Master internship project CWI's Intelligent and Autonomous Systems Group and Condi Food

Key words: Neural network, hyperspectral inspection systems, quality control,
Condifood supervisor:
Jacques van Munster van Heuven j.vanmunstervanheuven@condifood.com
CWI supervisor:
Eric Pauwels (CWI and VU) eric.pauwels@cwi.nl

Goal
Unpacking the neural network black box: making the Condi food hyperspectral inspection systems models robust.

Background
Condi Food (condifood.com) offers hyperspectral imaging (HSI) solutions for the quality control and/or safety inspection of food products. HSI is a non-destructive technique that combines spectroscopy with imaging ability. Thus, providing both spectral (chemical) and spatial information of samples. The spectral data recorded with this technique is combined with machine learning algorithms to determine the chemical composition or detect anomalies in the imaged food product. At Condi Food we develop and integrate complete and custom HSI systems for our clients using state of the art components. Therefore, we combine our knowledge of physics, chemistry, machine learning, engineering and programming to develop industrial solutions. Some examples of successful implementation by Condi Food are: (1) fruit sorting based on taste, (2) contaminants or foreign object detection in food and (3) freshness of food products.

Project
Condi Food has developed a number of systems that are in use in food processing plants. The systems are made with high-tech components and make use of various machine learning models to detect small quality features. The model is intended to detect a biological feature in a very large dataset that has a very skewed distribution. The occurrence of the feature is around 0.2% and the requirements for accuracy are high. Large in this respect is 100.000 objects per production day. These objects have at least 20 million relevant pixels with each 224 dimensions. This produces over 5 billion datapoints per day. As the dataset has a biological background, making a model that is robust is a challenge in itself.

The goal of the internship (under joint supervision of Condi Food and CWI) is making a contribution to the robustness of the hyperspectral inspection system, in particular the model, by unpacking the neural network black box. In the search for robustness you may be looking at particular events such as model drift or unraveling the underlying spectra that relate to biological aspects. The intern will learn about data analysis, preprocessing data, construction of models, hyperspectral imaging and hyperspectral inspection systems.

The challenges in the model construction are: unveiling the underlying decision making of the neural network and making the model robust. The model should be able to produce the same outcome irrespective of biological variety and challenging and changing environmental circumstances.

What you will learn
You will gain experience in building robust models for large skewed datasets with small features, hyperspectral imaging and working in a team. Furthermore, you will gain insight in world of a start-up and how to bridge the gap between theory and actual industrial applications.

We are looking for
A motivated and ambitious Master level student in Computer Science, Data Science, Machine Learning or Artificial Intelligence. A student with a pragmatic attitude who can apply his/her knowledge and expertise to solve real-world problems. Due to the nature of the project we expect you to have good theoretical knowledge and some practical experience with various machine learning methodologies for classification and signal processing. Academic supervision will be provided by CWI (www.cwi.nl).

Practical information
The research project can start on short notice and is for a minimum of 4 months. Condi food is located in Sassenheim, 20 minutes by bicycle from the Beta Campus and Central train station Leiden, 10 minutes on foot from train station Sassenheim and shares it office with cosine measurement systems (www.cosine.nl).

Sassenheim, oktober 2022