**"# Facemask-type-AI Project"**

**Team members**

1. Kary Sutariya (40193909)

2. Maryam Valipour (40224474)

3. Kanika Aggarwal (40195981)

**List of Folders:**

1. **Codes:**

* **Order of running Codes:**

For models evaluated with cross-validation methods (Models from Folders: “Models/resnet18/” and “Models/ old\_model\_cross\_validation/”):

1. Run file find\_mean\_std.py
2. Run file load\_data\_cross\_val.py
3. Run file training\_cross\_val.py
4. Run file evaluation\_cross\_val.py
5. Run file application.py

For models evaluated with training/test split (Models from Folders: “Models/ old\_model\_train\_test\_split/ver1”, “Models/ old\_model\_train\_test\_split/ver2”, and “Models/ old\_model\_train\_test\_split/ver3”):

1. training\_train\_test\_split.py
2. evaluation\_train\_test\_split.py
3. Run file application.py

* **Description of Codes:**

To run these codes, all required libraries which has been used in all the listed python files should be installed

1. find\_mean\_std.py:
2. By running this file, the mean and standard deviation of the images in the new balanced dataset will be saved into a folder named “saved tensors” in the “Codes” folder. The saved mean and std tensors are later used in the load\_data\_cross\_val.py file.
3. load\_data\_cross\_val.py:
4. By running this file, the images from the new dataset are split into train and test categories and the train and test images tensors and their corresponding labels in tensor format are saved into “saved tensors” folder in the “Codes” folder. These tensors are later used in the training\_cross\_val.py file
5. training\_cross\_val.py:

To run this file, at first:

5-1) for new model (resnet18) with k-fold cross validation, set: (model\_name = 'resnet18')

Saves trained models for each fold (total 10 models) in the Folder “Models/resnet18/”

5-2) for old model with cross validation: set: (model\_name ='old\_model\_cross\_validation')

By running, a PyTorch model is built by performing a 10-fold cross-validation evaluation (with random shuffling) on your AI to improve model performance in all the categories.

Trained models are saved for each fold (total 10 models) in the Folder “Models/ old\_model\_cross\_validation/”

1. evaluation\_cross\_val.py:

To run this file, at first:

6-1) for new model (resnet18) with k-fold cross validation, set: (model\_name = 'resnet18')

6-2) for old model with cross validation: set: (model\_name ='old\_model\_cross\_validation')

By running, it uses the saved models from “the training\_cross\_val.py” file to evaluate the model performance in the whole dataset, across the 5 categories, and the subclasses (male, female, young, and old).

Saved files: (Confusion matrix and Classification reports)

For the two models mentioned in 6-1 and 6-2, the confusion matrix and their classification reports are saved in different files.

6-1) new model (resnet18) with k-fold cross validation:

* + Confusion matrix across all the 10 folds, whole dataset, and all subclasses (female, male, young, and old) are saved in Folder “Figures/resnet18/”
  + Classification report:

for all folds: “Models/resnet18/classification\_report\_resnet18\_folds.txt”

for the whole dataset and the subclasses:

“Models/resnet18/classification\_report\_resnet18\_subclass.txt”

6-2) old model with cross validation:

* + Confusion matrix across all the 10 folds, whole dataset, and all subclasses (female, male, young, and old) are saved in Folder “Figures/old\_model\_cross\_validation/”
  + Classification report:

for all folds: “Models/resnet18/ classification\_report\_old\_model\_folds.txt”

for the whole dataset and the subclasses:

“Models/resnet18/ classification\_report\_old\_model\_subclass.txt”

1. training\_train\_test\_split.py:

This code is for phase 1 of the project:

By running, data is split into train, validation and test using training/test split from scikit-learn library. Three Pytorch models with different architectures are built. The model form class “net\_ver1” has the best performance.

Trained models for the three different architectures are saved in the three subfolders of this folder: “Models/old\_model\_train\_test\_split/”

1. evaluation\_train\_test\_split.py:

To run this file, at first:

8-1) set: “test\_dataset = 'new’ (default value )if you want to test on the test dataset from phase 2 of the project

8-2) set: “test\_dataset = 'old' if you want to test on the test dataset from phase 1 of the project

8-2) set: “load\_mode\_1=True”, “load\_mode\_2=True”, “load\_mode\_3=True” if you want to see the result of evaluation from Model ver1, ver2, and ver3, respectively (default: the first model is True and the second and third model are set to False)

Model with best performance is saved in folder: “Models/old\_model\_train\_test\_split/ver1”

For example for model saved in: “Models/old\_model\_train\_test\_split/ver1/”

* + Confusion matrix across whole dataset, and all subclasses (female, male, young, and old) are saved in “Figures/old\_model\_train\_test\_split/ver1”
  + Classification report:

for the whole dataset and the subclasses is printed in the console

1. application.py:

This file is for real time application. it reads images from “Data/SampleData/” Folder, which have not been used in either training or testing phase and predicts a label for each of the images.

1. data\_annotation.py: this file separated the new balanced dataset and saves a percentage of the data in the folder “Data/Data Part 2/Test”, which is used for evaluating the models
2. utils.py:

contains functions that are used across all the python codes.

annotation.csv: annotation file for the test dataset of the new balanced dataset the columns include the name of the image, gender, age, and label of each image

**Data:** contains images of five categories "Cloth mask", "No face mask", "Surgical mask", "N95 mask", "Mask worn incorrectly".

* **Subfolders:**

1. Data Part 1: This is the data used for phase 1 of the project.
2. Data Part 2: This is the data used for phase 2 of the project.

2-1) Whole: The whole dataset

2-2) Test: The separated dataset which is used in the test Phase (the data is separated by running the code .py)

1. SampleData: This is the data used for real-time application of the project used in file “application.py”
2. **Figures:**

* **Subfolders:** contains the confusion matrix for the whole dataset, each of the subclasses (female, male, young, and old), and the 10 folds when applicable.

1. resnet18
2. old\_model\_cross\_validation
3. old\_model\_train\_test\_split
4. **Models:**

* **Subfolders:** contains all the saved models (for each of the folds when applicable)

1. resnet18: “Models”
2. old\_model\_cross\_validation
3. old\_model\_train\_test\_split

To be able to test the code, please download the model and Data folders from the given google drive link and put it in the submission folder

Folder of Models:

<https://drive.google.com/drive/folders/11VV7r7eq0enIMyChXLBtzRg-WMw6YWUQ?usp=sharing>

Folder of Data:

<https://drive.google.com/drive/folders/1vzLxbhMTAiOjN81jkL6dpTOUXQNQPcyu?usp=sharing>