UNCERTAINTY QUANTIFICATION IN ENSEMBLE DEEP LEARNING

NSF REU - DEIM: Research Projects in Data-Enabled Industrial Mathematics

Embry-Riddle Aeronautical University
Rachel Swan
Dr. Berezovski

Nevada National Security Site
Jesse Adams, PhD
Maggie Lund, PhD

Support for the program has been provided by the National Science Foundation (NSF) through REU Award Number DMS - 2050754.

PROGRESS REPORT

INTRODUCTION

The Nevada National Security Site (NNSS) is a United States Department of Energy reservation located approximately 65 miles northwest of Las Vegas. NNSS specializes in defense nuclear nonproliferation and the pursuit of science. They cannot conduct full-scale nuclear tests, but instead conduct controlled underground experiments and computer simulations to verify that the arms are safe and usable.

NNSS is interested in using neural networks to analyze radiographic images, which is one method of data collection done during controlled experiments. For our neural network, we will be using a widely-used image dataset. However, we hope that our results can still be applied to the research done at NNSS.

PROJECT SCOPE

After background research is collected, a neural network will be created using Python and other packages like Keras/Tensorflow. The network architecture will be trained many times, with a confidence interval given for each trained architecture. Over time, the architecture will be made more sophisticated and an uncertainty quantification (UQ) approach will be implemented to verify the accuracy of the network. Finally, a probability model will be created with a UQ approach.

DATASET

The dataset used is synthesis, optical imaging, and absorption spectroscopy data for 179072 metal oxides [1]. The data contains 180902 images that are size 64 by 64, with 3 channel (RGB). Some of the images are blank, leading to 179072 total samples. The data is normalized from 0-1 for every channel. The color in the middle of each image represents the color of the printed sample, while the "coffee ring" around each sample is darker due to drying. Figure 1a presents an example of the images contained in the dataset.

Figure 1a: An example set of the images contained in the data set [1].

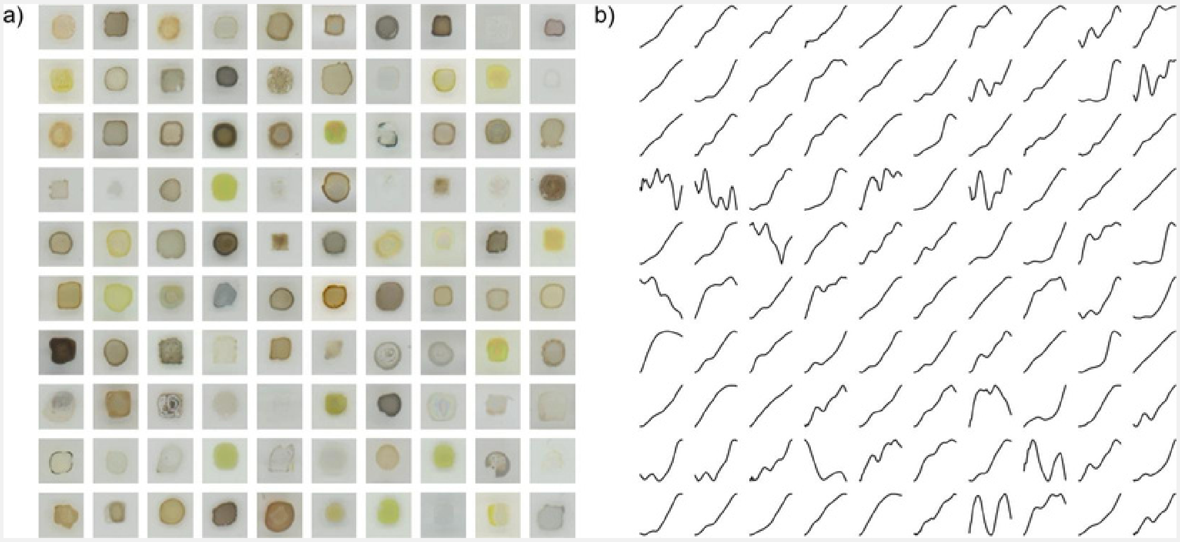

Figure 1b: An example set of spectra outputs from the data set [1].

The outputs of the data set are spectra that were originally 220 points. An example of the spectra are shown in Figure 1b. However, for computational simplicity and to reduce data oscillation, the neural network outputs 20 values with linear interpolation between each value.

INITIAL RESULTS

The initial strategy was to develop a simple neural network and experiment with the architecture. A ten image portion of our single neural network is shown in Figure 3. After deciding on a single architecture, an initial ensemble neural network was created and trained n = 2 times on 10,000 images. The neural network architecture contains a convolutional layer, a dense layer, a max pooling layer, a flatten layer, and three more dense layers. The convolutional and dense layers use three ReLU activation functions and one Softsign activation function. Then, the outputs of each neural network were averaged to generate results. The output of the ensemble neural network compared to the outputs of the individual neural network is shown in Figure 4.

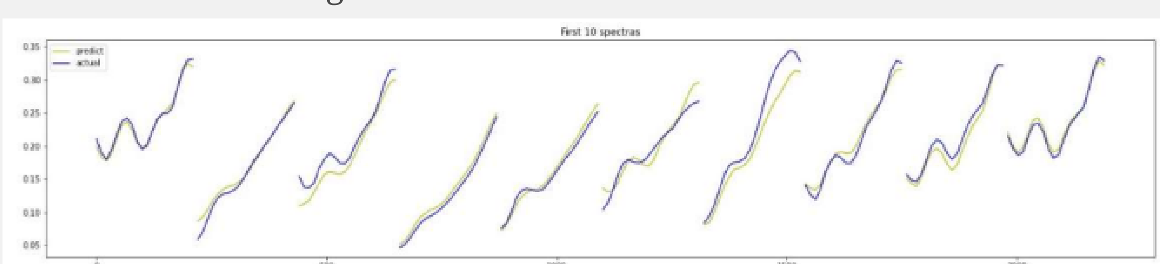

Figure 3: An example set of ten images produced by a single neural network. The yellow line represents the predicted spectra output, while the blue line represents the actual output.

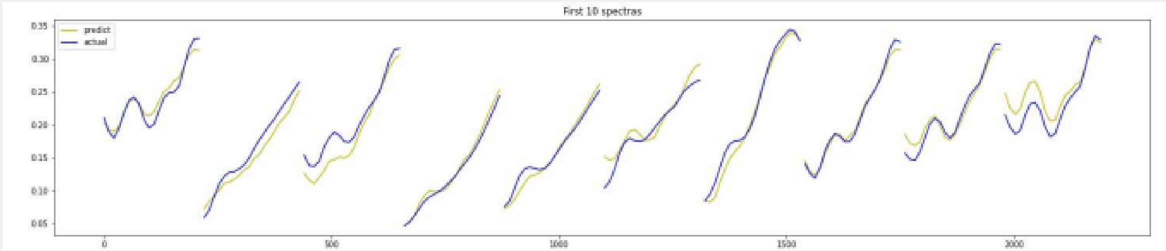

Figure 4: An example set of ten images produced by an ensemble neural network. The yellow line represents the predicted spectra output, while the blue line represents the actual output.

Additionally, the ensemble neural network was compared to the singular neural network. Figure 5 displays a scatter plot of the data points of the ensemble network compared to the single neural networks. The distance of each data point is shown in comparison to the predicted (black) line.

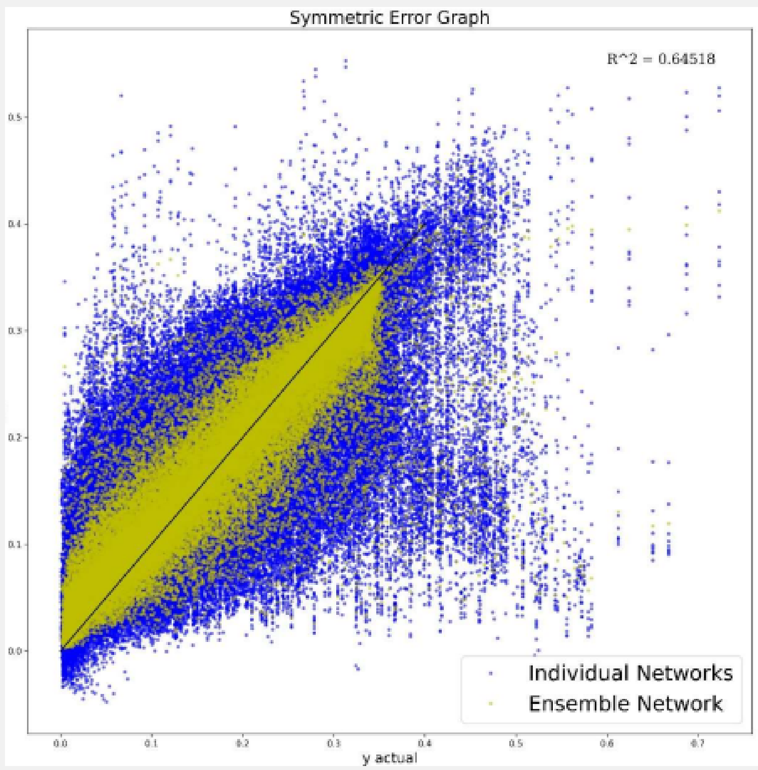

Figure 5: The error scatter plot of prediction error. The black line is the actual value. The yellow points are predicted values of the ensemble neural network. The blue points are predicted values of single neural networks.

LIMITATIONS

The network has greater trouble predicting spectra that have many curves. These spectra are usually produced by copper and niobium. Therefore, copper and niobium oxide data were excluded in the testing and validation of the model.

UNCERTAINTY QUANTIFICATION

While training the data, mean squared error (MSE) was used to evaluate the accuracy of the data. Figure 6 displays the mean squared error of the training data.

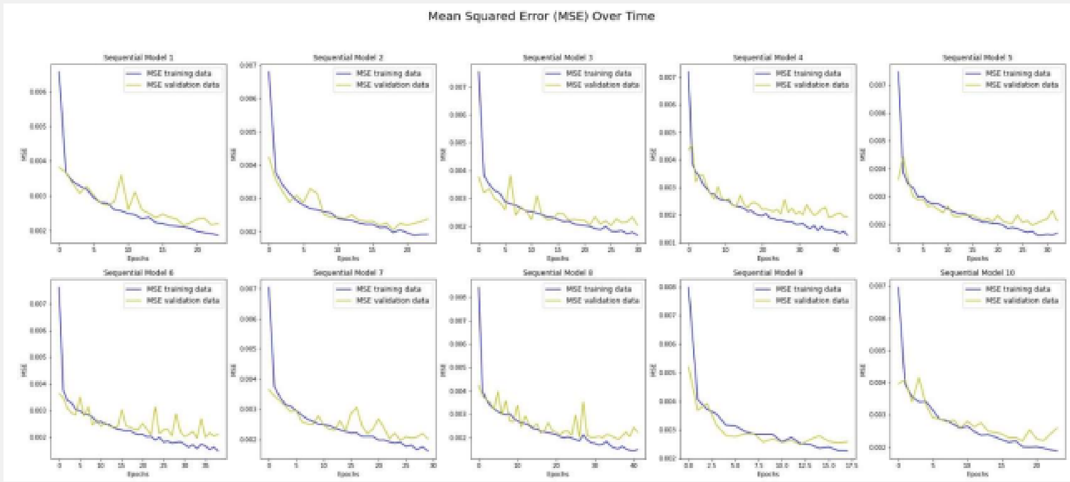

Figure 6: The MSE of each neural network. The blue line represents the training data error, while the yellow line represents the validation data error.

Each neural network had a final mean squared error of 0.003 to 0.001. This is a low a value, indicating that the network had only 0.1-0.3% error on the training data.

For each spectra, a 95% confidence interval was defined. Then, it was calculated what percentage of the neural network's prediction points were inside the 95% confidence interval. Figure 7 shows a sample confidence interval graph for a prediction. Most predictions had 7-8% of their points inside the 95% confidence interval, indicating that the network is unreliable.

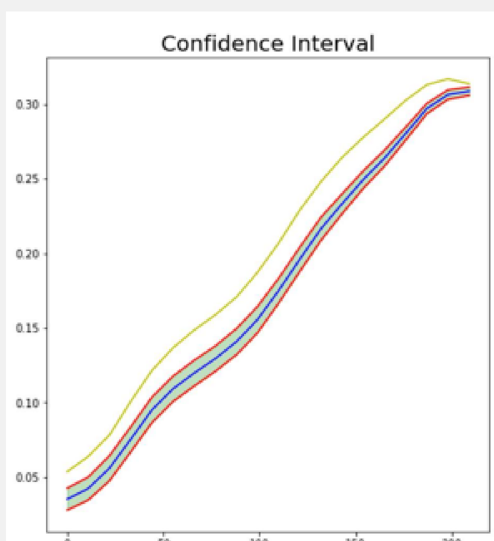

Figure 7: A sample of an expected graph (blue) with a 95% confidence interval defined around it (green). The prediction (yellow) is compared.

REFERENCES

[1] Stein, H. S., Soedarmadji, E., Newhouse, P. F., Guevarra, D. & Gregoire, J. M. Synthesis, optical imaging, and absorption spectroscopy data for 179072 metal oxides https://doi.org/10.6084/m9.fgshare.7502207 (2019).