Gowri Dinesh Nair

Registration Number: 200968001

B.Tech Data Science and Engineering

Section B, Semester V

Deep Learning Project:
Cataract Identification
Final Phase

As seen from the literature review, the best performing models for cataract classification were transfer learning models. Consequently, for this project, the models chosen were VGGNet, ResNet50, and InceptionV3. Keeping a CNN with five layers (a convolutional layer, a max pooling layer, a flattening layer, and two fully connected layers) as the baseline model, all models were trained for 10 epochs. The color fundus photographs used as training data were pre processed using Tensorflow's ImageDataGenerator by rescaling, flipping, rotating, and applying more transformations to the images. The models' performances were evaluated based on their respective accuracies and losses.

The CNN model presented an accuracy of 65% and a loss of 62.9%, but took the least training time out of all four models. This is explained by its relatively simpler architecture as compared to the transfer learning models. Despite being one of the best classifiers mentioned in the literature review, the ResNet model was the worst performing model, with the lowest accuracy score (53.3%) and highest loss value (68.6%). VGGNet proved to have better performance, with a higher accuracy of 83.2% and lower loss of 40.1%. However, it had the longest training time out of all the models (1.5 times the training time of the VGGNet and

Gowri Dinesh Nair

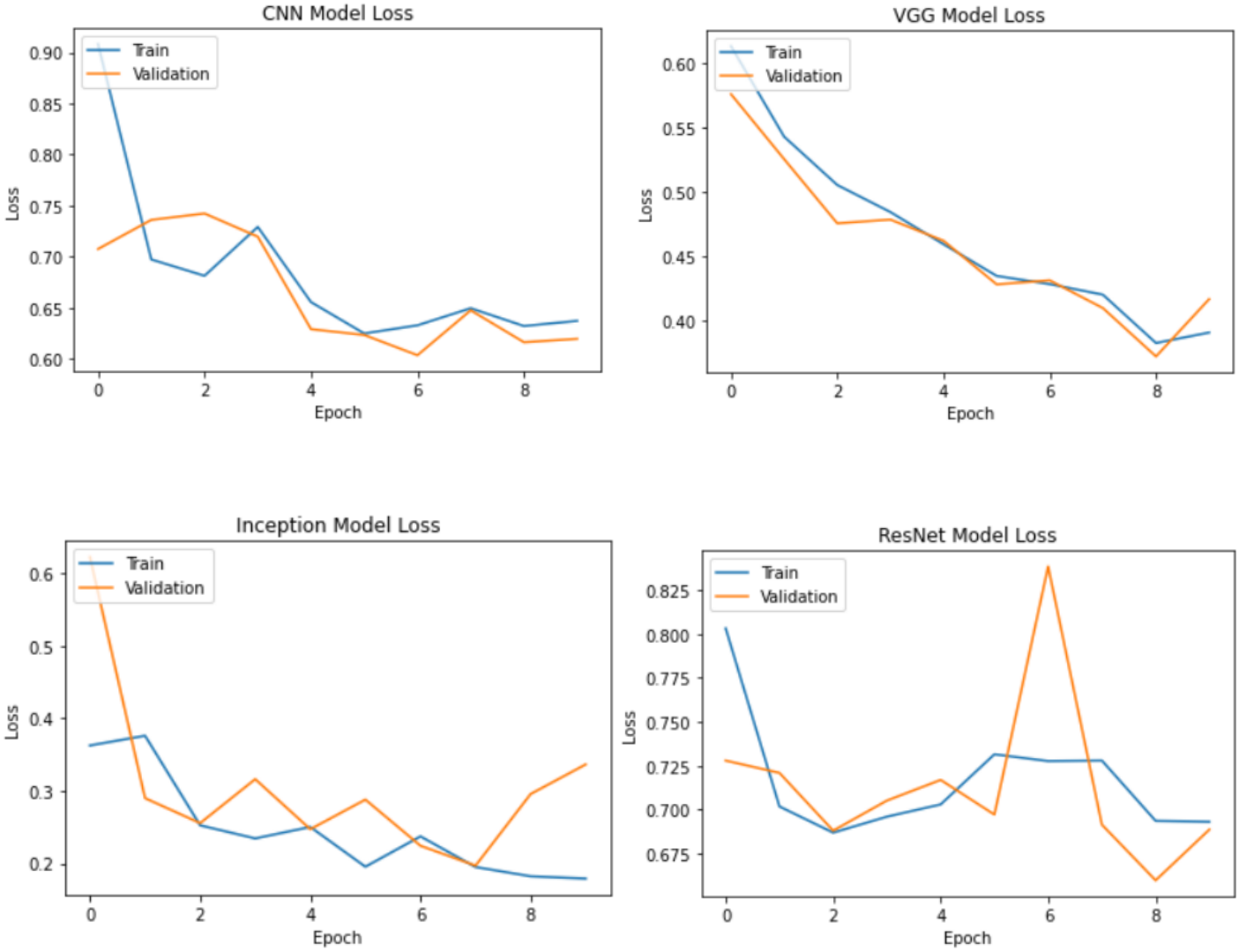
Registration Number: 200968001

B.Tech Data Science and Engineering

Section B, Semester V

InceptionV3 models). This is to be expected as VGGNet is established as a hefty model with a high number of parameters, 14,764,866 in this case. The best performing model was the InceptionV3 model, which had an accuracy of 89.6% and a loss value of 26.9%.

The models' loss plots are as shown below:



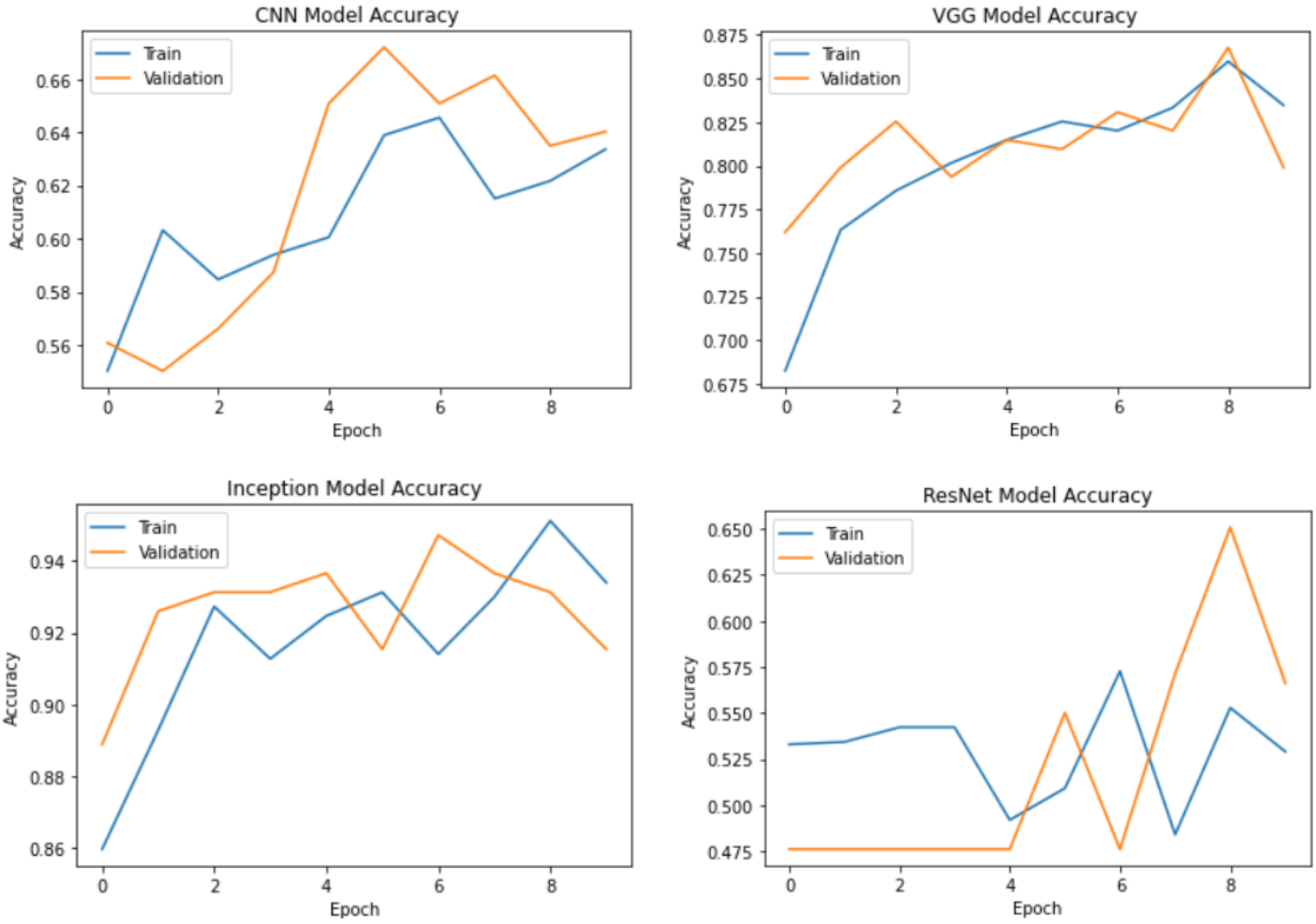
Gowri Dinesh Nair

Registration Number: 200968001

B.Tech Data Science and Engineering

Section B, Semester V

The accuracy plots are as shown below:



As seen from the above graphs, the CNN model and InceptionV3 models were the only ones to have validation loss values lower than those of their training loss values. The Resnet and VGGNet models' loss plots show otherwise. But since the ResNet model's loss constantly decreases even at the tenth training epoch, we can reason that this behavior is the result of our

model underfitting. The ResNet model also seems to perform poorly with the batches of the adam optimizer as seen by the unpredictable and inconsistent spikes in its accuracy graph. The performance is likely to improve if the number of training epochs were increased (this could not be confirmed as the Colab notebook's runtime kept getting disconnected and crashed when the number of epochs was increased beyond 10). However, the VGG models' behavior is likely due to overfitting, as indicated by its accuracy values decreasing and its loss values increasing beyond eight epochs. Thus, for optimal performance, this model's training should be terminated at 8 epochs. Finally, the InceptionV3 models' accuracies and losses were consistently increasing and decreasing, respectively. Thus, it is most compatible with the color fundus photographs and is the best-performing classifier out of all four models.

Considering that these models are to be used in healthcare institutions where the models' ability to correctly classify cataract is crucial, the accuracy of the deep learning model chosen should perform well accordingly. Additionally, training time must be taken into account so as to provide the required health care services for the patients in a timely manner. Since the InceptionV3 model has the highest accuracy rate and lowest loss value, it was the best classifier of cataract out of all four models. In terms of training time, the Inception model was the second-fastest to train after the CNN model, which makes it an appropriate choice for cataract classification out of all the models experimented with.