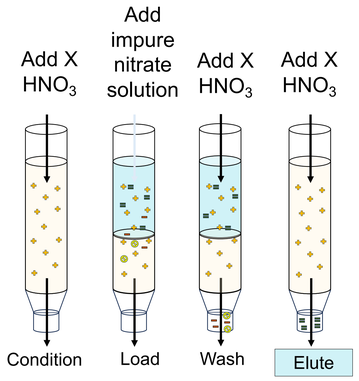
Convolutional Neural Networks for UV-Visible Spectroscopic Analysis of Chromatographic Effluents

Ashlyn Campbell[1], Camila C. Gonzalez[2], Saurabh Narain[2], and Tsuyoshi A. Kohlgruber[2]

I. Goals

- ✓ Model absorbance spectra of chromatographic effuents with unknown concentrations
- ✓ Determine modelling techniques to learn spectral patterns and predict analyte concentrations
- ✓ Support real-time monitoring and decision-making for experimentalists



II. Data Preparation

Nitric Acid (HNO_3): Non-linearly correlated; Peak heights in Nd^{3+}
Neodymium (Nd^{3+}): Overall linear with non-linear regions
Cerium (Ce^{4+}): Non-linearly correlated; limited literature

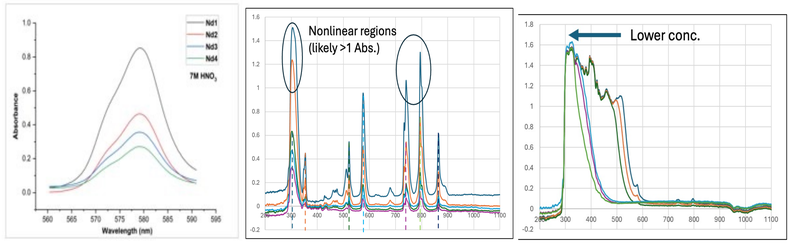


Fig. 1. Properties of raw or original spectra, left to right: Nd^{3+} changing concentration in HNO_3, Nd^{3+} non-linear regions, and a region of low concentration in Ce^{4+}.

1) Baseline Correction: Savitzky Golay (see Fig. 2)

$$Window\ Length = 189$$
$$Polynomial\ Order = 2$$
$$Derivative = 1^{st}$$

2) Min-Max Normalization

$$\boldsymbol{X'}, \boldsymbol{Y'};\ [0,1]$$

3) Band Selection: Mutual Information

$$M_{j,i} = \sum_{x_i}\sum_{y_j} p\,(x,y)\ log\left(\frac{p(x,y)}{p(x)p(y)}\right)dxdy$$

$$X_{Weighted} = X\ *\ Scale$$

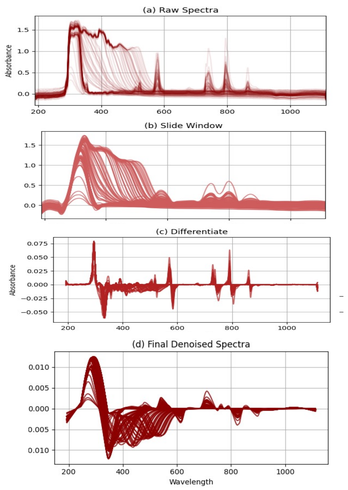


Fig 2. Analysis of Baseline Correction or Denoising: (a) Raw Spectra, (b) Slide Window, (c) Differentiate, and (d) Final Denoised Spectra

III. Modelling

Convolutional Neural Networks (CNNs) are deep learning models that use convolutional layers to extract spatial or temporal patterns from input data. In this work, a 1D CNN is trained on preprocessed spectral data to predict multiple target concentrations (Wang, 2022).

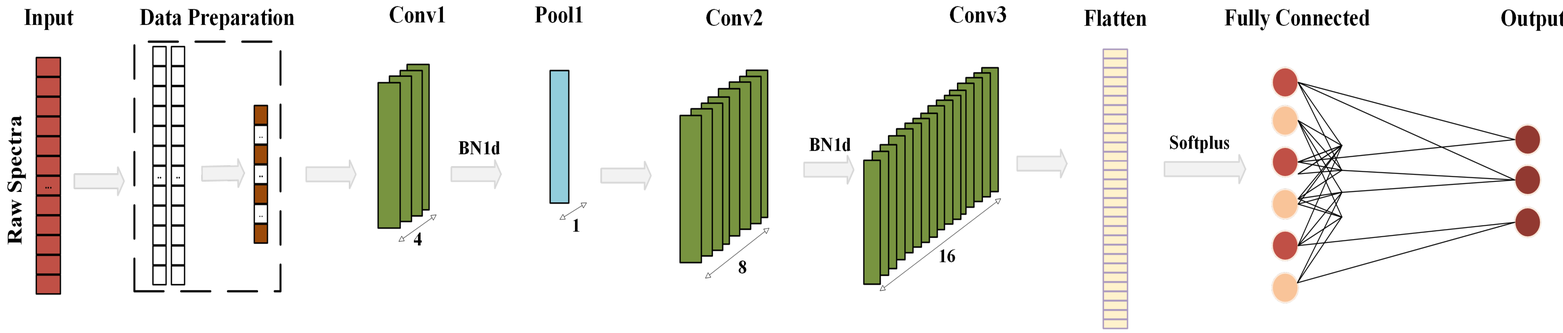


Fig. 3. *Architecture of UVVisCNN, a 1D convolutional neural network (CNN) used for predicted preprocessed UV-Vis spectra.*

IV. Preliminary Results

Model performance was evaluated using **RMSEP** (prediction accuracy), **RPD** (model precision), and **R^2** (fit quality).

Analyte	R^2	RMSEP	RPD
HNO_3	0.9686	0.3898	5.7752
Nd^{3+}	0.9361	0.0031	4.0461
Ce^{4+}	0.9621	0.0094	5.2585

Table 1. *UVVis CNN Performance*

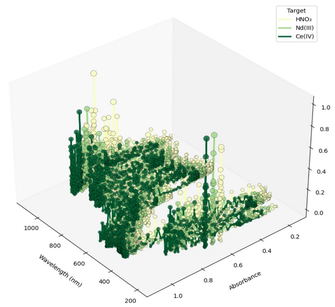


Fig. 5. *A saliency map showing key spectral regions influencing the UVVisCNN predictions*

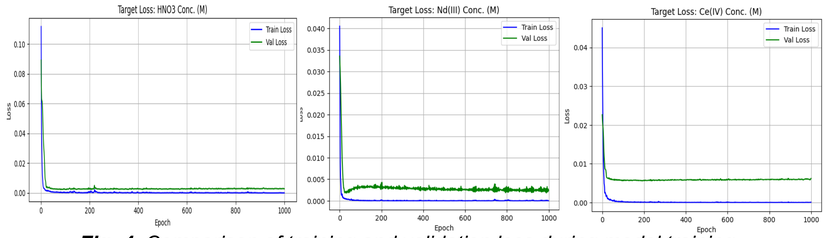


Fig. 4. *Comparison of training and validation loss during model training, illustrating the model's progress and its ability to generalize to unseen data.*

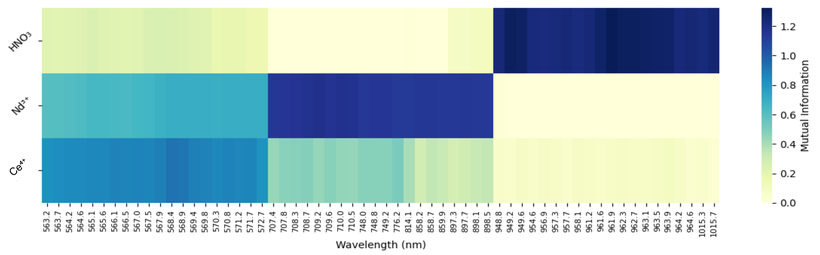


Fig. 6. *Mutual information scores of the top-selected wavelengths for each target, highlighting the spectral regions most relevant to the model's predictions.*

V. Future Work

- ❑ Apply modelling methods to new and more complex analytes.
- ❑ Explore the use of pretrained models (e.g., ResNet) for improved feature extraction and broader applicability (Kazemzadeh, 2022).

VI. Acknowledgements

I would like to sincerely thank the GEM fellowship, the Data Science Summer Institute and the incredible SPADE team for pairing me with amazing mentors.

References

Dayananda, B., Owen, S., Kolobaric, A., Chapman, J., & Cozzolino, D. (2023). Pre-processing applied to instrumental data in analytical chemistry: A brief review of the methods and examples. Critical Reviews in Analytical Chemistry, 54(8), 2745–2753. https://doi.org/10.1080/10408347.2023.2199864

Kazemzadeh, M., Hisey, C. L., Zargar-Shoshtari, K., Xu, W., & Broderick, N. G. R. (2022). Deep convolutional neural networks as a unified solution for Raman spectroscopy-based classification in biomedical applications. Optics Communications, 510, 127977 https://doi.org/10.1016/j.optcom.2022.127977

Wang, Y., Li, M., Ji, R., Wang, M., Zhang, Y., & Zheng, L. (2022). Mark-Spectra: A convolutional neural network for quantitative spectral analysis overcoming spatial relationships. Computers and Electronics in Agriculture, 192, 106624. https://doi.org/10.1016/j.compag.2021.106624

Lawrence Livermore National Laboratory

Prepared by LLNL under Contract DE-AC52-07NA27344.