EXPLORATION OF MACHINE LEARNING FOR FUTURE USE IN MEDICAL IMAGING

Margie Vollkommer, SenSIP Center, School of ECEE/BHSE, ASU

***Abstract* – Machine learning has emerged as a potentially life-changing tool for medical diagnosis. The aim of this project was to investigate the potential methods that might best lend themselves to this task. The eventual goal will be to improve the accuracy of ailment detection using machine learning methods trained on labeled data consisting of medical scans.**

***Index Terms*—Machine Learning, Medical Imaging, MedMNIST**

1. BACKGROUND

Machine learning, although a relatively new technology, has rapidly demonstrated its many uses. One potential application of machine learning is the diagnosis of ailments in medical imaging. Ideally, through classification methods, a computer program would be able to identify diseases or ailments when provided with an x-ray or other such scan. This would be very beneficial for overworked doctors, and could act as a potential crutch to aid in giving accurate diagnosis.

Although encouraging progress has been made so far, there are some limitations. Accurately labeled medical datasets are not in abundance, particularly ones large enough to accurately train a program [1]. Additionally, many datasets are sourced from small populations without much diversity, thereby providing the data with an implicit bias [1]. Many data sets also suffer from hidden stratification, where a model will perform well on the most common types of a certain condition, but consistently misdiagnose a less common subtype [2].

The dataset used in this project was the MedMNIST v2 set, as it is an open-source data set with thousands of 2D and 3D images. Our aim is to improve the accuracy of detection, ideally upwards of 90%. Past models using the MedMNIST data produced accuracy scores ranging from 0.531 to 0.998, depending on the data set [3]. Improving the accuracy of machine learning models on medical images would have the potential to save doctors a significant amount of time and money.

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

2. PROJECT DESCRIPTION

Initially, the effect of principal component analysis on the Organ A MedMNIST dataset was investigated. The variance was decreased from 100% to 25%, and the number of components was also manually decreased. This corresponded with a decrease in accuracy after the dataset was run through a logistic regression model as well as a neural network. The results are graphed in the figures below:

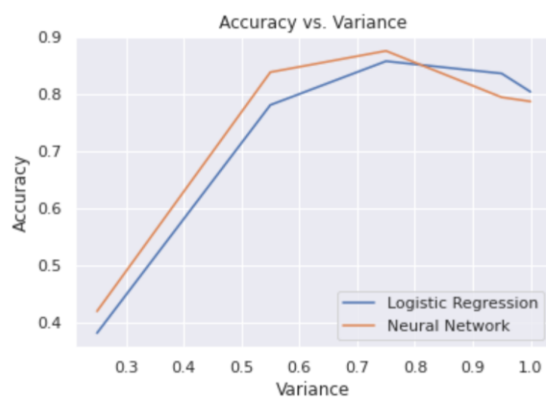

Figure 2: Accuracy vs. Variance of MedMNIST Organ A dataset post-PCA. As variance is decreased, the accuracy of both models decreases.

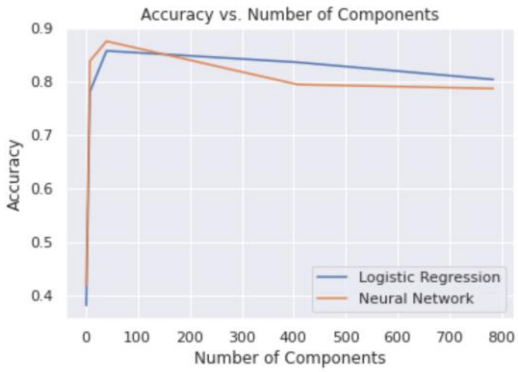

Figure 3: Accuracy vs. Number of Components of MedMNIST Organ A dataset post-PCA.

The results suggest that the variance in datasets composed of this type of medical image can be reduced to about 60% (or around 20 components) without a significant decrease in accuracy. This is promising, as decreasing the variance will allow algorithms to run faster.

Next, using PyTorch, we constructed a simple neural network and trained it on the same dataset. The network achieved 97.09% accuracy on 34,581 images.

In order to attempt the construction of a convolutional neural network (a technique advantageous for image analysis), the BloodMNIST dataset was used, as the images are 3-dimensional rather than black and white.

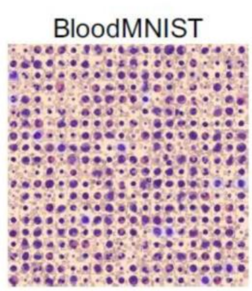

Figure 4: Images of human blood cells for use in medical classification through machine learning. (https://medmnist.com/)

Using PyTorch, we constructed a CNN with two convolutional layers and two pooling layers. Unfortunately, only 65.73% accuracy was achieved, but with adjustments to the layers of the network, we are confident that accuracy will improve.

By harnessing these results, we hope to eventually be able to construct an algorithm through machine learning that is capable of diagnosing medical conditions from images.

References:

[1] Barragán-Montero, Ana, Umair Javaid, Gilmer Valdés, Dan Nguyen, Paul Desbordes, Benoit Macq, Siri Willems et al. "Artificial intelligence and machine learning for medical imaging: A technology review." Physica Medica 83 (2021): 242-256.

[2] Oakden-Rayner, Luke, Jared Dunnmon, Gustavo Carneiro, and Christopher Ré. "Hidden stratification causes clinically meaningful failures in machine learning for medical imaging." In Proceedings of the ACM conference on health, inference, and learning, pp. 151-159. 2020.

[3] Yang, Jiancheng, Rui Shi, Donglai Wei, Zequan Liu, Lin Zhao, Bilian Ke, Hanspeter Pfister, and Bingbing Ni. "Medmnist v2: A large-scale lightweight benchmark for 2d and 3d biomedical image classification." arXiv preprint arXiv:2110.14795 (2021).

Ira A. Fulton Schools of Engineering
ARIZONA STATE UNIVERSITY

REU Site sponsored in part by NSF I/UCRC Award 1540040