Machine Learning for Medical Imaging

Margie Vollkommer
Advisors: Dr. Spanias, Vivek Sivaraman Narayanaswamy
8/5/2022 – Final Presentation
ASU SENSIP REU
Pre-Training at ASU:

- Matlab/Python Introduction
- Machine Learning Introduction
  - Clustering/K Means
  - Classification + Confusion Matrices
  - Regression and Prediction
- Neural Networks
- PCA
Statement of Problem:

- Potentially useful application of machine learning is the diagnosis of ailments when provided with an X-ray or CAT scan.
- Not enough accurately labeled datasets [1]
- Many existing datasets are not diverse, resulting in datasets that possess an implicit bias [1]
- Additionally, many have a hidden stratification, causing a model to only perform well on the most common varieties of a given condition [2]
Proposed Solution:

- MedMNIST v2 set was utilized in this project
  - Open-source dataset with thousands of 2D and 3D medical scans
  - Aimed to achieve accuracy of 90% or greater
  - Current accuracy is between 0.531 and 0.998, depending on the specific dataset
- Investigated different machine learning techniques that would be best suited to medical detection

Figure 1: Scans of human organs for use in medical classification through machine learning. (https://medmnist.com/)
Machine Learning Techniques Used:

- Neural Network on Organ A
- Logistic Regression on Organ A
- Principal Component Analysis on OrganAMNIST dataset
- Convolutional Neural Network on BloodMNIST Dataset
- Principal Component Analysis on BloodMNIST Dataset
Results:

- Significant decrease in accuracy for both logistic regression and neural network models at a variance of ~0.6 or ~100 components
- OrganAMNIST Neural Network accuracy on 34,581 images: 97.09%
- BloodMNIST Convolutional Neural Network Accuracy: 65.73%
Conclusions:

- PCA can be used to reduce the variance of some medical images of organs to ~0.6 and still maintain accurate results.
  - Will result in faster-running algorithms.
- A neural network with a hidden layer of size 500 can produce an accuracy of upwards of 97% on some medical data (images of organs).
- A convolutional neural network can produce an accuracy of 65% on some medical data (images of blood).
Next Steps:

- Improve upon the Convolutional Neural Network accuracy
- Investigate other machine learning techniques that might be useful
- Investigate other medical image datasets
Reflection:

- Learned a lot about machine learning techniques
- Improved my coding skills
- Improved my presentation skills
- Made new friends!
- Chipped my two front teeth while moving from summer housing back to Tooker
References:

