

Colorectal Tumor Classification with Neural Networks

...

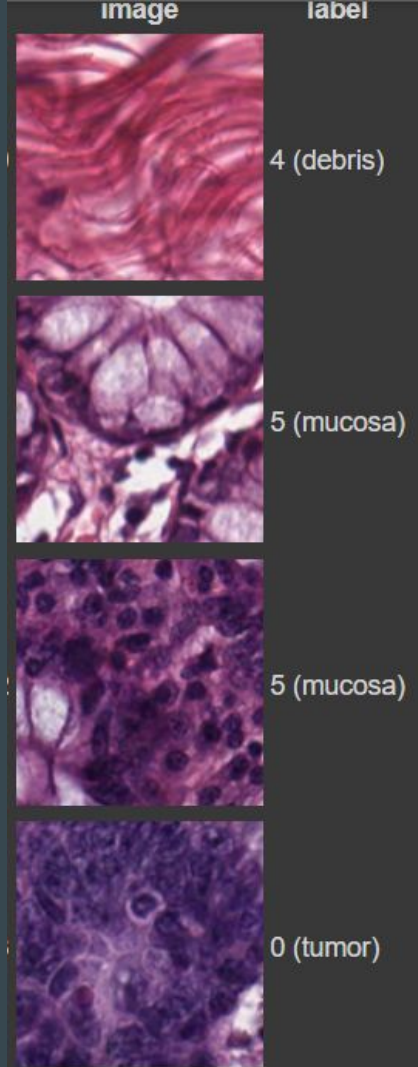
Katrina Wheeler
COMP 4449

Colorectal Cancer Information

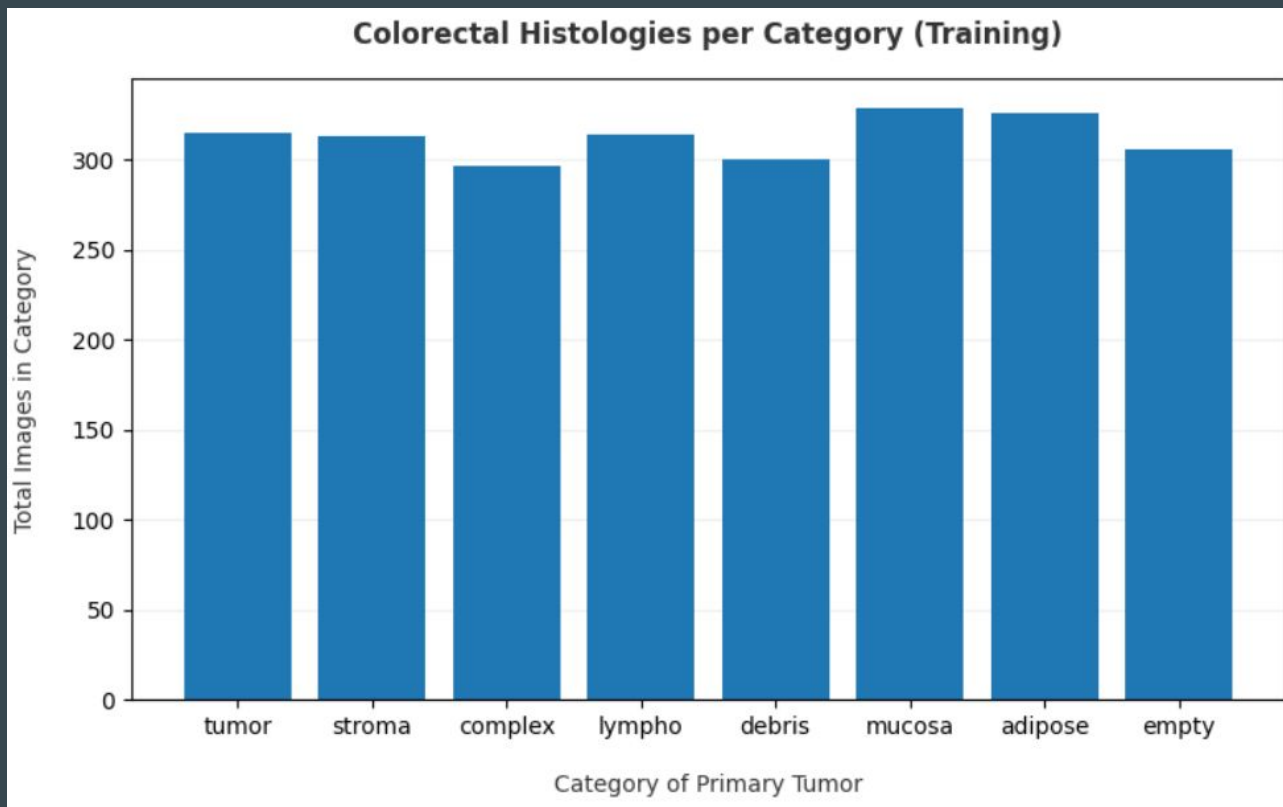
- Mortality rate of 33.33%
- Diagnosis in people under 55 years old is up from 11% in 1995 to 20%
- Early detection can increase survival rate by up to nearly 90%
- Diagnostic process is manual and risks human error

Goals of Analysis

- Utilize Convolutional Neural Network in Tensorflow to classify colorectal images as one of 8 different categories
- Hypertune algorithm for optimal accuracy of classification

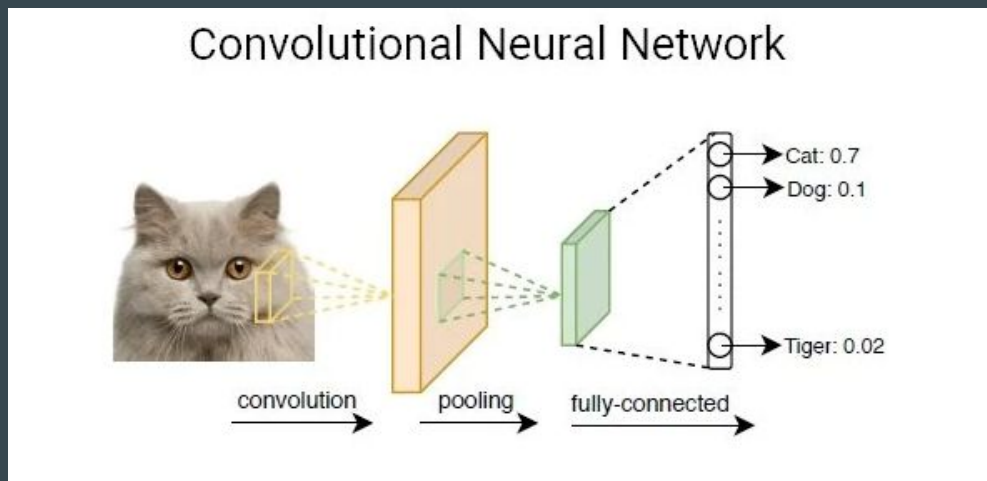


Distribution of Classification - Training Data



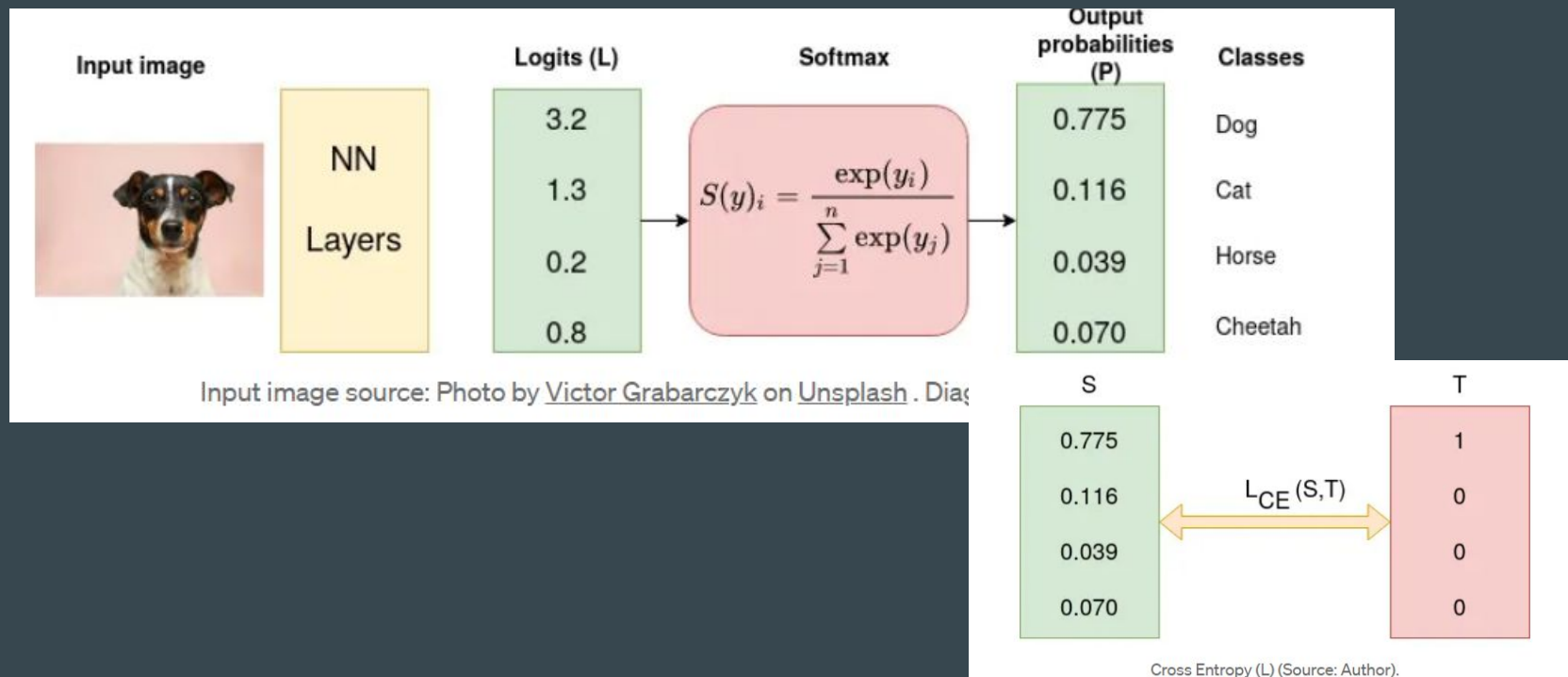
Convolutional Neural Network (CNN)

- Based on human neural and optic systems
- Can learn and train without much human intervention
- Do not require much preprocessing

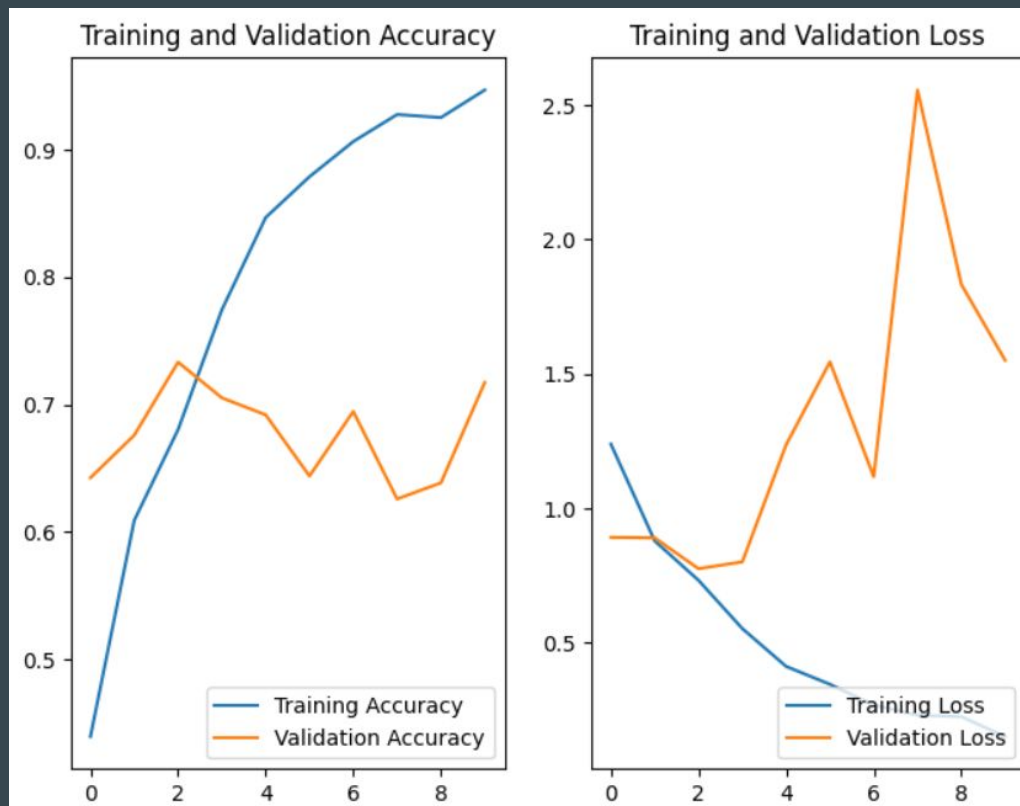


Cross Entropy Loss Function

- Often used as loss function in multi-class classification neural networks
- The model uses this function to optimize the model



Initial Base Model Scores



Overfitting

- There is a large difference between the accuracy of the training and validation data set, which could indicate overfitting
- To combat overfitting, I utilized a dropout layer, as well as initialized early stopping during my hypertuning
 - A dropout layer will drop certain outputs from the layer during the training process; otherwise features that were present in later input data may not be proportionally influential

Hyperparameter Tuning with Keras Tuner

- Number of units for the Dense layer
- Learning rate for the model
- Adding early stopping to assist with prevention of overfitting
- Adjusted rate for drop layer
- Validation accuracy of 74%

The optimal number of units in the first densely-connected layer is 64, the optimal learning rate for the optimizer is 0.001, and the optimal dropout rate is 0.4.

Final Accuracy Metrics & Additional Considerations

- Final test loss: .86, final test accuracy: 66%
- Hyperparameter tuning was limited due to time constraint
- Additional parameter tuning for the convolutional layers, additional depth for the dense layers, and additional learning rates could contribute to increased accuracy.

Image Sources

Boesch, Gaudenz. “A Complete Guide to Image Classification in 2023.” Viso.ai, 30 Jan. 2023,
viso.ai/computer-vision/image-classification/#:~:text=Today%2C%20the%20use%20of%20convolutional. Accessed 5 Nov. 2023.

Koech, Kiprono Elijah. “Cross-Entropy Loss Function.” Medium, 25 Feb. 2021,
towardsdatascience.com/cross-entropy-loss-function-f38c4ec8643e.