Face Mask Detection

INFT 22001

Professor Laith Adi

Ozair Rahman, Rhichard Koh, Caleb Mcanuff

Table of Contents

Introduction

The primary object is to successfully be able identify whether an individual is wearing a face mask, not wearing a face mask, and if the individual is wearing the face mask incorrectly.

Dataset

The dataset consists of two classes. The first class is images with individuals with masks on and the second class is individuals who are not wearing a face mask. Each class contains images and xml files that represent the annotation which correspond to each image. The xml annotations contain the name of the label, if it's with_mask, without_mask or mask_incorrectly, it also contains the dimensions of the boundary box itself as well as the location of the boundary box on the image itself.

Caveats to bear in with the dataset is that a majority of individuals in the dataset appear to be of asian ethnicity. This may skew results of the model when performing on other ethnicities.

Data Preprocessing

Before the data was given to the model, the data was split, augmented and normalized prior to feeding it to the model. The first thing that was done was to split the dataset into training, testing and validation partitions. The dataset is split into 70% for training data and 15% for validation and testing data respectfully.

Then, the all data partitions feed all the images into a TensorFlow imageDataGenerator pipeline so it simplifies how our data is processed and given to the model. Next, the image data is augmented so that it can make the model more robust by allowing it to account for multiple edge case scenarios. The augmentation that was done on all images was a horizontal flip, a random zoom of 20% as well as a random shear range of 20%. The data augmentation is then applied to the training, testing, and validation data. The images are also cropped to 128 x 128 pixels so that the pre-trained model can use the data to train on the dataset.

Model Building

Two pre-trained models are used to train for face mask detection, VGG19 and InceptionV3. Transfer learning is used so that the model can be trained on the dataset. Modifications to both models are made so that it is compatible with the dataset. Namely, a flatten layer is added so that it can convert the multi dimensional input a.k.a the image to a single vector so that it can be fed to a fully connected layer. A fully connected layer with a sigmoid activation function with 2 units is added so that it can then categorize the model into one of two categories, with_mask or without_mask.

The total number of parameters of the VGG19 model consisted of 20,040,770 parameters, 16,386 are trainable parameters, mainly, the fully connected layer. There are also 20,024,384 non-trainable parameters that are mainly due to the transfer learning that was applied from the VGG19 pre-trained model. Both models are trained for 20 epochs, each epoch lasting 9 steps.

The Inception V3 model contains a total number of 21,819,170. 16,386 of those parameters are trainable due to the fully connected layer. 21,802,784 non-trainable parameters are added due to transfer learning.

Performance Metrics

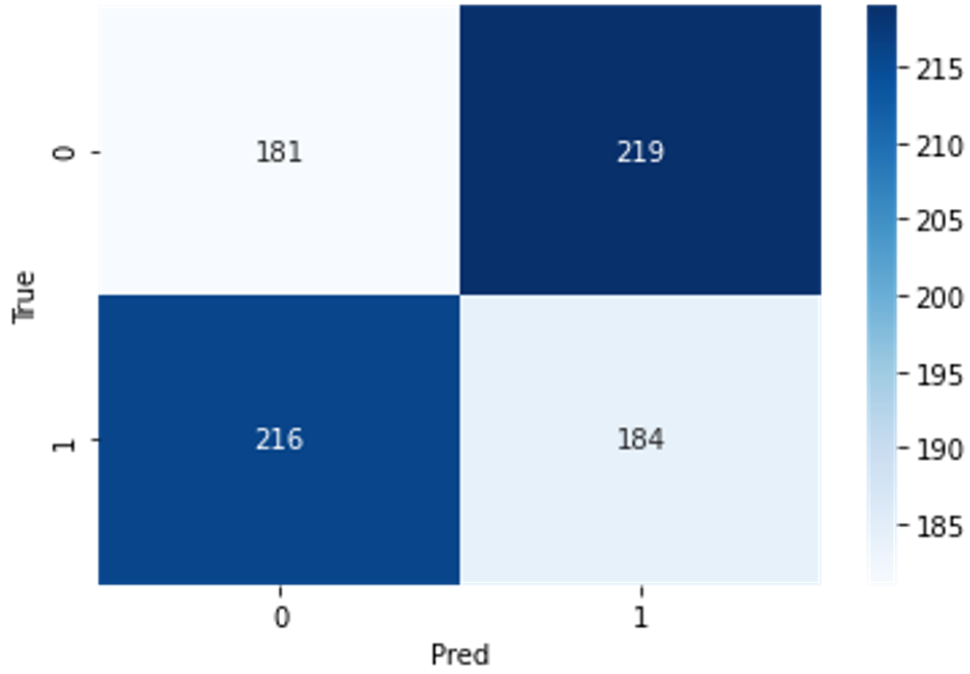
VGG19 Model Performance Metrics

The VGG19 model has a loss of 0.07 and accuracy 97.75% on the validation dataset. As shown below, the VGG19 model has an precision score of 0.46 for both mask and No mask, an recall score of 0.45 for Mask and 0.46 for No Mask, and an f1-score of 0.45 for the Mask category and 0.46 for the No Mask category.

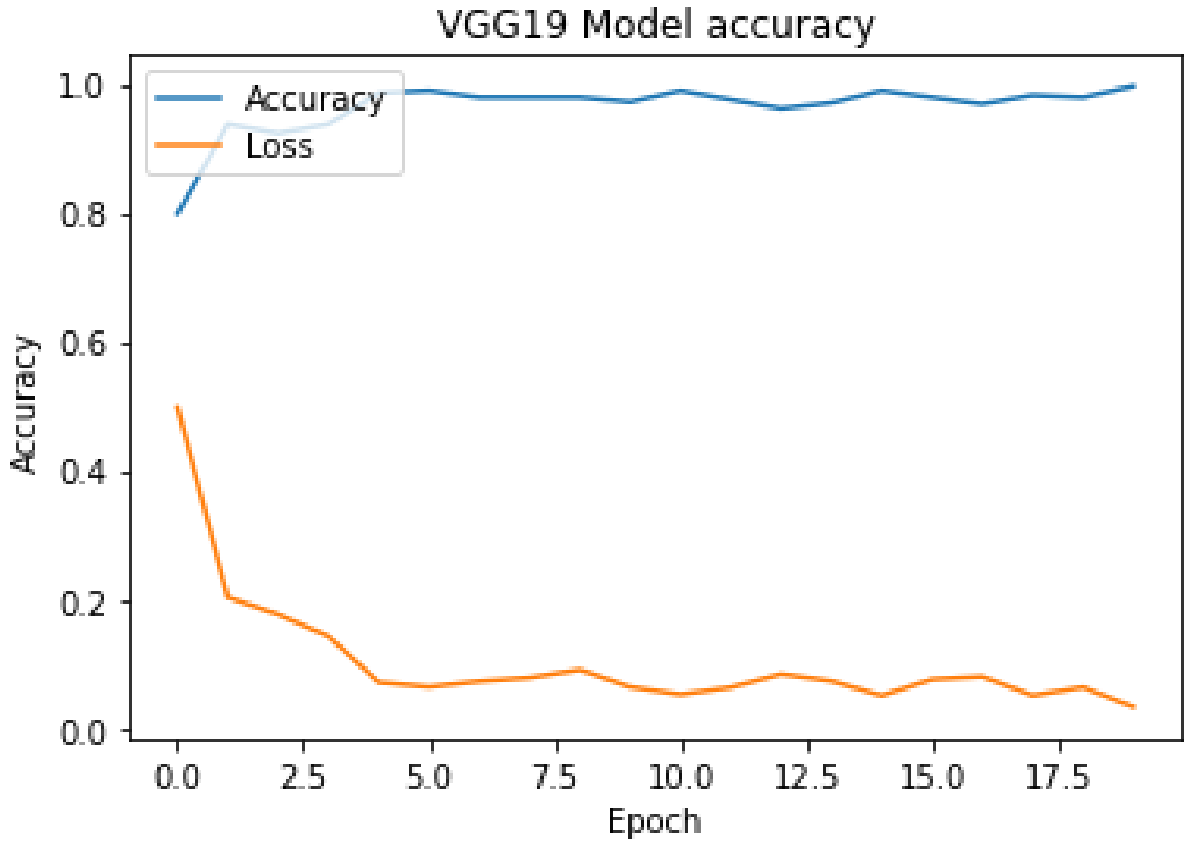
Classification Report

	precision	recall	f1-score	support
Mask	0.46	0.45	0.45	400
No Mask	0.46	0.46	0.46	400
accuracy			0.46	800
macro avg	0.46	0.46	0.46	800
weighted avg	0.46	0.46	0.46	800

Below is the confusion matrix of the VGG19 model. The VGG19 model got a total of 181 true positive images, 184 true negative images, 216 false negative images and 219 false positive.



Below is a line graph detailing the accuracy and loss of the VGG19 model. As explained above the VGG19 model has an accuracy of 97.75% and a loss of 0.07.



InceptionV3 Model

Below are the performance metrics for the InceptionV3 model. Across all categories, the performance metrics for the InceptionV3 model performs slightly better than the VGG19. After comparing both performance metrics and drawing that conclusion, it is decided to perform hypertuning to evaluate if the InceptionV3 model can perform even better than not performing hypertuning.

InceptionV3 model performance metrics without hyperparameter tuning

Below are the precision, recall and f1-score of the InceptionV3 model before performing any hypertuning parameters. The model receives a precision score of 0.49 for both Mask and No Mask categories. This score is slightly higher than the precision score of the VGG19 model which received a score of 0.46.

The InceptionV3 received a recall score of 0.48 for the Mask category and 0.49 for the No Mask category. Again, this is slightly higher than the recall score of the VGG19 model which received a score of 0.45 for the Mask category and 0.46 for the No mask category.

Lastly, InceptionV3 received an f1 score of 0.49 for both Mask and No Mask categories. Again, InceptionV3 received a slightly higher f1 score than the VGG19 model which received an f1 score of 0.45 for the Mask category and 0.46 for the No Mask category.

Classification Report

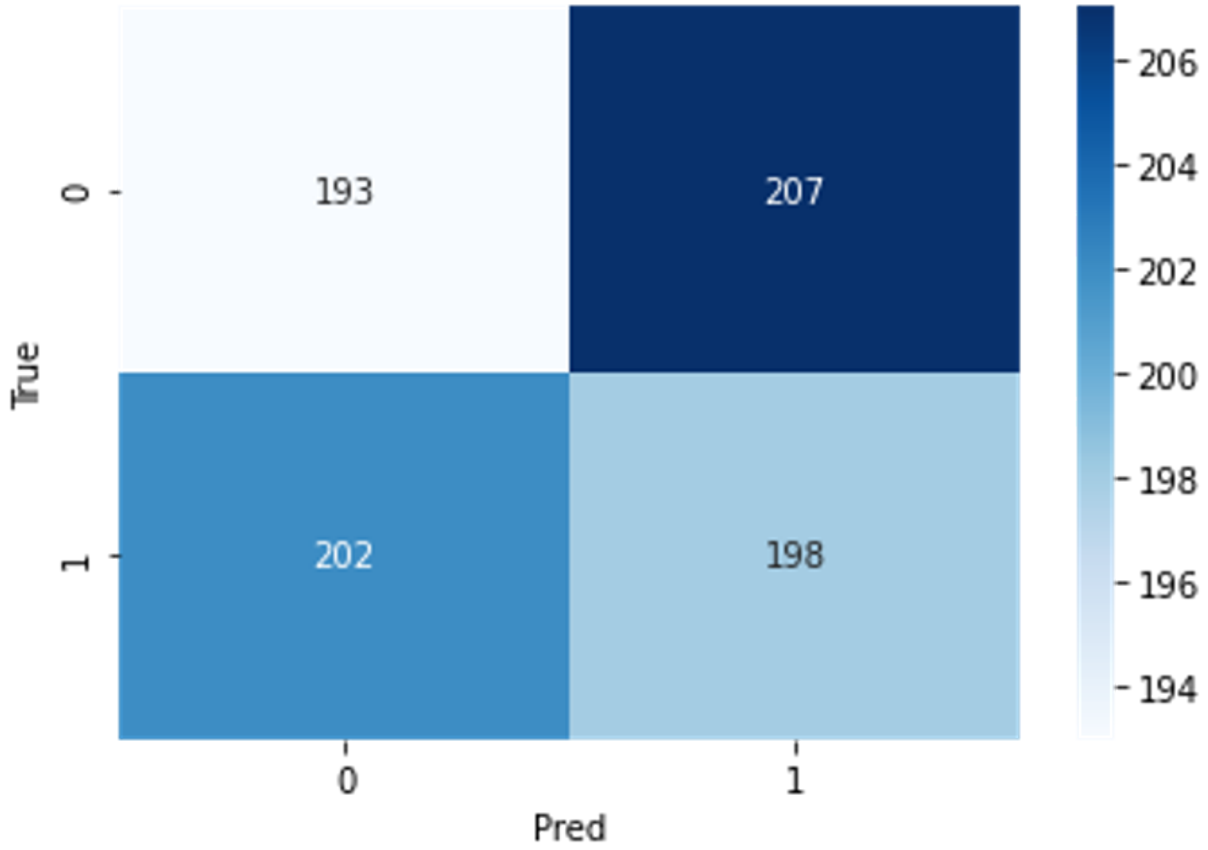
	precision	recall	f1-score	support
Mask	0.49	0.48	0.49	400
No Mask	0.49	0.49	0.49	400
accuracy			0.49	800
macro avg	0.49	0.49	0.49	800
weighted avg	0.49	0.49	0.49	800

Below is the confusion matrix for the InceptionV3 model before performing any hypertuning. The InceptionV3 model before performing hyperpertuning received 193 true positive images. This is higher than the true positive image score of the VGG19 model which received 181 true positive images.

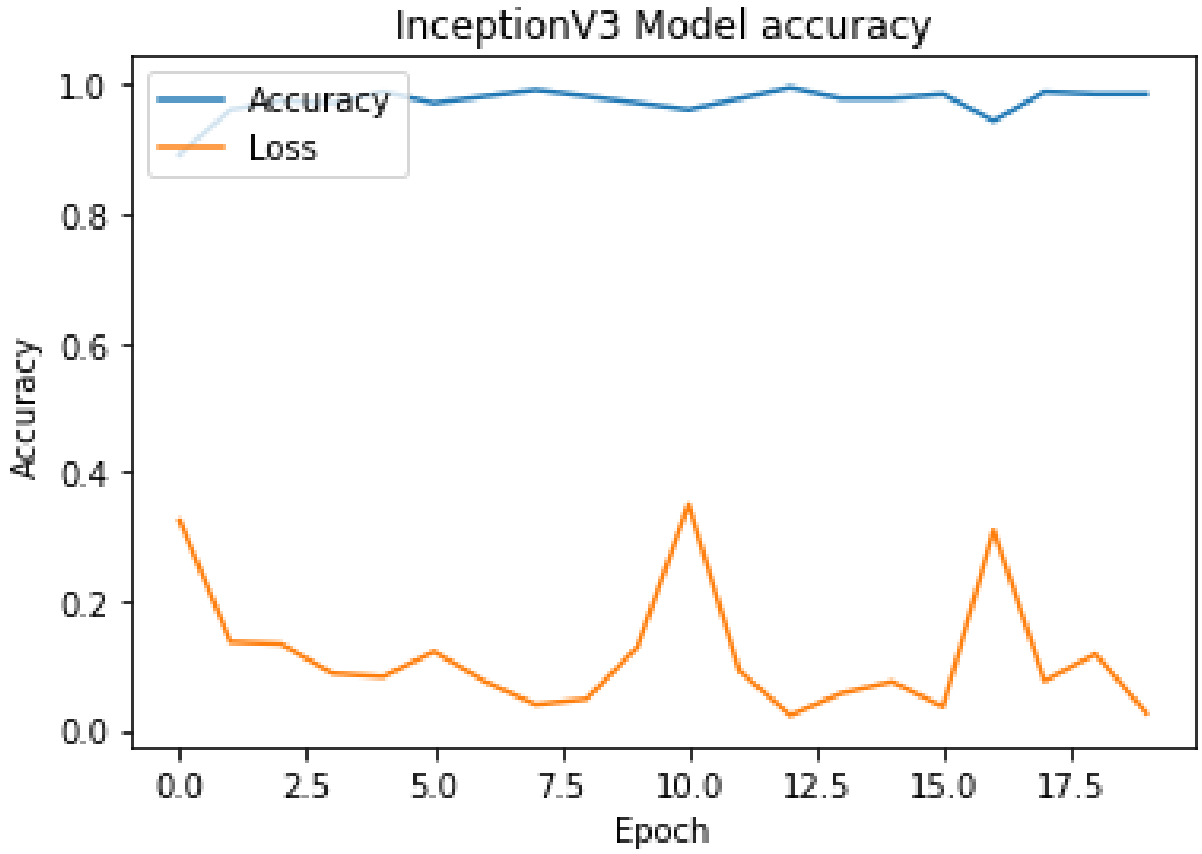
InceptionV3 received 198 true negatives images. Again, this is higher than VGG19 which received 184 true negative images.

InceptionV3 received 202 false negative images. Again, VGG19 received a higher score of 216 false negatives.

Lastly InceptionV3 received 207 false positives. Again, VGG19 received a higher score than InceptionV3 which received a score of 219 images.



Below is a line graph that displays the accuracy and loss of the InceptionV3 model before performing any hypertuning. InceptionV3 before performing hypertuning received an accuracy score of 99.37% and a loss score of 0.04. Compared to the VGG19 model, InceptionV3 performed better than the VGG19 model which received an accuracy score of 97.75% and a loss score of 0.07.



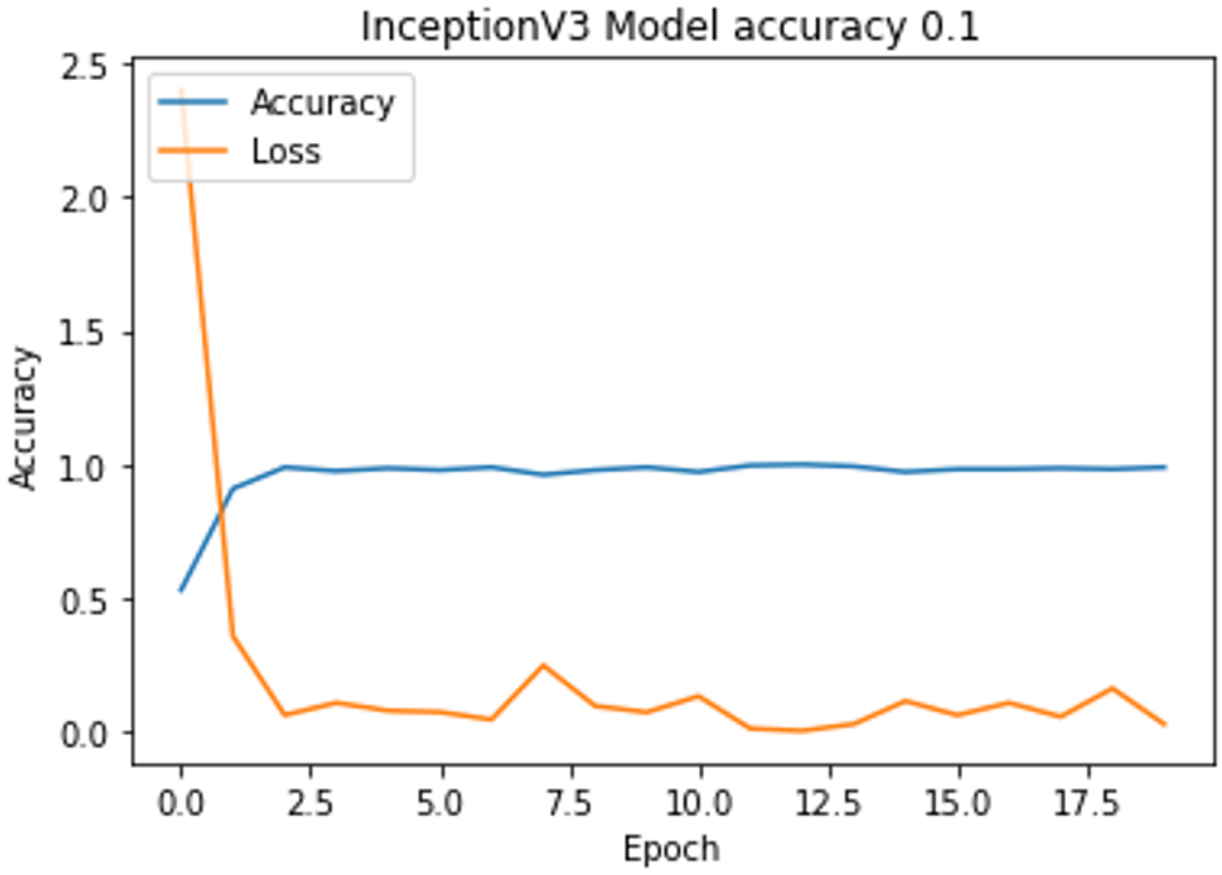
InceptionV3 model performance metrics with hyperparameter tuning

Below are line graphs that display the accuracy and loss of the InceptionV3 model after performing hypertuning. The parameters that were tuned was the learning rate. The InceptionV3 learning rate was changed three times once to 0.1, then 0.01, then 0.001.

InceptionV3 model learning rate set to 0.1

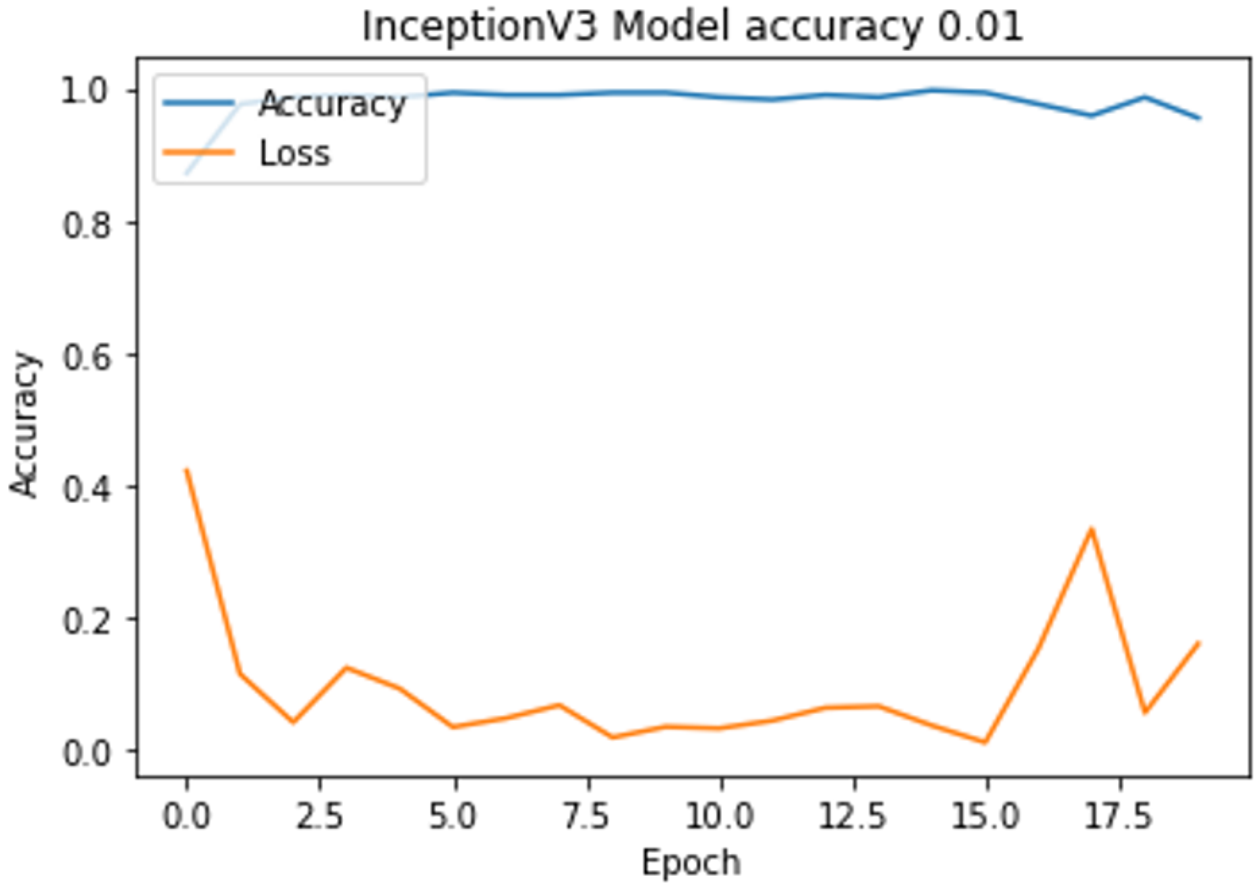
Below are all performance metrics of the InceptionV3 model where the learning rate was set to 0.1

Below is a line graph that displays the accuracy and loss of the InceptionV3 model where the learning rate is set to 0.1.



InceptionV3 model learning rate set to 0.01

Below is a line graph of the InceptionV3 model where the learning rate is set to 0.01 displaying the accuracy and loss.



InceptionV3 model learning rate set to 0.001

Below is a line graph of the InceptionV3 where the learning rate is set to 0.001 displaying the accuracy and loss.

Comparing all InceptionV3 models where hyperparameter tuning is performed, this model with the learning rate set to 0.001 before the best.