> 1.2678, Accuracy--> 30.00%
Epoch 2/10, Loss--> 0.6141, Accuracy--> 70.00%
Epoch 3/10, Loss--> 0.5093, Accuracy--> 80.00%
Epoch 4/10, Loss--> 0.3666, Accuracy--> 80.00%
Epoch 5/10, Loss--> 0.2223, Accuracy--> 80.00%
Epoch 6/10, Loss--> 0.3389, Accuracy--> 70.00%
Epoch 7/10, Loss--> 0.3949, Accuracy--> 80.00%
Epoch 8/10, Loss--> 0.4611, Accuracy--> 70.00%
Epoch 9/10, Loss--> 0.3016, Accuracy--> 80.00%
Epoch 10/10, Loss--> 0.2190, Accuracy--> 80.00%
```

```
Fold Number: 9
Epoch 1/10, Loss--> 1.2786, Accuracy--> 40.00%
Epoch 2/10, Loss--> 0.6600, Accuracy--> 70.00%
Epoch 3/10, Loss--> 0.5288, Accuracy--> 80.00%
Epoch 4/10, Loss--> 0.6348, Accuracy--> 50.00%
Epoch 5/10, Loss--> 0.1962, Accuracy--> 90.00%
Epoch 6/10, Loss--> 0.5170, Accuracy--> 70.00%
Epoch 7/10, Loss--> 0.3311, Accuracy--> 80.00%
Epoch 8/10, Loss--> 0.4226, Accuracy--> 70.00%
Epoch 9/10, Loss--> 0.3310, Accuracy--> 90.00%
Epoch 10/10, Loss--> 0.1516, Accuracy--> 90.00%
```

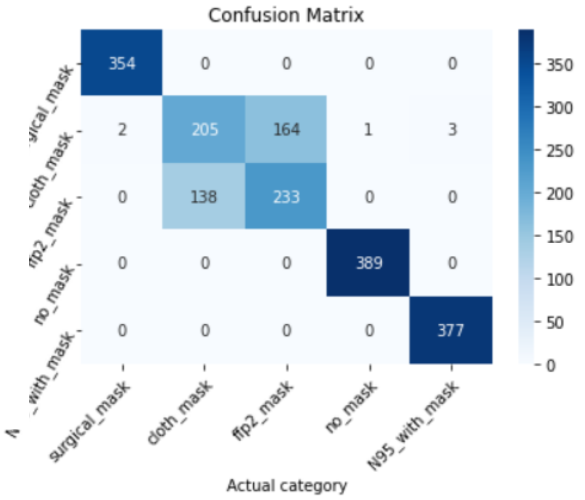
```
Fold Number: 10
Epoch 1/10, Loss--> 0.9789, Accuracy--> 50.00%
Epoch 2/10, Loss--> 0.3959, Accuracy--> 90.00%
Epoch 3/10, Loss--> 0.4036, Accuracy--> 80.00%
Epoch 4/10, Loss--> 0.5065, Accuracy--> 70.00%
Epoch 5/10, Loss--> 0.5157, Accuracy--> 60.00%
Epoch 6/10, Loss--> 0.2230, Accuracy--> 90.00%
Epoch 7/10, Loss--> 0.3157, Accuracy--> 90.00%
Epoch 8/10, Loss--> 0.3089, Accuracy--> 80.00%
Epoch 9/10, Loss--> 0.4085, Accuracy--> 70.00%
Epoch 10/10, Loss--> 0.4735, Accuracy--> 60.00%
```

Accuracy for each epoch for 10 folds

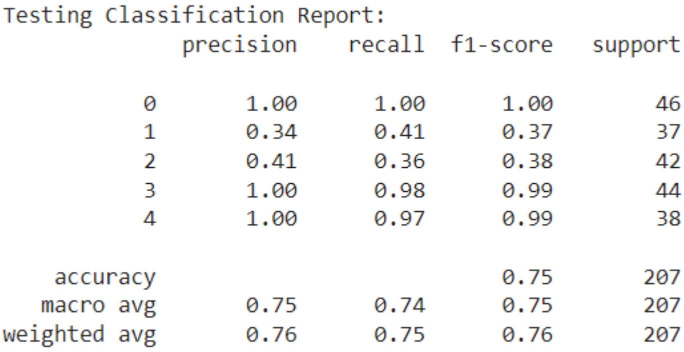


```
Training Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.99      1.00       356
           1       0.55      0.60      0.57       343
           2       0.63      0.59      0.61       397
           3       1.00      1.00      1.00       390
           4       1.00      0.99      1.00       380

    accuracy                           0.83      1866
   macro avg       0.83      0.83      0.83      1866
weighted avg       0.84      0.83      0.84      1866
```

Training Confusion Matrix:

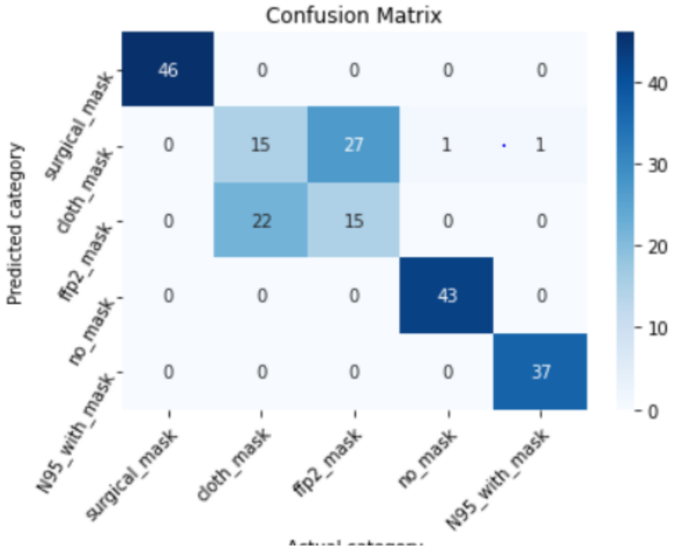


Testing Results After K-Fold on new Model:



```
Testing Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        46
           1       0.34      0.41      0.37        37
           2       0.41      0.36      0.38        42
           3       1.00      0.98      0.99        44
           4       1.00      0.97      0.99        38

    accuracy                           0.75       207
   macro avg       0.75      0.74      0.75       207
weighted avg       0.76      0.75      0.76       207
```

Testing Confusion Matrix:



Overall Performance Comparison

The performance of the new model has been compared to that of the old model and it showed 14% improvement in accuracy, 12% improvement in precision, 14% improvement in recall, and 14% improvement in F1-Score.

Metric	**Old Score**	**New Score**
Accuracy	**61%**	**75%**
Precision	**0.64**	**0.76**
Recall	**0.61**	**0.75**
F-1 Score	**0.62**	**0.76**

5. References

Reference Links to the datasets used:

[1] Face Masks Dataset-
https://www.kaggle.com/dataset/a71dfe0333dcabd1827ca3d6dcfd62d43785f83d3c38321b4113339a14f780e9

[2] Other References for Mask Dataset

- https://www.kaggle.com/ashishjangra27/face-mask-12k-images-dataset
- https://www.kaggle.com/andrewmvd/face-mask-detection
- https://www.kaggle.com/prasoonkottarathil/face-mask-lite-dataset
- https://www.kaggle.com/dhruvmak/face-mask-detection
- https://www.kaggle.com/sumansid/facemask-dataset
- https://www.kaggle.com/andrewmvd/face-mask-detection?select=images
- https://www.kaggle.com/omkargurav/face-mask-dataset
- https://www.kaggle.com/prithwirajmitra/covid-face-mask-detection-dataset

[3] Reference for N95 With Valve Images- Kaggle
https://www.kaggle.com/datasets/andrewmvd/face-mask-detection

[4] Reference for N95 With Valve Images - Kaggle
https://www.kaggle.com/datasets/wobotintelligence/face-mask-detection-dataset

[5] Manually Searching Google Images

- https://abc7news.com/n-95-mask-with-valve-kn95-n95-for-smoke/6378125/
- https://hartfordhealthcare.org/about-us/news-press/news-detail?articleid=26322&publicId=395
- https://theconversation.com/high-filtration-masks-only-work-when-they-fit-so-we-created-a-new-way-to-test-if-they-do-155987
- https://www.jracenstein.com/learn/expert-advice/why-you-shouldn%E2%80%99t-wear-a-mask-with-a-valve-for-covid-19/a136

- https://evaculife.com.au/product/adult-kn95-n95-p2-coronavirus-mask-white/
- https://www.sfchronicle.com/bayarea/article/Coronavirus-Bay-Area-officials-warn-some-N95-15208241.php
- https://sf.curbed.com/2018/11/9/18079866/fire-smoke-face-mask-find-oakland-san-francisco
- https://www.popsci.com/story/health/face-mask-valve-covid/
- And many more…

References Links for model design and evaluation

1. https://en.wikipedia.org/wiki/Convolutional_neural_network
2. https://machinelearningmastery.com/k-fold-cross-validation/
3. https://machinelearningmastery.com/confusion-matrix-machine-learning
4. https://www.analyticsvidhya.com/blog/2021/03/image-augmentation-techniques-for-training-deep-learning-models/
5. https://towardsdatascience.com/batch-normalisation-in-deep-neural-network-ce65dd9e8db