# An overview of implementing Multispectral Imaging coupled with machine learning for the assessment of microbiological quality and authenticity in foods

Anastasia Lytou[1], Lemonia-Christina Fengou[1], Nette Schultz[2], Jens Michael Carstensen[2], Yimin Zhang[3], Fady Mohareb[4], and George-John Nychas[1]

[1] Laboratory of Microbiology and Biotechnology of Foods, Department of Food Science and Human Nutrition, Agricultural University of Athens, Athens, Greece
[2] Videometer
Herlev, Denmark
[3] Lab of Beef Processing and Quality Control, College of Food Science and Engineering, Shandong Agricultural University
Tai'an, Shandong, 271018, P.R. China
[4] Cranfield University
Cranfield, UK

**Abstract** Multispectral Imaging is an increasingly applied technique for the estimation of several quality parameters across the food chain. The microbiological quality and safety as well as the detection of food fraud are among the most significant aspects in food quality and safety assessment. MSI analysis was performed using a VideometerLab instrument (Videometer A/S, Videometer, Herlev, Denmark), while more than 9000 food samples were examined in total, for the assessment of microbiological quality and the detection of food fraud. For estimating microbial populations, total aerobic counts (TAC) were determined. Several regression and classification algorithms were employed, including partial least squares regression (PLS-R), support vector machines (SVM), partial least squares discriminant analysis (PLS-DA), tree-based algorithms etc. The slope of the regression line, root mean squared error (RMSE), coefficient of determination (R-squared) and accuracy score were used as metrics

for the evaluation of models' performance. In adulteration case, the prediction of different levels of pork in chicken meat and vice versa yielded high accuracy scores i.e., over 90% , while, using the SVM algorithm, the presence of bovine offal in beef was successfully detected. Additionally, Random Forest algorithm was efficient (accuracy>93% ) in discriminating seabass and seabream fish fillets. Concerning microbiological quality, as indicated by the performance indices, the developed models exhibited satisfactory performance in predicting microbial load in different foods (RMSE<1.00, R-squared>0.80). Indicatively, MSI spectral data combined with PLS-R could satisfactorily predict TAC and *Pseudomonas* spp. counts on the surface of chicken fillets regardless of storage temperature and batch variation based on the performance metrics (R-squared: 0.89, RMSE: 0.88) while, this algorithm presented also satisfactory performance in estimation microbial populations in brown edible seaweed (R-squared: 0.80, RMSE: 0.90). However, in this case, selecting the appropriate analytical approaches and machine learning algorithms is still challenging.

**Keywords** Multispectral Imaging, Food Quality, Machine Learning, Food Fraud

## 1 Introduction

The interest in using optical technologies that are capable of real-time quality, safety and authenticity assessment has been continuously increasing [1]. Food industry, apart from stabilizing the products to avoid food losses and food waste, should also focus to the development of rapid analytical technologies for the estimation of the microbiological quality and freshness. The last few decades there has been a huge effort from stakeholders to investigate alternative methods that are suitable for online, real-time food quality/safety assessment [2]. In recent years, rapid development of non-invasive sensing technologies for food quality contributed to significant transformations in the supply chain [3]. The data acquired from sensors do not indicate anything without processing and conversion into useful information using pattern recognition or prediction models. Towards this direction, machine learning

algorithms such as, Partial least squares regression (PLS-R), Linear discriminant analysis (LDA), and Quadratic discriminant analysis (QDA) have been reported as reliable tools for the development of predictive models models for quality or adulteration assessment in meat [4], [5]. Moreover, deep learning approaches such as artificial neural networks (ANNs) and support vector machines (SVMs) have been employed, validated, and compared through available online platforms/tools (e.g., sorfML, Metaboanalyst), softwares (e.g., The Unscrambler) or programming languages (R, MatLab, Python), in an attempt to provide accurate predictive models for food spoilage assessment [6], [7]. This work is an overview of studies investigating Multispectral Imaging Analysis, by analyzing various foodstuffs, in an attempt to collect a satisfactory amount of MSI data which in combination with machine learning models can provide significant information about the quality and authenticity of foods.

## 2 Materials and Methods

The whole experimental procedure is briefly shown in Figure 1. The four main steps of the analytical process were 1. Samples' collection, 2. Microbiological analysis, 3. Multispectral Imaging Analysis and 4. Data analysis.

- Samples' collection: Various food samples (9000 samples in total) were collected during the last 5 years. In brief, poultry meat (2300), beef (400), pork (700), fish (1000), pineapple (400), leafy vegetables (500), seaweeds (500), shellfish (500) etc, were subjected to microbiological and MSI analysis, whereas 2000 samples were analysed covering different adulteration scenarios (i.e., chicken vs pork, beef vs offal etc.) In an attempt to increase the diversity of the samples and subsequently the size and the variability of the dataset, apart from the fresh samples, samples that were stored at different temperatures (0, 5, 10, 15°C) for certain time intervals, were also tested. In this way, samples with different microbiological populations and freshness levels were also analysed.

- Microbiological analysis: For the estimation of total aerobic

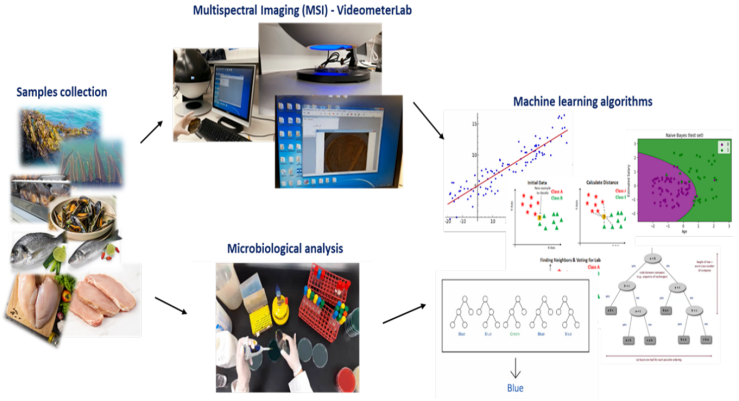


**Figure 1:** Schematic representation of the procedure from samples' collection to data analysis in brief.

counts (TAC), a specific quantity of food sample was transferred aseptically to a stomacher bag, diluted ten times using sterile maximum recovery diluent (MRD) and homogenized in a stomacher (Lab Blender, Seward Medical, London, UK) for 120 s at room temperature. The homogenate was then serially diluted in testing tubes and 0.1 mL of the appropriate dilution was spread in duplicate on the respective culture medium depending on the microbial group. After incubation, colonies were enumerated and their counts were logarithmically transformed (log CFU/g).

- Multispectral Imaging Analysis (MSI): Multi-spectral images (MSI) were captured using a Videometer-Lab instrument (Videometer A/S, Herlev, Denmark) that acquires images in 18 different non-uniformly distributed wavelengths from UV (405 nm) to short wave NIR (970 nm), namely, 405, 435, 450, 470, 505, 525, 570, 590, 630, 645, 660, 700, 850, 870, 890, 910, 940, and 970 nm. LED-based spectral imaging as illustrated in Figure 2 is a fast, non-destructive, and versatile technology for providing high contrast food chemical maps when combined with machine learning methodology. LEDs covering UV, Visual, and NIR wavelengths are sequentially strobed into an integrating sphere with a

superwhite coating. The food sample is placed in the opening of the lower half sphere and receives a very homogenous and diffuse illumination. The built-in calibration and exposure control ensures optimal dynamic range, reproducibility, and traceability.

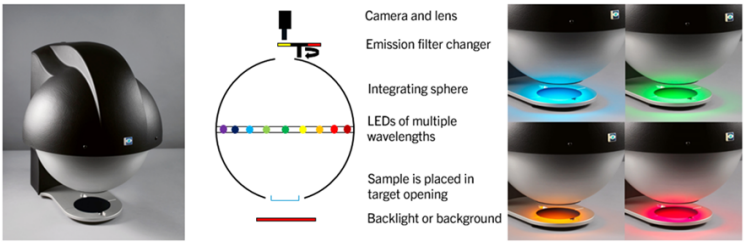


**Figure 2:** VideometerLab instrument used for spectral imaging of food systems. LED strobes of UV-Vis-NIR wavelengths are used to generate a spectral image. Reflectance and fluorescence modes may be combined in the same imaging sequence.

The spectral image, as illustrated in Figure 3, provides information about a rich set of important food compounds like plant and microbial metabolites, pigments, moisture, and lipids. Further it offers a way to measure or remove effects from physical food properties like scattering, specularity, translucency, and heterogeneity.

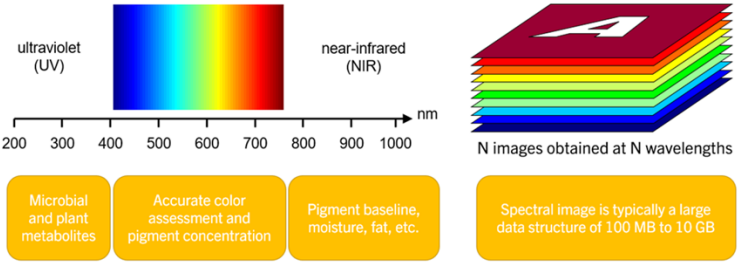


**Figure 3:** LED band-sequential imaging for MSI results in a spectral cube data structure that maps many food-related compounds.

- Data analysis: Various algorithms were employed in the analysis of the MSI data, including Partial Least Squares Regression (PLS-R), Support Vector Regression (SVM-R), tree-based algorithms (Random Forests Regression (RF-R) and Extra Trees) k-Nearest Neighbours' Regression (kNN-R), Linear Discrimination (LDA), Quadratic discrimination (QDA) etc. A part of the dataset was used for the training of the model, while an independent, external dataset was used for the validation (testing) of the model. The performance of the developed models was evaluated via the following metrics and indicess: root mean squared error (RMSE), correlation coefficient (r), overall accuracy, precision, and recall.

## 3 Results

Some indicative results of the MSI applications using various foods are presented below.

3.1 Estimation of microbial population in chicken fillets regardless of storage temperature and batch variation: A PLS-R model was developed by Spyrelli et al [8] for the estimation of microbial counts in chicken fillets. The model parameters and performance metrics (slope, R-squared, RMSE), for the estimation of the population of TAC and *Pseudomonas* spp. using MSI spectral data, are presented in Table 1.

**Table 1:** Performance metrics of PLS-R models estimating TAC and *Pseudomonas* spp. population of chicken fillets using MSI data.

| TAC | Number of samples | slope (a) | R-squared | RMSE |
|---|---|---|---|---|
| Calibration | 330 | 0.74 | 0.86 | 0.73 |
| Cross Validation | 330 | 0.73 | 0.84 | 0.78 |
| Prediction | 72 | 0.77 | 0.90 | 0.98 |
| *Pseudomonas* spp. | | | | |
| Calibration | 330 | 0.73 | 0.85 | 0.83 |
| Cross Validation | 330 | 0.71 | 0.83 | 0.88 |
| Prediction | 72 | 0.70 | 0.90 | 1.21 |

For TAC, the RMSE and R-squared values for model calibration and cross validation were 0.73 and 0.78 log CFU/cm$^2$, as well as 0.86 and

0.84, respectively, whereas the respective values for the prediction were 0.99 log CFU/cm$^2$ and 0.90, respectively. The predicted values were mostly observed within the area of ±1.0 log CFU/cm$^2$, which is considered microbiologically acceptable, while an overestimation for low counts (below 4.0 log CFU/cm$^2$) was evident. Concerning the PLS-R model assessing *Pseudomonas* spp. counts, RMSE and R-squared values were 0.83 log CFU/cm$^2$ and 0.85, respectively, for calibration, while for cross validation they were 0.87 and 0.83 log CFU/cm$^2$, respectively. For the prediction of *Pseudomonas* spp. counts, RMSE and R-squared values were estimated at 1.21 and 0.90 log CFU/cm$^2$ respectively.

3.2 Microbiological quality assessment of seaweed obtained from different geographical areas and harvest years: The prediction model development and validation for the MSI of *A. esculenta* from MI and SAMS samples, from different harvest years are presented below (Table 2), while the findings of this study have been extensively described in [9]. The performance of the model developed in separate for the samples from the different geographical areas was not satisfactory.

**Table 2:** Linear regression fit parameters between actual and predicted TAC values for the different datasets (*A. esculenta* MI, SAMS, MI+SAMS) acquired from MSI analysis.

| MI | slope (a) | R-squared | RMSE |
|---|---|---|---|
| Cross Validation | 0.67 | 0.67 | 0.96 |
| Prediction | 0.49 | 0.51 | 0.95 |
| SAMS | | | |
| Cross Validation | 0.79 | 0.79 | 1.18 |
| Prediction | 0.56 | 0.40 | 1.83 |
| MI+SAMS | | | |
| Cross Validation | 0.92 | 0.92 | 0.81 |
| Prediction | 0.84 | 0.81 | 1.04 |

Extended spectral differences have been observed among the years of harvesting suggesting that maybe the MSI is not suitable for efficient microbial population estimation due to the dependence of this method from the "colour" of the samples that can be misleading for the prediction model. In the case that data from SAMS and MI were

combined, performance statistics values were improved compared to those models developed for each origin in separate (R-squared: 0.80, RMSE: 1.04). Probably by enlarging the size of data, the model was trained/learned better (good performance statistics in cross validation) and the differences in products among the differences in years were more successfully incorporated into the model, while the significance of the visual features (colour related) was degraded.

3.3 Discrimination of fish fillet samples based on different fish species: Several machine learning algorithms were tested for their ability to classify fish fillets to the correct fish species. All the tested models yielded high accuracy scores (>90 % classified to the correct group) for images captured both from the skin and from the flesh side of the fillet (Table 3). Models developed using data from images captured from the skin side, exhibited even better performance (accuracy $>$ 96 % ).

**Table 3:** Accuracy scores ( % ) for the discrimination of fish fillets based on species (i.e., seabass, seabream) using different algorithms.

| Accuracy % | SVM | Extra trees | Random Forest |
|---|---|---|---|
| Skin | 98.39 | 97.85 | 96.77 |
| Flesh | 95.65 | 93.48 | 94.57 |

3.4 Detection of meat adulteration: In Table 4 the performance metrics for the external validation and the classification in five classes for the MSI data is presented. The developed models yielded high performances especially for the classes containing higher proportions of chicken (classes 0 and 25% ).

The classification models of SVMs for the detection of the adulteration of beef with bovine offal (bovine hearts) showed higher or equal performance in terms of accuracy scores for the respective cases compared with the pork-chicken adulteration scenario. The overall correct classification (accuracy) for the case of pork in chicken and offal in beef was 90 % and 100.00 % , respectively. These findings are part of results published before [10].

**Table 4:** Linear regression fit parameters between actual and predicted TAC values for the different datasets (*A. esculenta* MI, SAMS, MI+SAMS) acquired from MSI analysis.

| | | True class | | | |
|---|---|---|---|---|---|
| Pork in chicken | 0 % | 25 % | 50 % | 75 % | 100 % |
| Recall (% ) | 100 | 100 | 100 | 100 | 50 |
| Precision (% ) | 100 | 100 | 100 | 66.67 | 100 |
| Offal in beef | 0 % | 25 % | 50 % | 75 % | 100 % |
| Recall (% ) | 100 | 100 | 100 | 100 | 100 |
| Precision (% ) | 100 | 100 | 100 | 100 | 100 |

## 4 Conclusion

MSI data coupled with machine learning algorithms exhibit potential towards efficient detection of adulteration and microbial counts estimation and could be a rapid and non-invasive tool for the quality assessment in various foodstuffs.

This work has been funded by the project DiTECT (861915).

## References

1. H. Pu, K. M., and D. Sun, "Selection of feature wavelengths for developing multispectral imaging systems for quality, safety and authenticity of muscle foods-a review," in *Trends in Food Science and Technology*, 2015.
2. P. Tsakanikas, A. Karnavas, E. Panagou, and G. Nychas, "A machine learning workflow for raw food spectroscopic classification in a future industry," in *Scientific Reports*, 2020.
3. N. J. Watson, A. Bowler, A. Rady, O. Fisher, A. Simeone, J. Escrig, E. Woolley, and A. Adedeji, "Intelligent sensors for sustainable food and drink manufacturing," in *Frontiers in Sustainable Food Systems*, 2021.
4. A. Ropodi, D. Pavlidis, F. Mohareb, E. Panagou, and G. Nychas, "Multispectral image analysis approach to detect adulteration of beef and pork in raw meats," in *Frontiers in Sustainable Food Systems*, 2017.
5. Y. Kumar and S. Karne, "Spectral analysis: A rapid tool for species detection in meat products," in *Trends in Food Science and Technology*, 2019.

6. A. M. Jiménez-Carvelo, A. González-Casado, M. G. Bagur-González, and L. Cuadros-Rodríguez, "'alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity–a review," in *Food research international*, 2019.
7. S. Jaafreh, O. Valler, J. Kreyenschmidt, K. Günther, and P. Kaul, "In vitro discrimination and classification of microbial flora of poultry using two dispersive raman spectrometers (microscope and portable fiber-optic systems) in tandem with chemometric analysis," in *Talanta*, 2019.
8. E. Spyrelli, C. Papachristou, G.-J. Nychas, and E. Panagou, "Microbiological quality assessment of chicken thigh fillets using spectroscopic sensors and multivariate data analysis," in *Foods*, 2021.
9. A. Lytou, P. Tsakanikas, D. Lymperi, and G.-J. Nychas, "Rapid assessment of microbial quality in edible seaweeds using sensor techniques based on spectroscopy, imaging analysis and sensors mimicking human senses," in *Sensors*, 2022.
10. L.-C. Fengou, A. Lianou, P. Tsakanikas, F. Mohareb, and G.-J. Nychas, "Detection of meat adulteration using spectroscopy-based sensors," in *Foods*, 2021.