Introduction

The goal of this project is to build deep learning models to categorize cancerous vs non-cancerous skin images. Specifically to first build a deep convolutional neural network for binary classification of cancerous skin. Then use transfer learning to use a pre-trained kaggle model for the same goal.

Analysis

The data included 84 training images (42 cancerous, 42 noncancerous), and 184 testing images (42 cancerous, 42 noncancerous). Image directories are to be reorganized to be into their respective training and testing file directories. Images were loaded into a keras image dataset with a batch size 16 and image size of 170 by 170 pixels.

Methods

The first model I manually built with keras and tensorflow included 5 convolutional layers, and 3 maxpooling layers. I used relu as my activation function in my convolution layers with kernel_regularizer set as l2 regularization with a lambda of 0.001. After the convolution and max pooling layers, I flatten the nodes then fully connect them with a 64 dense layer with relu activation. I dropped 50% of the least contributing nodes. A 1 dimensional layer with sigmoid activation was added to get the final probabilities. I compile the model with an Adam optimizer with exponential decay learning rate, and use binary cross entropy loss function. I took the metrics of accuracy, precision, and recall. The model was fit on the training data then tested on the test data set with 12 epochs. Additionally, an early stopping function with a patience value of 5 epochs was added to the model fitting.

For the transfer learning model, I used the EfficientNetB0 keras base model. I then just repeated te same as above with the final flattening, fully connected dense layer with relu activation, dropout of 0.3, and 1 dimensional layer with sigmoid activation. I used adam optimizer with learning rate of 0.001 and binary cross entropy loss function with accuracy measurements. The model was fit with 5 epochs.

Analysis

The model shows overall improvement in training accuracy, reaching 89% by the 12th epoch, but experiences significant fluctuations in validation performance, especially in epochs 2, 4, and 5, where validation accuracy drops to around 25%. These dips suggest potential issues with overfitting or difficulties in generalizing to the validation set. Despite this, validation accuracy stabilizes at around 81% by the end, with recall at 83%, indicating the model is reasonably good at identifying cancerous images. Training precision improves steadily, aligning with recall, and suggests the model is becoming better at distinguishing cancerous from non-cancerous images. However, the early-stage overfitting and occasional validation performance drops highlight areas for further improvement, such as enhancing regularization or using data augmentation strategies.

The second model shows strong training performance, with accuracy reaching 90% by the end of the fifth epoch. However, the validation accuracy fluctuates significantly, dropping as low as 40% in the second epoch before stabilizing around 80%. This suggests the model is struggling with generalization and may be overfitting to the training data, as indicated by the large gap between training and validation losses.

Comparing both models, the first model demonstrates more stable performance across epochs, with a steady increase in validation accuracy. In contrast, the second model, while initially showing high training accuracy, experiences more significant variability in validation performance. This suggests the first model may have better generalization, while the second model could benefit from better regularization or more robust validation techniques.

Conclusion

One of the main things I learned from this assignment was how difficult it is to fine tune models with so many hyperparameters, possibilities for layers, normalizations, regularizations, etc. In the future I hope to spend more time tuning the model to gain more accurate results. Additionally to have more default values that I know work well with specific data and types of data.