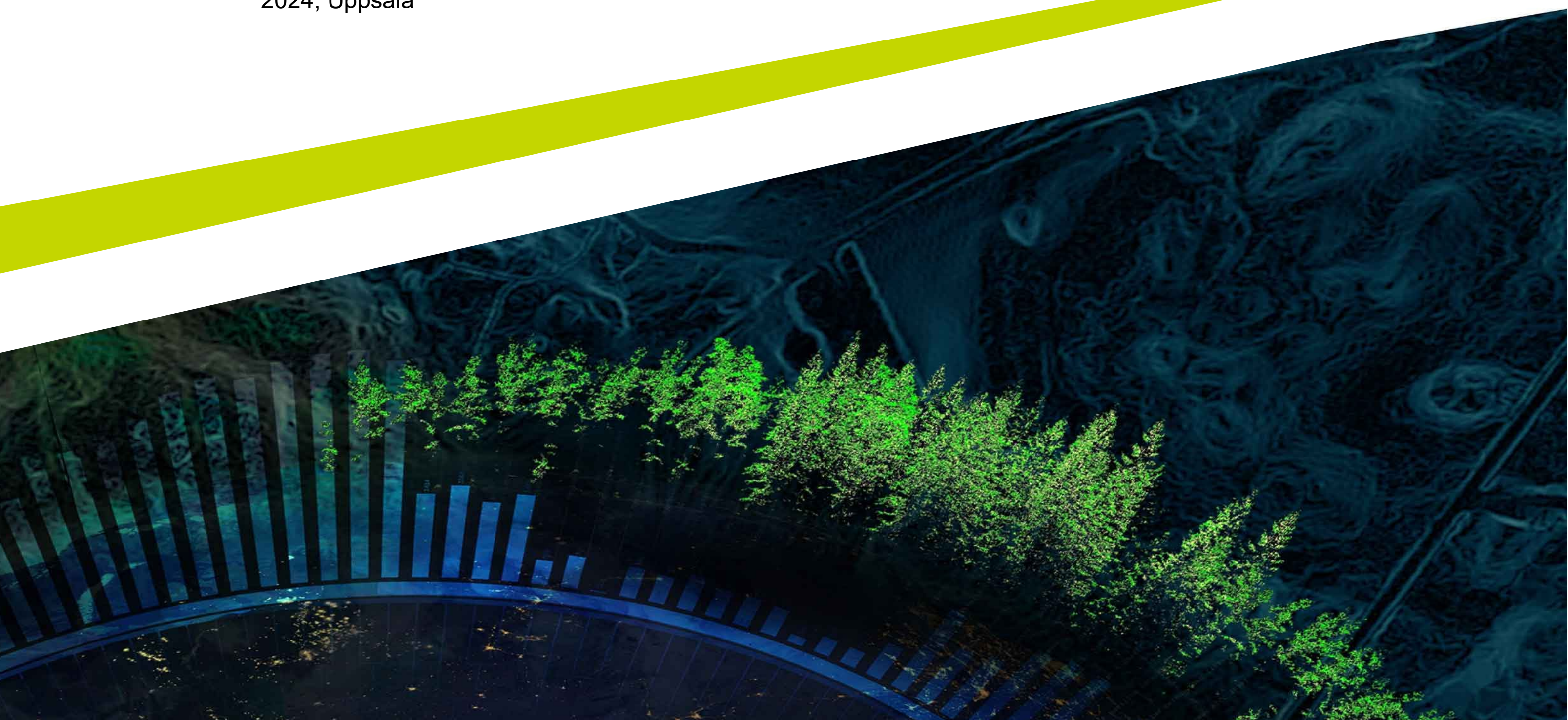
Cheese quality authentication through infrared- and Raman spectroscopy for industrial use

Lovisa Hambäck

Independent project • 15 credits
Swedish University of Agricultural Sciences, SLU
Department of Molecular sciences
Bachelor's degree in food science
Molecular Sciences, 2024:22
2024, Uppsala



Cheese quality authentication through infrared- and Raman spectroscopy for industrial use

Lovisa Hambäck

Supervisor:	**Åse Lundh, Swedish University of Agricultural Sciences, Department of molecular sciences**
Examiner:	Monika Johansson, Swedish University of Agricultural Sciences, Department of Molecular Sciences

Credits:	15
Level:	G2E
Course title:	Självständigt arbete I livsmedelsvetenskap
Course code:	EX0876
Programme/education:	Bachelor's degree in food science
Course coordinating dept:	Department of molecular sciences
Place of publication:	Uppsala
Year of publication:	2024
Title of series:	Molecular Sciences
Part number:	2024:22
Copyright:	All featured images are used with permission from the copyright owner.
Keywords:	Cheese quality, Image analysis, Infrared spectroscopy, Raman spectroscopy, Hyperspectral imaging

Swedish University of Agricultural Sciences
Faculty of Natural Resources and Agricultural Sciences
Department of Molecular Sciences

Abstract

Cheese fraud, which include mislabelling and adulteration, is prevalent, particularly in high-value cheeses such as PDO and PGI cheeses. Novel imaging techniques such as infrared (IR) spectroscopy and Raman spectroscopy has emerged as potentially non-destructive, fast, and accurate ways to authenticate cheeses. These techniques, coupled with chemometric analysis, have been used to gain detailed insights into the chemical and physical properties of cheese. They have also been improved to analyse spatial information, with near-infrared hyperspectral imaging (NIR-HSI) and spatial Raman. Despite their potential, widespread industrial adaption is limited.

This essay explores the application of IR and Raman spectroscopy, coupled with spatial techniques, in cheese authentication. It evaluates their advantages, like being able to analyse samples through their packaging, as well as its disadvantages, such as expensive instruments and complicated data analysis. Necessary steps towards implementing the techniques in industry is integrating advanced analytic tools, such as effective machine learning algorithms, and creating a large public dataset to improve the accuracy and applicability of these imaging techniques for industrial usage.

Keywords: Cheese quality, Image analysis, Infrared spectroscopy, Raman spectroscopy, Hyperspectral imaging

Table of contents

Abbreviations

ANN	Artificial neural network
HSI	Hyperspectral imaging
IR	Infrared
NIR	Near-infrared
MIR	Mid-infrared
PR	Parmigiano Reggiano
PDO	Protected Designation of Origin
PGI	Protected Geographical Indication

1. Introduction

During the last 30 years, dairy production has increased worldwide by over 59% and as more cheese is being bought, the rate of cheese food fraud has followed (FAO n.d.). The need for cheese quality authentication has therefore increased, both for retailers and regulating agencies, to authenticate that the cheeses comply with legislation and specifications. Cheese is the most common dairy product in food fraud, with fraudulent documentation along with adulteration accounting for 94% of cheese frauds (Montgomery et al. 2020). This can include having a lower protein or fat content than the labelling says, substitution of milk species, addition of illegal substances, or higher salt content. The most commonly adulterated cheeses are mozzarella, ricotta, hard cheese, feta, butter cheese and Parmigiano Reggiano (Abedini et al. 2023). Many of these are protected by a certification such as PDO or PGI and are thus subject to additional regulation such as geographical region of the manufacturer, region of the feed and processing steps, making them more expensive than regular cheese.

Traditionally, cheese authenticity has been evaluated by physicochemical methods such as GC, HPLC, PCR etc. The problem with these techniques is that they are time-consuming, chemical-dependent, they require trained personnel, and they produce waste through being destructive techniques. Novel imaging techniques which are non-destructive, fast with no preparation have therefore gained interest for authenticating cheese quality during the 21st century (Lei & Sun 2019). Imaging techniques coupled with chemometric analysis utilize different kinds of wavelengths to analyse physical and chemical properties have been successfully applied to many different sectors, from pharmaceuticals to agri-food products. Concerning cheese quality evaluation, the most common wavelength regions are the infrared (IR) region, the visible region, and the X-ray region. Raman spectroscopy, not relating to a certain wavelength region, has also been used. All these techniques only acquire spectral information and therefore they have been developed to measure spatial information through hyperspectral imaging (HSI) or other technical additions to the light sources.

The main challenges of all the above mentioned imaging techniques is that they are expensive, require extensive training and calibration as well as complicated data analysis (Hebling e Tavares et al. 2022). Therefore, none of the techniques have yet been fully implemented for industrial use in the food sector, although they have been in use in other sectors, mainly the medical sector, for a longer period.

This essay will therefore explore industrial application of cheese authentication through imaging techniques. Since the X-ray region is used mostly for evaluating physical properties such as cheese eye development or microstructure, IR and Raman spectroscopy as well as their spatial counterparts are the most common for cheese authentication and will thus be the focus. Through the literature provided by peer-reviewed journals available online this essay will start with the description of the techniques, followed by different applications and finally analysis of future challenges and possibilities.

2. Spectroscopic imaging methods

Spectroscopic methods are when absorbance of photons is measured to generate an absorbance spectrum specific to the sample. The photons can be described by their wavelength, which is usually measured in nanometres (nm), or cm^{-1}, and their wavelengths are associated to different regions of the electromagnetic spectrum (Figure 1). The different wavelengths will interact differently with the material as certain spectral regions are associated with certain types of interaction. Shorter wavelengths will interact with smaller parts of the material, such as electrons, while broader wavelengths will interact with larger material, such as chemical bonds (Dufour 2009).

An image can therefore be described by spatial and spectral resolution where spatial resolution is the measure of the smallest object distinguishable in the image and spectral resolution is the selection of wavelengths measured (Khan et al. 2018). The spatial information is therefore related to the sensors of the instruments while the spectral information is related to the choice of wavelengths.

In order to properly describe the spectral and spatial information, data preprocessing followed by data analysis is required. These steps can greatly affect the result of the prediction or classification accuracy and therefore they are briefly described at the end of this chapter.

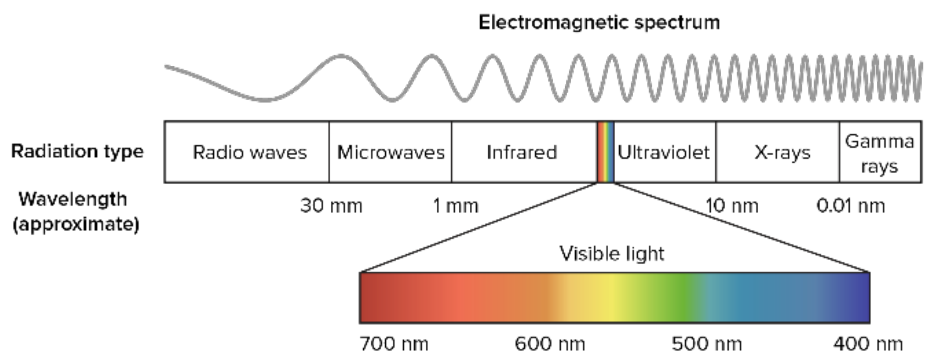


Figure 1. The electromagnetic spectrum. Image acquired from https://www.khanacademy.org/science/biology/photosynthesis-in-plants/the-light-dependent-reactions-of-photosynthesis/a/light-and-photosynthetic-pigments

2.1 Spectroscopy techniques

2.1.1 Infrared spectroscopy

The spectral bands from infrared waves are derived from vibrations of chemical bonds which is why infrared spectroscopy is sometimes referred to as vibrational spectroscopy (Dufour 2009). The infrared region is divided into three smaller regions: near-infrared (NIR) (800-2500 nm), mid-infrared (MIR) (2500-25000 nm) and far IR (25000-100000 nm). NIR is the most extensively researched region for food quality analysis. It has traditionally been used to measure chemical composition (Dufour 2009) with different wavelengths being associated with different chemical composition, as illustrated in figure 2, but it can also be used for texture and sensory analysis (Woodcock et al. 2008). Its disadvantage is a weak sensitivity to components present at low concentrations (Woodcock et al. 2008). MIR has also been employed for various applications and has its main implementation in identification of organic and organometallic molecules (Woodcock et al. 2008).

Both NIR and MIR can be improved through Fourier transform-NIR and -MIR (FTIR and FT-MIR) which has the advantages of being faster and more sensitive, but is consequently more expensive (Thain 2022).

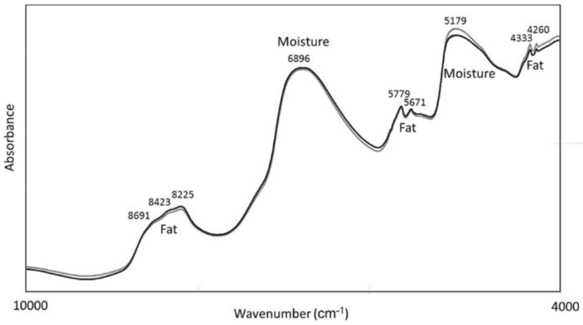


Figure 2. An example of a NIR spectrum collected from cheese. Image acquired from Ayvaz et al. (2021)

2.1.2 Raman spectroscopy

Raman spectroscopy is not related to a certain wavelength region, but is a technique where the sample is exposed to a laser and the information is gathered from inelastically scattered photons (Jin et al. 2016) (Figure 3). Specific chemical bonds generate specific "Raman shifts", and Raman spectroscopy can thus be used to measure chemical properties. Its advantages include speed and the ability to be used for water-rich samples, as the information from the water molecules do not overlap with the information of other molecules (Jin et al. 2016). One of the major drawbacks of Raman spectroscopy is a band overlap between fat and protein components, making quantification of these compounds difficult but certain wavelengths can be selected to reduce the overlapping information (Smith et al. 2017).

2.2 Spatial imaging

2.2.1 Infrared-hyperspectral imaging

Hyperspectral imaging (HSI) is often described as a major upcoming research area of interest for various applications, among them food quality analysis. HSI combines spatial and spectral information by relating several optical bands in broad spectral range to each pixel (Khan et al. 2018) (Figure 3). The huge advantage of HSI compared to regular spectroscopy is its ability to acquire spatial information and thus being able to examine heterogenous samples (Khan et al. 2018). However, the multidimensionality of HSI images creates the problem of handling large amounts of data, together with low-speed and high-cost calculations (Siche et al. 2016).

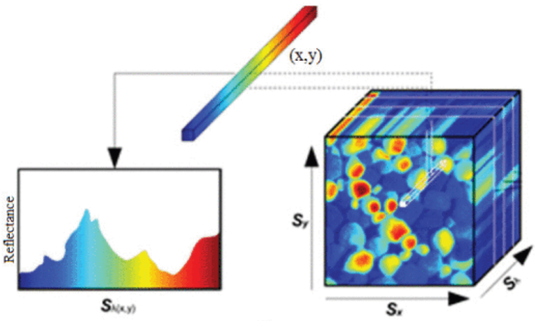


Figure 3. Hyperspectral image data, where the spatial information is represented by the "pixels" or S_x and S_y, and the spectral information by S_λ. Image adapted from Khan et al. (2018)

2.2.2 Raman spectroscopy and spatially offset Raman spectroscopy

To capture spatial information through Raman spectroscopy, two techniques called confocal Raman spectroscopy and spatially offset Raman spectroscopy (SORS) are used. These are both relatively new techniques and have not been widely used for food analysis. SORS have the main strength of being able to collect data below the surface, up to several millimetres, but is a relatively new technique with few applications (Conti et al. 2016; Parker 2017).

2.3 Data preprocessing

In order to improve the ability of the computer to properly analyse the information, data preprocessing is crucial and can improve the accuracy of cheese product classification from 40% to 98% (Zhang et al. 2023). For NIR, data preprocessing is necessary to adhere the signal data to Lambert-Beerz's law for quantification of the compounds.
Preprocessing techniques are divided into reference-dependent and reference-independent, where reference-independent is the most common for cheese quality analysis since a preexisting reference rarely exists. Common reference-independent techniques are multiplicative signal correction, baseline correction, standard normal variate correction and normalization, as well as differentiation by Savitzky-Golay routine (Rinnan et al. 2009). The choice of technique depends on the type of spectral information as well as on the later classification algorithm. The techniques can also be combined, although a combination rarely leads to significant improvements of accuracy (Rinnan et al. 2009).

2.4 Algorithms and data analysis

After the data preprocessing, the classification or prediction model is created. This step involves making a calibration set, constructing a model, validating the model and finally, predicting unknown samples.

Classification is the process of grouping different objects into different classes based on certain parameters. These classes can be predefined, or created as the classification algorithm learns the data. Models can be defined as linear or nonlinear, with linear models assuming that the signals depend exclusively on the concentration of the desired analyte. Despite this not always being the case, linear models are by far the most used, with the most common ones being principal component regression (PCR) and partial least squares regression (PLSR).

One common nonlinear model is artificial neural networks (ANNs) which is a type of machine learning that uses nodes or "neurons" in hidden layers to learn by iteration. The amount of neurons in the hidden layers should be optimised in order to reduce the risk of overfitting as well as saving time, since the model can stabilize after a certain amount of neurons (Zhang et al. 2023). ANNs can be modelled in different ways and the interested reader is referred to the review by Zareef et al. (2020).

When applying the data analysis in an industrial setting, there are several other aspects, besides receiving the highest accuracy, that must be considered. For instance, the resulting analytical information should be easily interpretable, since it will be used by less data-skilled operators. The models should be useful for longer time periods and they should be easily updated when deviations emerge, as normally happens (Blanco Romía & Alcalà Bernàrdez 2009). Recalibration can be very expensive, and therefore, transfer algorithms has to be coupled with a robust database to translate data between different machines and techniques (van den Berg & Rinnan 2009).

3. Applications of infrared and Raman spectroscopy for cheese

Cheese is a heterogenous food product mainly consisting of water, fat, and protein, with minor quantities of carbohydrates and minerals. The product is the result of a combination of raw material, processing parameters, and maturation. Therefore, different kinds of cheeses can vary substantially in their physical and chemical composition, allowing certain properties to be attributed to certain cheeses and classifying them accordingly.

This chapter explores the use of IR and Raman spectroscopy to quantify key chemical components and to classify cheeses based on their imaging fingerprint.

3.1 Measuring chemical composition

Food fraud in cheese often involves mislabelling concerning fat and protein content (Montgomery et al. 2020). Additionally, moisture content is a critical parameter for regulatory compliance in certain cheeses. Thus, measuring these three attributes is key in monitoring compliance. Since the chemical composition changes during maturation, measuring macromolecular composition can also be used for monitoring the maturation process (Malegori et al. 2021).

NIR is the most extensively used method for measurement of macromolecular composition, with root mean standard error of validation ($RMSE_{VAL}$)-values typically ranging from 1.13 to 1.39 for fat, protein and moisture, although values can vary between 0.07 and 4.79 (Bittante et al. 2022). Raman spectroscopy has been far less utilized for cheese, and the results have been similar to NIR (Stocco et al. 2024), although the study by Stocco et al. (2024) had poor NIR-results compared to other studies.

Since the chemical composition can vary throughout the cheese, spatial information can enhance predictions by producing chemical maps of cheeses. The accuracy difference between NIR and NIR-HSI is small with the difference of RMSE being -0.04 to 0.35 (da Silva Medeiros et al. 2024). Spatially offset Raman spectroscopy has also successfully predicted fat and protein content with a mean absolute error of 1.29 and 0.93 g/g cheese, respectively (Arroyo-Cerezo et al. 2023).

3.2 Authentication

Cheese authentication involves classifying cheeses based on geographical origin, milk species, farming practices etc., as well as identifying adulteration. A prerequisite is that the sampled cheeses vary enough in chemical and/or physical properties since that is the basis for the spectroscopic analysis. However, the great potential of spectroscopic analysis is the classification based on multiple parameters, without the need for specific biomarkers.

3.2.1 Determination of origin

Some cheeses have been granted Protected Designation of Origin (PDO) meaning that they must be produced, processed, and developed in a specific geographical area and they are therefore usually premium products. Spectroscopic techniques have been used to classify several different PDO cheeses, among them Parmigiano Reggiano (PR), Gruyére and Prato (Li Vigni et al. 2020; Tolentino et al. 2023; Wu et al. 2023; Stocco et al. 2024).

A comparison between NIR and Raman for PR identification showed that Raman was slightly better at identifying the cheese type (Stocco et al. 2024). Fusion of NIR and Raman data improved accuracy in some cases, though inconsistently, likely due to a small dataset affecting the classification model (Stocco et al. 2024). For Prato cheese compared with the similar, but shorter matured, mozzarella cheese, MIR separated the cheeses with 92% accuracy (Tolentino et al. 2023).

It is interesting to note that the study conducted by Tolentino et al. (2023) also had 16 of 20 labelled Prato cheeses from the supermarket classified as mozzarella by the model, showing that samples from the supermarket may not be the most reliable choice for building a model. The studies conducted on PR however, used PR provided specifically by the PR consortium (Li Vigni et al. 2020; Stocco et al. 2024).

3.2.2 Identification of species origin

Another common food fraud related to cheese production is the substitution of milk from one species with another species, often due to cheaper milk. Species affect the chemical composition, particularly the fatty acid profile, and should thus be suitable for classification (Paszczyk & Łuczyńska 2020). Few studies have focused on cheese samples, but Halloumi, a Cypriote goat and sheep milk cheese, has been successfully classified according to species origin with 97% accuracy using FTIR (Tarapoulouzi et al. 2020). Later, NIR-HSI was also shown to be able to predict species origin in grated halloumi, although with a less accurate result of 76% (Tarapoulouzi et al. 2024). Spatially offset Raman spectroscopy has also been used for successful differentiation of cheeses with different animal species origin at an accuracy of 84-86% (Ostovar pour et al. 2021; Arroyo-Cerezo et al. 2023).

3.2.3 Prediction of feeding and farming system

Certain PDO/PGI cheeses have regulations on feed types, such as silage or grazing. Feeding system significantly alter the fat and protein profile of cheese, impacting yield and cheese flavour (Soryal et al. 2004).

NIR spectroscopy has shown over 90% high accuracy in predicting feeding regimes in milk samples, highly outranking the use of visible light and sensorial description (Coppa et al. 2012; Andueza et al. 2013; Bergamaschi et al. 2020). MIR has also been used successfully to authenticate feeding restrictions in PDO specification for milk with a success rate of 80-90% depending on the specification of the feeding regime (Coppa et al. 2021). The accuracy is higher for milk samples than for cheese, suggesting that the techniques may find better use for PDO producers wanting to monitor their raw milk providers (Bergamaschi et al. 2020). The samples which are more difficult to predict are those with lower grass quality and/or higher maize silage quantity (Andueza et al. 2013).

The validation of organic ingredients can also be important for the consumer since organic cheeses tend to be more expensive. When predicting organic and non-organic cheeses, NIR has achieved a 84% accuracy, which is quite well but far below traditional GC methods (Liu et al. 2018). Liu et al. (2018) hypothesized that this can be due to the nature of the chemical markers of organic milk. The overlap of different feeding regimes in organic and conventional farming practices makes it difficult to find specific differential chemical markers (Capuano et al. 2013). Studies have also found it difficult to differentiate between lowland and upland farms due to the large effect of feeding regimes (Coppa et al. 2012).

3.2.4 Validation of maturation time

Maturation length impacts flavour and aroma of cheese (Weimer 2007), but is an expensive process due to storage and monitoring. Therefore, validating maturation time is important to ensure that consumers are paying more to receive a better product.

NIR predicted the age of cheddar cheese with an R^2 of 0.94 (Downey et al. 2005), while NIR-HSI predicted the age of a traditional Swedish hard cheese with an R^2 of 0.76 (Priyashantha et al. 2020). NIR also outperformed Raman in determining PR cheese maturation (Stocco et al. 2024). NIR-HSI has also been used to monitor

Formagetta cheese ripening by visualising the dehydration, proteolysis and lipolysis occurring during ripening (Malegori et al. 2021).

3.2.5 Adulteration

Common cheese adulterations include adding vegetable oil, curd, starch, other food additives and not adhering to regulations (Abedini et al. 2023). For example, grated PR cheese must have a rind content below 18% (w/w), moisture content between 25% - 35%, and less than 25% of particles smaller than 0.5 mm (Parmigiano Consortium 2018). Although few cheese wheels in total have more than 18% rind, this limit could be exceeded due to mixing errors or different sedimentation velocities and spectroscopic analysis can be continuously used to asses when the rind content surpasses the limit (Alinovi et al. 2019). Using NIR, rind content has been successfully predicted with a R^2 of 0.875, where a removal of wavelengths from 1335-1933 nm improved the accuracy slightly (Alinovi et al. 2019). To obtain spatial information as well, since grated cheese is a highly heterogenous product, NIR-HSI has been tested and the technique further improved the accuracy to an R^2-value of 0.98, accurately predicting all unknown samples with a rind content exceeding 18% (Calvini et al. 2022). However, Calvini et al. (2022) did not include cheese samples of different maturation age, which may have affected the robustness.

Another type of adulteration associate to grated cheese is the addition of cellulose and other anti-caking agents without declaring it on the package. Other starches can be added as well to increase weight and volume of cheeses (Visconti et al. 2020). A model based on NIR was able to correctly predict samples containing both starches and cellulose with a specificity of 100% and a slightly lower sensitivity (Visconti et al. 2020). Even digital images taken by a digital camera have been able to differentiate between different adulterations of grated cheese and although the precision and sensitivity was much lower (70-90%) it is an interesting method since the cost of equipment is much lower (Visconti et al. 2023).

Adulteration of fat source by substitution with a cheaper alternative than milk fat, e.g. vegetable oil, can be done to lower the economic cost of producing cheeses. Raman spectroscopy has been able to detect margarine, corn oil and palm oil in ultrafiltered white cheese with R^2-values of 0.97-0.99 and a limit of detection around 3.4% (Genis et al. 2021). FTIR has also been successful at predicting soybean oil in butter cheeses with an R^2-value of 0.998 (Leite et al. 2019), whereas NIR acquired an R^2 of 0.94 (Medeiros et al. 2023). These values show that all three techniques can be used to identify fat adulteration in cheese at high accuracy.

3.3 Prediction through packaging

Developing techniques to analyse cheese samples through packaging would allow usage to be much quicker, cheaper, and sustainable, minimizing food waste. Although packaging materials can significantly affect spectroscopy signals, spectral filtering can be applied for accurate result (Ding et al. 2023). However, studies on cheese authentication using IR and Raman spectroscopy for packaged cheese are limited.

NIR has been used on packaged cheese for various applications. It was used to evaluate the shelf-life of Crescenza cheese (Cattaneo et al. 2005), to measure chemical components (Pi et al. 2009) and, coupled with HSI, to detect *Clostridium* contamination (Reis et al. 2023).

Raman spectroscopy has also demonstrated potential, identifying non-authentic PR through plastic packaging with 100% accuracy and predicting rind content in grated cheese samples with a 5% error margin (Li Vigni et al. 2020). Although spatially offset Raman spectroscopy has been used to a lesser extent in research associated to cheese, it shows considerable promise (Ostovar pour et al. 2021).

4. Analysing future challenges

While the results from many articles highlight the potential of spectroscopic techniques for authenticating cheese quality, their adaption in industry remains limited. Reviews on the subject frequently underline certain challenges: lowering the cost of the instruments, development of robust portable devices, development of machine learning tools for data analytics, and the development of large datasets which account for seasonal and regional differences (Woodcock et al. 2008; Lei & Sun 2019; Lu et al. 2020; Pu et al. 2021; Yakubu et al. 2022).

4.1 Challenges

4.1.1 Cost of instruments

The high cost of spectroscopic instruments, especially those adapted from the medical sector, remains a significant barrier even though the price has been decreasing continuously. The total cost can vary depending on the model and features, typically starting at $10,000 and going up to several hundred thousand dollars.

Portable NIR and Raman devices could potentially provide cheaper and more user-friendly alternatives to tabletop versions. As of right now, there are portable NIR/Raman devices which perform with similar accuracy as tabletop versions and is thereby a real possibility for the industry (Liu et al. 2018; Müller-Maatsch & van Ruth 2021; Medeiros et al. 2023; Wu et al. 2023).

4.1.2 Data calibration and standardization

Calibrating different data across devices is a major bottleneck for the widespread use of imaging techniques (Müller-Maatsch & van Ruth 2021). In the pharmaceutical industry, successful attempts at calibration transfer for Raman instruments has been done, both local transfer between two instruments and global transfer, which could potentially pave the way for a similar methodology in the food sector (Steinbach et al. 2017). Deep learning tools are expected to play a vital role in facilitating this process (Müller-Maatsch & van Ruth 2021).

4.1.3 Handling large data sets

The large data sets generated by techniques like HSI, and spatially offset Raman spectroscopy pose challenges in data handling and processing time. One solution is the selection of certain important wavelengths, which has been used for some studies to improve accuracy (Siche et al. 2016; Alinovi et al. 2019). However, the wavelength selection differs between studies, showing that choice of data analysis tools and type of cheese can affect wavelength selection. Therefore, a standardisation of method together with a comparison of wavelength selection for various cheeses is necessary (Kamruzzaman et al. 2016).

4.2 Solutions and future directions

4.2.1 Database construction

While deep learning often has been mentioned as the future of data analysis in this field (Lei & Sun 2019; Abedini et al. 2023), it requires extensive training on high-quality datasets. Constructing a large and robust database to be used by machine learning requires the data to encompass a wide range of factors such as cows, geographical origin, season, age etc. (Hebling e Tavares et al. 2022). Few studies mentioned in this essay contain more than 100 samples, with only one having approximately 1000 samples (Bergamaschi et al. 2020). It would be beneficial to create a public dataset comprising of data from various sources, as has been done by (Zedda et al. 2024) to be able to develop machine learning tools more effectively.

4.2.2 Use of advanced analytics models

Nonlinear models such as ANNs and probabilistic neural networks (PNNs) have shown promise in improving accuracy over traditional methods such as PLSR (Vásquez et al. 2018; Zhang 2020). Often, the classification algorithms are chosen based on the authors knowledge and preference instead of the goal of the analysis which limits the improvement pace (Ballabio & Todeschini 2009). Since machine learning techniques can handle complex data relationships better, researchers in the field of spectroscopic techniques should collaborate with machine learning experts in order to improve their prediction capabilities.

4.2.3 Data fusion

Fusion of data from different spectroscopic techniques, such as NIR and Raman, has been repeatedly examined to improve accuracy. However, fusion of data can introduce excessive noise (Stocco et al. 2024) which can negatively impact the accuracy. To properly benefit from the data-fusion the fused data must be complementary, which is the problem for the fusion of NIR and Raman, since they measure similar molecular interactions.

5. Conclusion

This essay showcases the broad range of applications for IR and Raman spectroscopy in cheese quality authentication. The high accuracy of prediction which has been shown, along with the ability of portable instruments to carry out equally good results, proves that the technique can be valuable for industrial use. However, there needs to be a compilation of larger databases, a standardisation of data preprocessing and classification, as well as more robust calibration algorithms to fully attain to the industry’s standards.

Further research should focus on techniques which are able to capture spatial data as well as below-surface-data, such as NIR-HSI and spatially offset Raman spectroscopy. While these two have been equally good at predicting chemical composition compared to the normal spectroscopic techniques, the spatial instruments have been better at authentication. There should also be a focus on standardising data analysis and incorporating machine learning algorithms. Future collaboration between chemists and data scientists is therefore necessary to further improve this field.

References

Abedini, A., Salimi, M., Mazaheri, Y., Sadighara, P., Alizadeh Sani, M., Assadpour, E. & Jafari, S.M. (2023). Assessment of cheese frauds, and relevant detection methods: A systematic review. *Food Chemistry: X*, 19, 100825. https://doi.org/10.1016/j.fochx.2023.100825

Alinovi, M., Mucchetti, G. & Tidona, F. (2019). Application of NIR spectroscopy and image analysis for the characterisation of grated Parmigiano-Reggiano cheese. *International Dairy Journal*, 92, 50–58. https://doi.org/10.1016/j.idairyj.2019.01.010

Andueza, D., Agabriel, C., Constant, I., Lucas, A. & Martin, B. (2013). Using visible or near infrared spectroscopy (NIRS) on cheese to authenticate cow feeding regimes. *Food Chemistry*, 141 (1), 209–214. https://doi.org/10.1016/j.foodchem.2013.02.086

Arroyo-Cerezo, A., Jiménez-Carvelo, A.M., González-Casado, A., Ruisánchez, I. & Cuadros-Rodríguez, L. (2023). The potential of the spatially offset Raman spectroscopy (SORS) for implementing rapid and non-invasive in-situ authentication methods of plastic-packaged commodity foods – Application to sliced cheeses. *Food Control*, 146, 109522. https://doi.org/10.1016/j.foodcont.2022.109522

Ayvaz, H., Mortas, M., Dogan, M.A., Atan, M., Yildiz Tiryaki, G. & Karagul Yuceer, Y. (2021). Near- and mid-infrared determination of some quality parameters of cheese manufactured from the mixture of different milk species. *Journal of Food Science and Technology*, 58 (10), 3981–3992. https://doi.org/10.1007/s13197-020-04861-0

Ballabio, D. & Todeschini, R. (2009). Chapter 4 - Multivariate Classification for Qualitative Analysis. I: Sun, D.-W. (red.) *Infrared Spectroscopy for Food Quality Analysis and Control*. Academic Press. 83–104. https://doi.org/10.1016/B978-0-12-374136-3.00004-3

van den Berg, F. & Rinnan, Å. (2009). Chapter 5 - Calibration Transfer Methods. I: Sun, D.-W. (red.) *Infrared Spectroscopy for Food Quality Analysis and Control*. Academic Press. 105–118. https://doi.org/10.1016/B978-0-12-374136-3.00005-5

Bergamaschi, M., Cipolat-Gotet, C., Cecchinato, A., Schiavon, S. & Bittante, G. (2020). Chemometric authentication of farming systems of origin of food (milk and ripened cheese) using infrared spectra, fatty acid profiles, flavor fingerprints, and sensory descriptions. *Food Chemistry*, 305, 125480. https://doi.org/10.1016/j.foodchem.2019.125480

Bittante, G., Patel, N., Cecchinato, A. & Berzaghi, P. (2022). *Invited review:* A comprehensive review of visible and near-infrared spectroscopy for predicting the chemical composition of cheese. *Journal of Dairy Science*, 105 (3), 1817–1836. https://doi.org/10.3168/jds.2021-20640

Blanco Romía, M. & Alcalà Bernàrdez, M. (2009). Chapter 3 - Multivariate Calibration for Quantitative Analysis. I: Sun, D.-W. (red.) *Infrared Spectroscopy for Food Quality Analysis and Control*. Academic Press. 51–82. https://doi.org/10.1016/B978-0-12-374136-3.00003-1

Calvini, R., Michelini, S., Pizzamiglio, V., Foca, G. & Ulrici, A. (2022). Evaluation of the effect of factors related to preparation and composition of grated Parmigiano Reggiano cheese using NIR hyperspectral imaging. *Food Control*, 131, 108412. https://doi.org/10.1016/j.foodcont.2021.108412

Capuano, E., Boerrigter-Eenling, R., van der Veer, G. & van Ruth, S.M. (2013). Analytical authentication of organic products: an overview of markers. *Journal of the Science of Food and Agriculture*, 93 (1), 12–28. https://doi.org/10.1002/jsfa.5914

Cattaneo, T.M.P., Giardina, C., Sinelli, N., Riva, M. & Giangiacomo, R. (2005). Application of FT-NIR and FT-IR spectroscopy to study the shelf-life of Crescenza cheese. *International Dairy Journal*, 15 (6), 693–700. https://doi.org/10.1016/j.idairyj.2004.07.026

Conti, C., Realini, M., Colombo, C., Botteon, A. & Matousek, P. (2016). Contrasting confocal with defocusing microscale spatially offset Raman spectroscopy. *Journal of Raman Spectroscopy*, 47 (5), 565–570

Coppa, M., Martin, B., Agabriel, C., Chassaing, C., Sibra, C., Constant, I., Graulet, B. & Andueza, D. (2012). Authentication of cow feeding and geographic origin on milk using visible and near-infrared spectroscopy. *Journal of Dairy Science*, 95 (10), 5544–5551. https://doi.org/10.3168/jds.2011-5272

Coppa, M., Martin, B., Hulin, S., Guillemin, J., Gauzentes, J.V., Pecou, A. & Andueza, D. (2021). Prediction of indicators of cow diet composition and authentication of feeding specifications of Protected Designation of Origin cheese using mid-infrared spectroscopy on milk. *Journal of Dairy Science*, 104 (1), 112–125. https://doi.org/10.3168/jds.2020-18468

Ding, F., Zuo, C., García-Martín, J.F., Ge, Y., Tu, K., Peng, J., Xiao, H., Lan, W. & Pan, L. (2023). Non-invasive prediction of mango quality using near-infrared spectroscopy: Assessment on spectral interferences of different packaging materials. *Journal of Food Engineering*, 357, 111653. https://doi.org/10.1016/j.jfoodeng.2023.111653

Downey, G., Sheehan, E., Delahunty, C., O'Callaghan, D., Guinee, T. & Howard, V. (2005). Prediction of maturity and sensory attributes of Cheddar cheese using near-infrared spectroscopy. *International Dairy Journal*, 15 (6), 701–709. https://doi.org/10.1016/j.idairyj.2004.06.013

Dufour, É. (2009). Chapter 1 - Principles of Infrared Spectroscopy. I: Sun, D.-W. (red.) *Infrared Spectroscopy for Food Quality Analysis and Control*. Academic Press. 1–27. https://doi.org/10.1016/B978-0-12-374136-3.00001-8

FAO (u.å.). *Milk Production*. *DPP*. https://www.fao.org/dairy-production-products/production/en/ [2024-05-10]

Genis, D.O., Sezer, B., Durna, S. & Boyaci, I.H. (2021). Determination of milk fat authenticity in ultra-filtered white cheese by using Raman spectroscopy with multivariate data analysis. *Food Chemistry*, 336, 127699. https://doi.org/10.1016/j.foodchem.2020.127699

Hebling e Tavares, J.P., da Silva Medeiros, M.L. & Barbin, D.F. (2022). Near-infrared techniques for fraud detection in dairy products: A review. *Journal of Food Science*, 87 (5), 1943–1960. https://doi.org/10.1111/1750-3841.16143

Jin, H., Lu, Q., Chen, X., Ding, H., Gao, H. & Jin, S. (2016). The use of Raman spectroscopy in food processes: A review. *Applied Spectroscopy Reviews*, 51 (1), 12–22. https://doi.org/10.1080/05704928.2015.1087404

Kamruzzaman, M., Makino, Y. & Oshita, S. (2016). Parsimonious model development for real-time monitoring of moisture in red meat using hyperspectral imaging. *Food Chemistry*, 196, 1084–1091. https://doi.org/10.1016/j.foodchem.2015.10.051

Khan, M.J., Khan, H.S., Yousaf, A., Khurshid, K. & Abbas, A. (2018). Modern Trends in Hyperspectral Image Analysis: A Review. *IEEE Access*, 6, 14118–14129. https://doi.org/10.1109/ACCESS.2018.2812999

Lei, T. & Sun, D.-W. (2019). Developments of nondestructive techniques for evaluating quality attributes of cheeses: A review. *Trends in Food Science & Technology*, 88, 527–542. https://doi.org/10.1016/j.tifs.2019.04.013

Leite, A.I.N., Pereira, C.G., Andrade, J., Vicentini, N.M., Bell, M.J.V. & Anjos, V. (2019). FTIR-ATR spectroscopy as a tool for the rapid detection of adulterations in butter cheeses. *LWT*, 109, 63–69. https://doi.org/10.1016/j.lwt.2019.04.017

Li Vigni, M., Durante, C., Michelini, S., Nocetti, M. & Cocchi, M. (2020). Preliminary Assessment of Parmigiano Reggiano Authenticity by Handheld Raman Spectroscopy. *Foods*, 9 (11), 1563. https://doi.org/10.3390/foods9111563

Liu, N., Parra, H.A., Pustjens, A., Hettinga, K., Mongondry, P. & van Ruth, S.M. (2018). Evaluation of portable near-infrared spectroscopy for organic milk authentication. *Talanta*, 184, 128–135. https://doi.org/10.1016/j.talanta.2018.02.097

Lu, Y., Saeys, W., Kim, M., Peng, Y. & Lu, R. (2020). Hyperspectral imaging technology for quality and safety evaluation of horticultural products: A review and celebration of the past 20-year progress. *Postharvest Biology and Technology*, 170, 111318. https://doi.org/10.1016/j.postharvbio.2020.111318

Malegori, C., Oliveri, P., Mustorgi, E., Boggiani, M.A., Pastorini, G. & Casale, M. (2021). An in-depth study of cheese ripening by means of NIR hyperspectral imaging: Spatial mapping of dehydration, proteolysis and lipolysis. *Food Chemistry*, 343, 128547. https://doi.org/10.1016/j.foodchem.2020.128547

Medeiros, M.L. da S., Freitas Lima, A., Correia Gonçalves, M., Teixeira Godoy, H. & Fernandes Barbin, D. (2023). Portable near-infrared (NIR) spectrometer and chemometrics for rapid identification of butter cheese adulteration. *Food Chemistry*, 425, 136461. https://doi.org/10.1016/j.foodchem.2023.136461

Montgomery, H., Haughey, S.A. & Elliott, C.T. (2020). Recent food safety and fraud issues within the dairy supply chain (2015–2019). *Global Food Security*, 26, 100447. https://doi.org/10.1016/j.gfs.2020.100447

Müller-Maatsch, J. & van Ruth, S.M. (2021). Handheld Devices for Food Authentication and Their Applications: A Review. *Foods*, 10 (12), 2901. https://doi.org/10.3390/foods10122901

Ostovar pour, S., Afshari, R., Landry, J., Pillidge, C., Gill, H. & Blanch, E. (2021). Spatially offset Raman spectroscopy: A convenient and rapid tool to distinguish cheese made with milks from different animal species. *Journal of Raman Spectroscopy*, 52 (10), 1705–1711. https://doi.org/10.1002/jrs.6179

Parker, A.W. (2017). Spatially Offset Raman Spectroscopy. I: Lindon, J.C., Tranter, G.E., & Koppenaal, D.W. (red.) *Encyclopedia of Spectroscopy and Spectrometry (Third Edition)*. Academic Press. 143–148. https://doi.org/10.1016/B978-0-12-409547-2.12151-1

Parmigiano Consortium (2018). For Parmigiano Reggiano cheese: Parmigiano Reggiano cheese production standards

Paszczyk, B. & Łuczyńska, J. (2020). The Comparison of Fatty Acid Composition and Lipid Quality Indices in Hard Cow, Sheep, and Goat Cheeses. *Foods*, 9 (11), 1667. https://doi.org/10.3390/foods9111667

Pi, F., Shinzawa, H., Ozaki, Y. & Han, D. (2009). Non-destructive determination of components in processed cheese slice wrapped with a polyethylene film

using near-infrared spectroscopy and chemometrics. *International Dairy Journal*, 19 (10), 624–629. https://doi.org/10.1016/j.idairyj.2009.05.002
Priyashantha, H., Höjer, A., Saedén, K.H., Lundh, Å., Johansson, M., Bernes, G., Geladi, P. & Hetta, M. (2020). Use of near-infrared hyperspectral (NIR-HS) imaging to visualize and model the maturity of long-ripening hard cheeses. *Journal of Food Engineering*, 264, 109687. https://doi.org/10.1016/j.jfoodeng.2019.109687
Pu, Y., Pérez-Marín, D., O'Shea, N. & Garrido-Varo, A. (2021). Recent Advances in Portable and Handheld NIR Spectrometers and Applications in Milk, Cheese and Dairy Powders. *Foods*, 10 (10), 2377. https://doi.org/10.3390/foods10102377
Reis, M.M., Dixit, Y., Carr, A., Tu, C., Palevich, F., Gupta, T. & Reis, M.G. (2023). Hyperspectral imaging through vacuum packaging for monitoring cheese biochemical transformation caused by *Clostridium* metabolism. *Food Research International*, 169, 112866. https://doi.org/10.1016/j.foodres.2023.112866
Rinnan, Å., Nørgaard, L., Berg, F. van den, Thygesen, J., Bro, R. & Engelsen, S.B. (2009). Chapter 2 - Data Pre-processing. I: Sun, D.-W. (red.) *Infrared Spectroscopy for Food Quality Analysis and Control*. Academic Press. 29–50. https://doi.org/10.1016/B978-0-12-374136-3.00002-X
Siche, R., Vejarano, R., Aredo, V., Velasquez, L., Saldaña, E. & Quevedo, R. (2016). Evaluation of Food Quality and Safety with Hyperspectral Imaging (HSI). *Food Engineering Reviews*, 8 (3), 306–322. https://doi.org/10.1007/s12393-015-9137-8
da Silva Medeiros, M.L., Moreira de Carvalho, L., Madruga, M.S., Rodríguez-Pulido, F.J., Heredia, F.J. & Fernandes Barbin, D. (2024). Comparison of hyperspectral imaging and spectrometers for prediction of cheeses composition. *Food Research International*, 183, 114242. https://doi.org/10.1016/j.foodres.2024.114242
Smith, G.P.S., Holroyd, S.E., Reid, D.C.W. & Gordon, K.C. (2017). Raman imaging processed cheese and its components. *Journal of Raman Spectroscopy*, 48 (3), 374–383. https://doi.org/10.1002/jrs.5054
Soryal, K.A., Zeng, S.S., Min, B.R., Hart, S.P. & Beyene, F.A. (2004). Effect of feeding systems on composition of goat milk and yield of Domiati cheese. *Small Ruminant Research*, 54 (1), 121–129. https://doi.org/10.1016/j.smallrumres.2003.10.010
Steinbach, D., Anderson, C.A., McGeorge, G., Igne, B., Bondi, R.W. & Drennen, J.K. (2017). Calibration Transfer of a Quantitative Transmission Raman PLS Model: Direct Transfer vs. Global Modeling. *Journal of Pharmaceutical Innovation*, 12 (4), 347–356. https://doi.org/10.1007/s12247-017-9299-4
Stocco, G., Gómez-Mascaraque, L.G., Deshwal, G.K., Sanchez, J.C., Molle, A., Pizzamiglio, V., Berzaghi, P., Gergov, G. & Cipolat-Gotet, C. (2024). Exploring the use of NIR and Raman spectroscopy for the prediction of quality traits in PDO cheeses. *Frontiers in Nutrition*, 11. https://doi.org/10.3389/fnut.2024.1327301
Tarapoulouzi, M., Kokkinofta, R. & Theocharis, C.R. (2020). Chemometric analysis combined with FTIR spectroscopy of milk and Halloumi cheese samples according to species' origin. *Food Science & Nutrition*, 8 (7), 3262–3273. https://doi.org/10.1002/fsn3.1603
Tarapoulouzi, M., Logan, N., Hardy, M., Montgomery, H., Haughey, S.A., Elliott, C.T. & Theocharis, C.R. (2024). A Pre-Trial Study to Identify Species of Origin in Halloumi Cheese Utilising Chemometrics with Near-Infrared and Hyperspectral Imaging Technologies. *Analytica*, 5 (1), 17–27. https://doi.org/10.3390/analytica5010002

Thain, S. (2022). *IR Spectroscopy and FTIR Spectroscopy: How an FTIR Spectrometer Works and FTIR Analysis*. *Analysis & Separations from Technology Networks*. http://www.technologynetworks.com/analysis/articles/ir-spectroscopy-and-ftir-spectroscopy-how-an-ftir-spectrometer-works-and-ftir-analysis-363938 [2024-05-23]

Tolentino, I.C. de S.A., Jesus, J.C. de, Conceição, D.G., Onelli, R.R.V., Reis, L.C.C., Santos, L.S. & Ferrão, S.P.B. (2023). Use of MIR spectroscopy associated with chemometric techniques to verify the authenticity of prato cheese. *International Dairy Journal*, 142, 105643. https://doi.org/10.1016/j.idairyj.2023.105643

Vásquez, N., Magán, C., Oblitas, J., Chuquizuta, T., Avila-George, H. & Castro, W. (2018). Comparison between artificial neural network and partial least squares regression models for hardness modeling during the ripening process of Swiss-type cheese using spectral profiles. *Journal of Food Engineering*, 219, 8–15. https://doi.org/10.1016/j.jfoodeng.2017.09.008

Visconti, L.G., Martínez Vargas, S., Rodríguez, M.S., Di Anibal, C.V. & Delrieux, C. (2023). Determination of grated hard cheese adulteration by digital image analysis and multivariate analysis. *International Dairy Journal*, 138, 105539. https://doi.org/10.1016/j.idairyj.2022.105539

Visconti, L.G., Rodríguez, M.S. & Di Anibal, C.V. (2020). Determination of grated hard cheeses adulteration by near infrared spectroscopy (NIR) and multivariate analysis. *International Dairy Journal*, 104, 104647. https://doi.org/10.1016/j.idairyj.2020.104647

Weimer, B.C. (2007). *Improving the Flavour of Cheese*. Elsevier.

Woodcock, T., Fagan, C.C., O'Donnell, C.P. & Downey, G. (2008). Application of Near and Mid-Infrared Spectroscopy to Determine Cheese Quality and Authenticity. *Food and Bioprocess Technology*, 1 (2), 117–129. https://doi.org/10.1007/s11947-007-0033-y

Wu, X., Shin, S., Gondhalekar, C., Patsekin, V., Bae, E., Robinson, J.P. & Rajwa, B. (2023). Rapid Food Authentication Using a Portable Laser-Induced Breakdown Spectroscopy System. *Foods*, 12 (2), 402. https://doi.org/10.3390/foods12020402

Yakubu, H.G., Kovacs, Z., Toth, T. & Bazar, G. (2022). The recent advances of near-infrared spectroscopy in dairy production—a review. *Critical Reviews in Food Science and Nutrition*, 62 (3), 810–831. https://doi.org/10.1080/10408398.2020.1829540

Zareef, M., Chen, Q., Hassan, M.M., Arslan, M., Hashim, M.M., Ahmad, W., Kutsanedzie, F.Y.H. & Agyekum, A.A. (2020). An Overview on the Applications of Typical Non-linear Algorithms Coupled With NIR Spectroscopy in Food Analysis. *Food Engineering Reviews*, 12 (2), 173–190. https://doi.org/10.1007/s12393-020-09210-7

Zedda, L., Perniciano, A., Loddo, A. & Di Ruberto, C. (2024). Understanding cheese ripeness: An artificial intelligence-based approach for hierarchical classification. *Knowledge-Based Systems*, 295, 111833. https://doi.org/10.1016/j.knosys.2024.111833

Zhang, Z.-Y. (2020). Rapid Discrimination of Cheese Products Based on Probabilistic Neural Network and Raman Spectroscopy. *Journal of Spectroscopy*, 2020, e8896535. https://doi.org/10.1155/2020/8896535

Zhang, Z.-Y., Jiang, M.-Q. & Xiong, H.-M. (2023). Optimized identification of cheese products based on Raman spectroscopy and an extreme learning machine. *New Journal of Chemistry*, 47 (14), 6889–6894. https://doi.org/10.1039/D2NJ06350F

Acknowledgements

I would like to thank my supervisor, Prof. Åse Lundh, as well as my father, Prof. Peter Hambäck, for patiently reading through this, and many less revised drafts and giving good comments on improvement. I would also like to thank the rest of my course mates, who have been greatly supportive throughout the period of writing. I am gladly looking forward to seeing your results as well.

6. Publishing and archiving

Approved students' theses at SLU are published electronically. As a student, you have the copyright to your own work and need to approve the electronic publishing. If you check the box for **YES**, the full text (pdf file) and metadata will be visible and searchable online. If you check the box for **NO**, only the metadata and the abstract will be visible and searchable online. Nevertheless, when the document is uploaded it will still be archived as a digital file. If you are more than one author, the checked box will be applied to all authors. You will find a link to SLU's publishing agreement here:

- https://libanswers.slu.se/en/faq/228318.

☒ YES, I/we hereby give permission to publish the present thesis in accordance with the SLU agreement regarding the transfer of the right to publish a work.

☐ NO, I/we do not give permission to publish the present work. The work will still be archived and its metadata and abstract will be visible and searchable.