Transforming Hyperspectral Images Into Chemical Maps: A Novel End-to-End Deep Learning Approach

By Ole-Christian Galbo Engstrøm

Data Analytics in food and agricultural science

Current approaches to chemical map generation from hyperspectral images are based on models such as partial least squares (PLS) regression, generating pixel-wise predictions that do not consider spatial context and suffer from a high degree of noise. This study proposes an end-to-end deep learning approach using a modified version of U-Net and a custom loss function to directly obtain chemical maps from hyperspectral images, skipping all intermediate steps required for traditional pixel-wise analysis. This study compares the U-Net with the traditional PLS regression on a real dataset of pork belly samples with associated mean fat reference values. The dataset contains 122 pork bellies for cross-validation and an additional 60 for testing.

The U-Net is trained to generate chemical maps by minimizing a custom multi-objective loss function including (1) an MSE term for the mean of the pixel-wise fat predictions, (2) a term for the squared spatial gradient magnitude to encourage smoothly varying chemical maps, (3) a term for pixel-wise deviations from the 0-100% range, and (4) an L2-regularization term to combat potential overfitting.

The U-Net obtains a lower test set root mean squared error than that of PLS regression on the task of mean fat prediction. At the same time, U-Net generates fine detail chemical maps where 99.91% of the variance is spatially correlated. Conversely, only 2.37% of the variance in the PLS-generated chemical maps is spatially correlated, indicating that each pixel-wise prediction is largely independent of neighboring pixels. Additionally, while the PLS-generated chemical maps contain predictions far beyond the physically possible range of 0-100%, U-Net learns to stay inside this range.

While the accuracy of pixel-wise predictions can not be individually evaluated, each pork belly is divided into 15 sections for which finger pressure reference values are obtained. Finger pressure is known to correlate negatively with fat content. The negative correlation is retained in the predictions of both U-Net (Pearson rho = -0.91) and PLS (Pearson rho = -0.96). This indicates that both models make accurate predictions when averaging pixel-wise predictions over larger areas. For U-Net in particular, this emerging behavior indicates that it has accurately

learned to distinguish between low and high fat areas in the images of pork bellies despite only being trained to estimate the mean over the entire belly.

Thus, the findings of this study indicate that U-Net is preferred over PLS for chemical map generation.