



Contribution ID: 17

Type: **not specified**

Exploring CNN architectures for NIR based chemometric tasks - the Deep Tutti-Frutti application.

Friday, 30 May 2025 10:25 (20 minutes)

Convolutional Neural Networks (CNNs) have been increasingly used to build NIR based chemometric models with applications ranging from chemical sample analysis to food quality control. In the latter, NIR spectroscopy combined with CNNs enable rapid, non-destructive SOTA predictions of important quality parameters such as dry matter content in fruit [1].

The lack of a standard CNN architecture optimized for spectral data, makes the use of these algorithms a non-trivial task. In this work we approach this problem by performing a joint neural architecture search (NAS) and hyperparameter optimization (HPO) to find an appropriate architecture for the task of dry matter (DM) prediction in fruit. DM is a critical quality indicator, influencing nutritional value, shelf life, and overall consumer acceptance. To improve model robustness, we used a dataset that includes four fruit types (apples, kiwis, mangoes, and pears). Variations in fruit morphology, cultivar, and origin introduces substantial spectral variability, posing challenges for traditional chemometric methods. The multi-fruit spectra were acquired with the same spectrometer model (but different devices) and pre-processed using the second derivative to enhance feature extraction. By training a deep learning model on a multi-fruit dataset we aimed at improving its generalization capabilities.

Our joint NAS and HPO process uses Bayesian optimization via the Tree-structured Parzen Estimator (TPE) [2] where different CNN configurations are evaluated for DM prediction. Models were trained under different data distribution assumptions, initialization strategies and from shallow networks, with a single convolutional layer and one dense layer, to more complex configurations featuring multiple convolutional and dense layers. The HPO included the number and size of convolutional filters, dropout rates, L2 regularization strengths and number of units in dense layers using a cross-validation strategy and aiming at minimizing RMSE. Furthermore, the integration of dual-task models that performed both regression (for DM prediction) and classification (for fruit type identification) allowed the CNNs to leverage shared information between tasks. This allowed to train CNN global models (trained on all fruit types) that performed better than individual PLS models (single fruit models) and global and Locally Weighted PLS models [3].

A second study [4] addressed one of the main criticisms deep learning models face in chemometrics, i.e. their “black box” nature. We show that the use of model-agnostic explainability techniques such as Regression Coefficients [5], LIME [6], and SHAP [7] can be used to reliably identify the spectral bands that most strongly influenced the CNNs’ predictions. Our analysis demonstrated that the key wavelengths identified by the CNNs coincided with those recognized by classical methods like PLS VIP scores and theoretical expectations. Moreover, we show that the CNN leans towards using the domain invariant features across fruit samples (akin di-CovSel [8]). An in-depth analysis of the learned convolution kernels also helps us understand that CNNs perform a type of data driven preprocessing for different samples in the dataset.

We are currently extending the lessons learned here to other datasets to further validate our methodology.

Type of presentation

Contributed Talk

Primary author: PASSOS, Dário (University of Algarve)

Presenter: PASSOS, Dário (University of Algarve)

Session Classification: Session

Track Classification: Spring Meeting