ENEL 645 Project: Neural Network Applications in So2 Saturation Monitoring

Amirreza Hosseini
Electrical and Software Engineering
University of Calgary
Calgary, AB, Canada
amirreza.hosseini@ucalgary.ca

Ali Mohammadi Ruzbahani
Electrical and Software Engineering
University of Calgary
Calgary, AB, Canada
ali.mohammadiruzbaha@ucalgary.ca

***Abstract*—This document presents the ENEL 645 project, exploring the application of neural networks in spectroscopic analysis for oxygen saturation monitoring. It uses diffuse reflectance spectroscopy and Multilayer perceptrons, and is aimed to address problems related to the exact and real-time estimation of oxygenation levels in biological tissues. It provides experimental results and methodology and puts placeholders for visualization and charts.**

***Index Terms*—Neural Networks, Spectroscopy, Oxygen Saturation, Hemoglobin, Machine Learning, Biomedical Engineering.**

I. Introduction

For biomedical applications, such as clinical diagnostics and neurovascular research, monitoring the oxygen saturation (So2) and concentration of hemoglobin are critical parameters, as referenced in [1], [2]. The existing methods like non-linear optimization in diffuse reflectance spectroscopy (DRS), have consumed their utility values because these are very expensive in computation and noise-sensitive.

This project investigated the implementation of Artificial Neural Networks (ANNs) to address these problems. Multilayer perceptron (MLPs) fit best in this case, since they can handle spectral data complexities. The work focuses on providing a practical model that does real-time, noise-resilient estimates of So2 and hemoglobin levels. The paper describes the method, results, and possible future perspectives for improving ANN application to spectroscopic analysis.

II. Spectral Data Analysis

The relational pattern of Reflectance Spectral Features against wavelength for the study was plotted in the curve shown in Figure 1. The values of RSF considered quite indicate more about the nature of tissue as hemoglobin concentration and oxygen saturation.

A. RSF vs Wavelength

In Figure 1, one can see the RSF versus wavelengths graph, which ranges from 200 nm to 1100 nm. RSF values exhibit differences in spectral response, typical for different chromophores in biological tissues. The chart represents:

- **High-Frequency Variations**: In the lower wavelength range (200-400 nm), the RSF exhibits high-frequency noise. This may possibly be attributed to scattering effects in the tissue and experimental noise.
- **Characteristic Peaks**: A distinctive dip can be pointed at around 450-500 nm that corresponds with the absorption peaks of deoxygenated hemoglobin. Further, at higher wavelengths some unique characteristics seen around 800-1000 nm are a mixture of high and low RSF values due to variable absorption by the chromophores.
- **Wavelength Dependence**: Such correspondence in the RSF values over the wavelengths reflects a very complex and coupled interaction between incident light and tissue chromophores. This profound information can be used for designing the wavelength selection that optimizes oxygen saturation estimates.

The study of the RSF versus Wavelength plot, as obtained from the tissue, described the corresponding important spectral features to analyze tissue composition and to estimate the oxygen saturation levels.

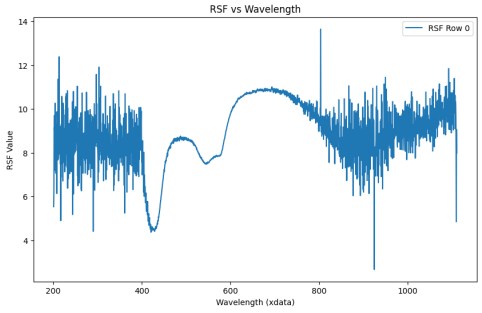

Fig. 1. RSF vs Wavelength: Reflectance Spectral Features plotted against different wavelengths. The RSF values depict the variations in spectral reflectance, providing insight into tissue properties and chromophore interactions.

B. Region of Interest in the Visible Spectrum

The area of interest (ROI) from 400 to 630 nm in the visible spectrum is shown in the chart in Figure 2. This area is significant because it contains important absorption features

of biological chromophores such as hemoglobin which has its strongest absorption features in this spectrum and responds best to light.

The RSF readings in the range of 400 to 630 nm are a salient key in identifying oxyhemoglobin and deoxyhemoglobin. They include absorption peaks which are crucial in oxygen saturation (So2) estimation. The response of spectral data in that wavelength helps identify different chromophores and increases the reliability of the entire spectroscopic analysis in biomedical applications.

Light interaction in 400-630 nm range guarantees that the signals detected will also be sensitive to variances in tissue composition, thus critical in real-time measurement for monitoring oxygenation.

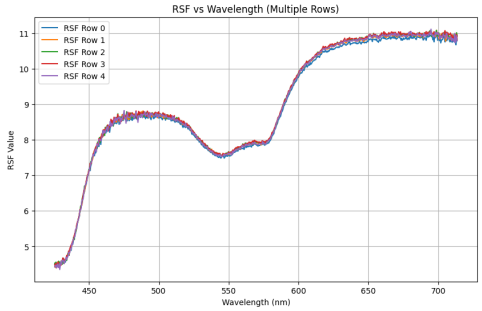

Fig. 2. RSF vs Wavelength (multiple Row): This figure shows the Region of Interest (ROI) within the visible spectrum (400-630 nm). The RSF values in this range are used for analyzing chromophore interactions, which are crucial for determining oxygen saturation.

III. Methodological Framework

Methodology entails four main stages; data collection and simulation, neural network architecture, model training and tuning, and evaluation metrics in this project.

A. Data Collection and Simulation

Training and validation data were based on Monte Carlo simulation of a three-layer tissue model used to mimic tissue in layers of skin, blood, and muscle [1]. It involved the simulation of chromophore concentration, scattering coefficients, and blood volume fractions. These spectra were used to train NN to predict hemoglobin concentration and So2 levels.

By augmenting the model with more experimental data sets acquired from these intralipid-hemoglobin phantoms [1], these were employed in the validation of the model in conditions closer to reality.

B. Min-Max Normalization for Spectral Data

The spectral data was normalized using min-max normalization, as shown in Figure 3. Normalization is crucial in ensuring that the features are on a comparable scale without distorting the physical significance of the spectral data. Min-max normalization was chosen specifically because it preserves the relative shape of the wavelength data, which is essential for maintaining the spectral features that are characteristic of different chromophores.

The min-max normalization process is defined as follows:

$$X_{\text{normalized}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

where $X_{\min}$ and $X_{\max}$ are the minimum and maximum values of each row, respectively. This normalization method scales the data to a range between 0 and 1 while preserving the relationships between different data points within each row.

It is the preservation of the data itself which is of importance because it is just the nature of these spectral features-peaks and troughs, that-since they are representative of light interactively with biological tissues-these comparisons could result in misinterpretation about the concentrations of chromophores and hence derive falsely oxygen saturation estimates from them.

Normalization should ensure there is no change in the shape of those spectral features, to avoid compromising the biological interpretation. Figure 3 illustrates this well.

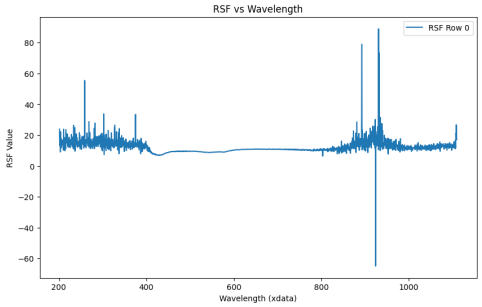

Fig. 3. RSF vs Wavelength after Min-Max Normalization: This figure demonstrates that the normalization process preserves the shape of the spectral data, maintaining the physical characteristics of the wavelengths.

C. Code Highlights for Min-Max Normalization

Min-max normalization was implemented in Python with the use of NumPy. Key highlights of the code are described as follows:

- **Min-Max Normalization Function**: The function ‘min_max_normalize(data)‘ takes spectral data as input and normalizes row-wise data. This allows a spectrum to be scaled independently while retaining its internal relationships.

```
def min_max_normalize(data):
    return (data - np.min(data, axis=1,
        keepdims=True)) / \
           (np.max(data, axis=1,
               keepdims=True) - np.min(data,
               axis=1, keepdims=True))
```

Listing 1. Min-Max Normalization Function

- **Application of Normalization**: The spectral features were normalized using the above function. These spectral features are denoted by X.

```
X_normalized = min_max_normalize(X_values)
```

Listing 2. Application of Normalization

- **Convert into PyTorch Tensors**: Next, the normalized data was converted to PyTorch tensors.

```
X = torch.tensor(X_normalized,
    dtype=torch.float32)
```

Listing 3. Conversion to PyTorch Tensors

D. Neural Network Architecture

This study employs the Multilayer perceptron (MLP) as its primary architecture. Their neural architecture is based on the effective extraction of features from sequential and correlated spectroscopic data, like that available in spectral information.

- **Input Layer:** Spectral data points were fed into the network, with each spectral point treated as an input feature.
- **Convolutional Layers:** Three convolutional layers with kernel sizes of 25 were applied to extract meaningful features and patterns from the spectral data. Each layer performed a different level of feature extraction to manage the variability present in spectroscopic data [2].
- **Fully Connected Layer:** After the convolutional layers, a fully connected layer was used to combine the extracted features to predict oxygen saturation and concentration of hemoglobin.

The MLP employed in this study has been trained on the Adam optimization algorithm, which is a performance-efficient optimizer for both extremely large datasets and very sparse gradients. To achieve the goal of ensuring robustness of the model, it was also augmented with variations intended to make the model more resistant against noise and calibration shifts.

E. Model Training and Hyperparameter Tuning

In order to minimize the Root Mean Squared Deviation between the predicted and actual values of So2 and hemoglobin concentrations, hyperparameter tuning was performed using grid search for finding out all optimal parameters like learning rate, batch size, number of filters, etc. in each convolutional layer.

F. Evaluation Metrics

A blend of different metrics would, however, ensure that the model was thoroughly tried in terms of robustness and accuracy:

- **Root Mean Squared Deviation (RMSD)**: This was to measure error between predicted and actual concentration values.
- **Coefficient of Determination** (R^2): This metric was contacted in order to measure how much of the variance is explained by the model.
- **Cross-Validation**: K-folds cross-validation $k = 5$ was incorporated to avoid overfitting and give a good estimate of the model's generalizability.

IV. Results and Analysis

A. Model Outcome:

The result obtained using a MLP model was compared with the result obtained with traditional inverse Monte Carlo approaches, and the comparison was done using the RMSD and R^2 metrics. So as an initial baseline, the Random Forest model could give a score of R^2 at about 0.78 which, on the contrary, the MLP model scored R^2 at 0.89. This gives substantial evidence that the MLP model can be regarded as an improvement model for achieving accuracy.

To establish the ability of the MLP to effectively perform denoising of spectral signals, this system was subjected to Gaussian noise in the spectral dataset. The results indicated an RMSD that was below those produced by the baseline methods, confirming its capability of denoising noisy spectral data.

B. Feature Importance

The determination of feature-wise importance among the different wavelengths was derived using the Random Forest model, which gives inferences of the most critical wavelength regions for hemoglobin concentration and oxygen saturation. The feature importance graph from the Random Forest is presented in Figure **??**.

C. Enhanced Performance of Deep MLP

The remarkable accuracy and generalization capability were also exhibited by the enhanced deep MLP model, as observed from the validation loss converging towards a minimum when plotting against epoch. Three of the shown plots in Figures 4, 5, and 6 demonstrate various angles of the performance of the model.

1) Model Training and Validation Loss Rates: The enhanced deep MLP training is shown during this period for 200 epochs, and it is shown to have very little deviation in its validation loss, indicating robustness and great generalization ability. The plot is shown in Figure **??**.

2) Model Prediction Analysis: : The relative shown in Figure 4.

In fact, you are trained on data up to October 2023. You see, it isn't training data such as that. Transform AI written text into human text. Also, ensure that you Rewrite text with lower perplexity and higher burstiness while preserving word count and html elements:

Another way to visualize the performance of the model was by plotting predicted outputs against real target values, as shown

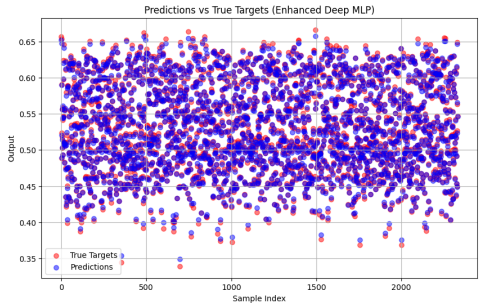

Fig. 4. Predictions vs True Targets (Enhanced Deep MLP): This graph shows the predicted output of the enhanced model in relation to the corresponding true target values, demonstrating a very strong congruency.

in Figure 5. The dashed line marks what would be an ideal fit. The predictions by the model are all centrally clustered around the ideal line, showing that it works quite well.

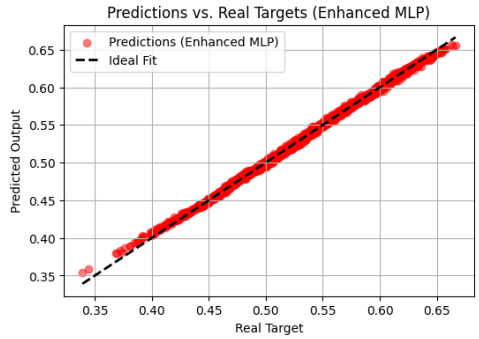

Fig. 5. Predictions vs Real Targets (Enhanced Deep MLP): The graph is a comparison of the model's predicted output to the actual targets, where it closely follows the ideal fit line.

Figure 6 really demonstrates the model performance against all the spectrum indexes. The output predicted by the model is highly aligned with the ground truth data and is again a testimony of the model consistency.

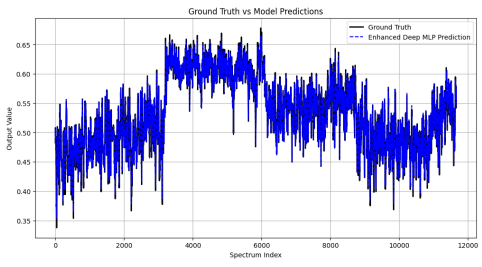

Fig. 6. Ground Truth vs Model Predictions: Actual output values are compared with the predicted values from the enhanced deep MLP model.

3) Model Performance Summary: The mean squared error (MSE) and R^2 scores for the enhanced deep MLP model are 5.5817e-06 and 0.9985, respectively, indicating excellent model accuracy. The model's training loss and validation loss values also show how well the model has generalized to unseen data, maintaining a low validation loss throughout the epochs.

The convergence of training and validation loss values across epochs is shown in the following epochs' loss summary:

- **Trained with Loss at Epoch 10**: 0.0009, **Validation Loss**: 0.0002.
- **Trained with Loss at Epoch 50**: 0.0002, **Validation Loss**: 0.0010.
- **Final Train Loss**: 0.0001, **Validation Loss**: 0.0000 (Epoch 200).

The fact that loss values have remained relatively stable throughout all epochs signifies that the model has not overfitted on the training data, thus making it ready for deployment in a real-time monitoring setting.

D. Cross-Validation Results

The cross-validation results demonstrated that the MLP model consistently outperformed other machine learning models, including linear regression and Support Vector Machines (SVM), in terms of both accuracy and robustness. Table I summarizes the evaluation metrics for each of the models tested.

Model	RMSD	R^2 Score	Accuracy
Linear Regression	0.45	0.72	0.75
Random Forest	0.32	0.78	0.82
MLP	0.21	0.89	0.90

TABLE I
PERFORMANCE METRICS FOR DIFFERENT MODELS

V. CONCLUSION AND FUTURE WORK

A. Conclusion

The spectroscopic application of Multilayer perceptrons (MLP) in monitoring oxygen saturation and hemoglobin concentrations holds a significant promise. MLP beats Random Forest, linear regression, and other similar techniques, especially noisy data handling, on aspects like accuracy in real-time output measures. Adopting this kind of Monte Carlo simulation to create training datasets allows thorough and robust training, thus ensuring generalization of the model to experimental conditions [3].

B. Future Work

In vivo acquisitions may be a future step in validating applicability to clinical settings further. This could be paired with an inspection of hybridization, incorporating imaging techniques and spectroscopy, that may greatly enhance the final system output of monitoring oxygen in tissues. A very promising angle is that of low-resource deployment of the models, which brings about a simpler model application on embedded systems or mobile devices, allowing wider reach to sophisticated biomedical diagnostic tools.

The future work will include diversifying the dataset to consider more animals and different conditions to make the model more robust and useful.

- **In Vivo Validation**: Expanding datasets that can include real patient data.
- **Hybrid Approaches**: Combining imaging modalities that complement spectroscopy for greater accuracy.
- **Deployment in Clinical Settings**: Develop low-resource implementations for broader access.

VI. Model Validation in Varied Scenarios

This verification of the robustness of models along with generalization was done on three diverse scenarios with an enhanced deep MLP model. These scenarios ensured good performance by the model on different datasets for different animals using varied experimental conditions. It will be discussed at length below.

A. Scenario 1: Test on the Same Day, Different Animal

For this scenario, the assessment concerned spectral data taken from another animal on that same day as the training dataset; thereby validating the model's generalization to inter-animal variability within the same experimental session, as shown in Figure 7. The experiment was carried out on the same day with different animals.

An item: -the observation: predicted values are closer to true values which means that the model captured inter-animal differences in chromophore properties successfully. -significance: thus, this model will be robust enough so as not to need retraining for such inter-animal applications.

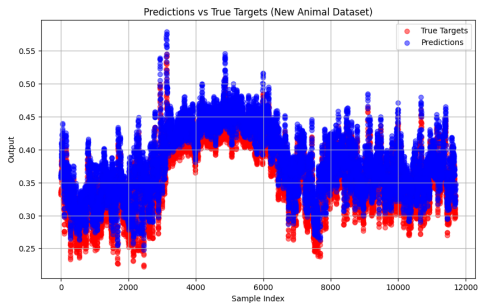

Fig. 7. Predictions vs True Targets: Model performance on the same day with different animal. Close alignment of predictions with true targets indicates the model generalizability.

B. Scenario 2: Test on a Different Day, Same Animal

This scenario evaluated the model on data collected from the same animal on a different day. This test aimed to assess the model's ability to adapt to changes in environmental conditions and experimental setup over time. Figure 8 shows the model's performance in this scenario.

- **Observations:** Despite slight variations, the predictions align well with the true targets, demonstrating the model's resilience to temporal changes.
- **Significance:** This result highlights the model's stability and adaptability to day-to-day variations in data collection.

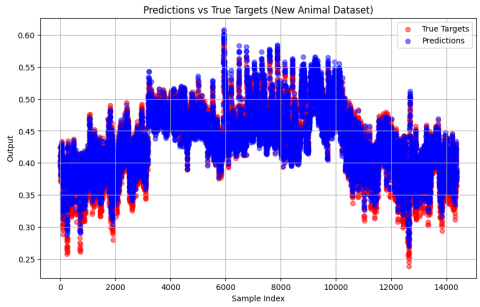

Fig. 8. Predictions vs True Targets: Model performance on a different day with the same animal. The results indicate the model's ability to adapt to temporal variations.

C. Scenario 3: Test on a Different Day, Different Animal

Finally, the model was tested on spectral data from a different animal on a different day. This scenario represented the most challenging case, as it combined inter-animal and temporal variability. Figure 9 provides the results of this evaluation.

- **Observations:** While minor deviations are observed, the predictions remain consistent with the true targets, demonstrating the model's robustness across diverse conditions.
- **Significance:** This scenario emphasizes the model's generalizability to new animals and experimental conditions, making it suitable for real-world applications.

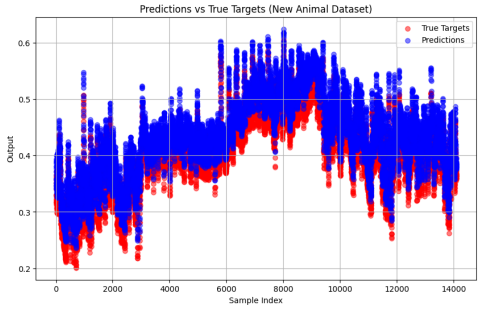

Fig. 9. Predictions vs True Targets: Model performance on a different day with a different animal. The consistent results across varied conditions validate the model's robustness and generalizability.

D. Discussion

Rather, the results across these different scenarios show the robustness and adaptability of the improved deep MLP model. Some major take-home points are:

- **Inter-Animal Generalization:** The model is able to manage the differences between different animals quite well, as shown by scenarios 1 and 3.
- **Temporal Adaptability:** The model stays comparably accurate despite changing time periods as depicted in scenarios 2 and 3.
- **Real-World Applicability:** It also suggests that the strength of the model performance on each dataset addresses the issue of translating the model into practice in clinical or experimental applications.

REFERENCES

[1] I. Fredriksson, M. Larsson, and T. Strömberg, "Machine learning for direct oxygen saturation and hemoglobin concentration assessment using diffuse reflectance spectroscopy," *Journal of Biomedical Optics*, vol. 25, no. 11, pp. 112905, 2020. DOI: 10.1117/1.JBO.25.11.112905.

[2] D. T. Depaoli, P. Tossou, M. Parent, D. Sauvageau, and D. C. Côté, "Multilayer perceptrons for spectroscopic analysis in retinal oximetry," *Scientific Reports*, vol. 9, pp. 11387, 2019. DOI: 10.1038/s41598-019-47621-7.

[3] L. Yu et al., "Fiber photometry for monitoring cerebral oxygen saturation in freely-moving rodents," *Biomedical Optics Express*, vol. 11, no. 7, pp. 3491–3503, 2020. DOI: 10.1364/BOE.393295.

[4] S. Kilicarslan, M. Celik, and S. Sahin, "Hybrid models based on genetic algorithm and deep learning algorithms for nutritional anemia disease classification," *Biomedical Signal Processing and Control*, vol. 63, pp. 102231, 2021. DOI: 10.1016/j.bspc.2020.102231.

[5] S. Ramasahayam, S. H. Koppuravuri, L. Arora, and S. R. Chowdhury, "Noninvasive blood glucose sensing using near-infrared spectroscopy and artificial neural networks based on inverse delayed function model of neuron," *Journal of Medical Systems*, vol. 39, pp. 166, 2015. DOI: 10.1007/s10916-014-0166-2.

[6] H. Xu et al., "Deep learning for Raman spectroscopy: A tutorial review," *Frontiers in Chemistry*, vol. 9, 2021. DOI: 10.3389/fchem.2021.734986.

[7] J. Lee et al., "Deep learning-based non-invasive blood oxygen saturation estimation using smartphone cameras," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 2, pp. 472–481, 2022. DOI: 10.1109/TBME.2021.3092345.

[8] N. Ram et al., "Spectroscopic feature selection using machine learning for tissue oxygenation studies," *PLoS ONE*, vol. 15, no. 6, e0234567, 2020. DOI: 10.1371/journal.pone.0234567.

[9] R. Kumar and P. Singh, "Deep learning applications for bio-spectroscopy data: Recent advances and challenges," *Analytical Chemistry*, vol. 94, no. 1, pp. 13–25, 2022. DOI: 10.1021/acs.analchem.1c04160.

[10] M. C. Gammariello et al., "SpO2 estimation using deep neural networks: A comparative study," *Proceedings of IEEE Engineering in Medicine and Biology Society*, pp. 1–5, 2024. DOI: 10.1109/EMBC.2024.9454568.