Application of Machine Learning for Detection of Food Spoilage

Fabio Knoll
Data Science and Intelligent Analytics
FH Kufstein Tirol
Kufstein, Austria
2010837227@stud.fh-kufstein.ac.at

ABSTRACT

Food spoilage is a problem that has affected all of humanity since its existence, as the affected food either has undesirable odor or taste characteristics, or consumption is harmful to health. The signs of spoilage are sometimes difficult or impossible to detect by humans and can only be identified by special instruments and analytical methods. Machine learning could optimize existing methods here or enable completely new approaches. In this paper the current state of research regarding the application of machine learning methods for the detection of food spoilage is investigated based on a review of pertinent literature. The research shows that machine learning is used in combination with a wide variety of analysis techniques and provides good predictive performance.

KEYWORDS

food spoilage, machine learning, data science, classification, detection, prediction

1 Introduction

Industrialization and globalization have given humanity, and especially the inhabitants of highly developed countries, unprecedented access to an incredible amount and variety of different foods. This is only possible through the use of special treatments that slow down the spoilage of foods and increase the shelf life [1]. However, no matter which type of food and which treatment applied, spoilage cannot be prevented completely.

Most of this spoilage is caused by microorganisms that infest the food, colonize it and produce chemicals that are toxic or have a repugnant taste or smell [2].

Food spoilage is also one of the main reasons for food loss along the entire supply chain, mainly caused by errors in harvesting, transportation and storage of food [2]. Another reason is the disposal of perfectly edible food by end consumers, because it exceeds the minimum shelf life and they are not able to evaluate the condition of the food themselves and want to be on the safe side [2]. The consumption of spoiled food may lead to severe health risks and diseases like food poisoning potentially even leading to death [1].

The development and use of food spoilage detection methods and technologies are therefore of great importance and the use of machine learning methods might lead to an improvement in detection accuracy. Therefore, a literature search was conducted using Google Scholar, selecting the most recent and well-cited papers in order to find out, to what extent machine learning methods are currently used in research in this area.

2 Food Spoilage

The spoilage of food can have several causes, like the breakdown of complex molecules or the consumption by insects and rodents, but the most common cause are microorganisms or microbes [2]. These microbes use the food as a carbon and energy source and thereby cause chemical reactions leading to deterioration of the food [1]. There is a huge variety of different types of microbes found on spoiled foods, which can be more or less selective regarding the type of food they consume and a food item can be colonized by one or more species depending on them being competitors or not. On a high level, theses microbes can be divided into bacteria and fungi [2], and fungi can be divided into yeasts and molds [1]. Bacteria is the more dangerous type of microbes, as the spoilage of the affected food is often not visually recognizable and the presence of dangerous toxins and spores is only detected in the event of an outbreak of a disease [1]. In the following, the most relevant types of bacteria and fungi with regard to spoilage are indicated and explained.

2.1 Bacteria

According to the overview provided by [1], there are several different groups or species of bacterias. There is the group of spore-forming bacteria, whose spores are able to survive at high temperatures, some of them being thermophilic and growing at temperatures up to 55°C and therefore often found in heat-treated foods. Other thermophile bacteria are *Bacillus* and *Geobacillus* producing a flat sour spoilage in canned foods. Another well-known group of bacteria are *Lactic Acid Bacteria* (LAB) and some of them are used for the fermentation of foods e.g. yogurt. Under low oxygen and temperature and in an acidic environment they are causing spoilage of a variety of foods, as for example greening of and slime production on meat, as well as gas formation in cheeses, pickles and packaged meat and vegetables. Some species of the group of Pseudomonas comprise up to 40% of naturally occurring bacteria on fruits and vegetables and are responsible for about half of post-harvest rot of cold-stored products. Further species of *Pseudomonas* cause spoilage of animal products such as meat, fish and milk. *Enterobacteriaceae* are often found in soil, on plants and in digestive tracts of animals and present in many foods. They include not only many spoilage organisms but also human pathogens such as *Salmonella* and *Yersina*.

2.2 Fungi

According to Rawat [1], Fungi are eukaryotic microbes that includes yeasts, molds and mushrooms. He states that the subset of the yeasts life in specialized, usually liquid environments and do not produce toxins, thus they are often used for fermentations in bread and alcoholic beverage production. They colonize with or without oxygen on foods with high sugar or salt content and spoil foods like maple syrup, pickles, sauerkraut, meet, cheese and fruits and juices with a low pH. Although most spoilage yeasts do not cause human infections, the affected food is often undesirable for consumption. The other subset of fungi causing food spoilage are molds, which are filamentous fungi without large fruiting bodies. Rawat indicates that they are important for the degradation of dead plant and animals, but are also found on many foods, requiring oxygen for their metabolic process. They grow on and through solid substrates and their spores are spread through the air. Unlike yeasts, they can produce toxic and carcinogenic mycotoxins and harm human health, but some species are also useful for production of fermented soy products, antibiotics and blue cheese. Molds spoil a wide variety of food and non-food items, but the most affected are fruits, vegetables, grains, beans, nuts, spices and some processed foods like jam and dairy products. [1]

2.3 Microbial Count

In microbiology, there are a number of techniques for biochemical identification and enumeration of microorganisms. The most conventional method is the viable plate count or standard plate count, in which the piece of food is first diluted to a certain concentration, the diluted liquid is then pipetted into a sterile petri dish, and then a liquefied agar medium is added. The petri dish is then placed in an incubator at a specific temperature for 24h-48h so that the microorganisms grow into visible cultures that can finally be counted. This method is very labor and time consuming and leads to waste of petri dishes and other disposables in the laboratory, but it is still widely used worldwide. By using special culture media, it is possible to identify and count specific species of microbes. [3]

3 Detection using Machine Learning

There are already a number of studies that use different machine learning models to predict the number of bacteria present on the food or to overall classify the freshness of the food. Here, most studies focus on the analysis of a specific type of food e.g. meat [4–7], rather than on the development of generalized model for all types of food.

It is important to distinguish the methods used to collect the underlying data for model prediction. The more classical methods are based on special sensors and chemical techniques and usually have the disadvantage of being slow, inefficient, impractical, and often destroy the food during analysis, making them unsuitable for use in production and by the end consumer [6]. Estelles-Lopez et al. used machine learning models for meat spoilage prediction and mainly utilized classical methods such as High Performance Liquid Chromatography (HPLC), Fourier Transformed Infrared Spectroscopy (FT-IR), Gas Chromatography coupled Mass Spectrometry (GC-MS), Multi-Spectral Imaging (MSI) and Electrical Noses on samples stored under different packaging types such as aerobic packaging (AIR) and modified atmosphere packaging (MAP) [4]. But there are also imaging methods such as Computer Vision (CV) that visually assess food quality based on visual features such as color, shape, size and surface texture features [6]. These methods have the advantage of fast, nondestructive and efficient prediction of food quality or spoilage and are also suitable for application in production or even by the consumer, but are limited to the identification of external quality and spoilage factors [6]. In the following, some of the methods are explained in more detail and results obtained from machine learning models in combination with these methods are shown.

3.1 Fourier Transformed Infrared Spectroscopy

According to the definition provided by [8], FT-IR is a method of infrared (IR) spectroscopy, where infrared radiation is used to change the vibrational behavior of molecules depending on wavelength and leads to an energy absorption, if the frequency corresponds to the molecular vibrational mode of the molecule or chemical group. The most popular and widely used frequency range is the mid IR spectrum, which ranges between 4000 and 400 cm^{-1} and covers the vibrational frequencies of most molecules. By IR irradiating a substance in a continuous spectrum and measuring the IR beam before and after the substance, a complex resonance-absorption curve is obtained, which corresponds to a fingerprint of the substance. Thus, when a bacterial cell is irradiated, its chemical composition can be identified. [8]

In Figure 1, an example FT-IR spectrum can be seen.

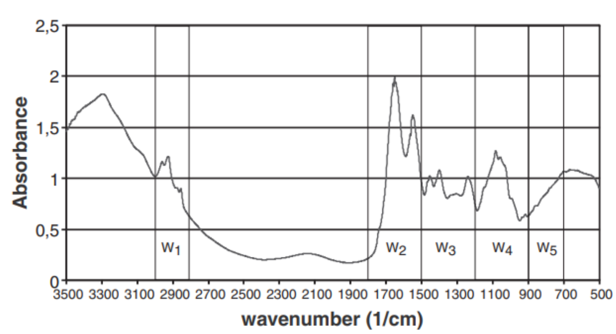

Figure 1: Representative FT-IR spectrum from a bacterial cells sample [8]

In the study of Estelle-Lopez et al. small portions of beef meat were analyzed with a FT-IR spectrometer and the collected spectra data were used as inputs to several machine learning models to predict bacterial counts on the meat [4]. The best predictions for all types of bacteria and environments resulted from a Random Forest Regressor model, with the best accuracy of 92.3% for LAB for samples stored under both AIR and MAP packaging. However, compared to all the other data collection methods FT-IR performed worst. In another study, the FT-IR method was used in combination with a neural network to classify samples of beef fillets into the categories fresh, semi-fresh and spoiled and an overall correct classification accuracy of 90.5% was reached [5].

3.2 High Performance Liquid Chromatography and Gas Chromatography coupled to Mass Spectrometry

The analytical methods HPLC and GC-MS have also been utilized in the study of Estelles-Lopez et al., but the exact explanation of these methods would require extensive chemical background knowledge and is beyond the scope of this paper. The output of both methods are peaks on a spectrum and in the work of Estelle-Lopez et al. these peaks are used as input into the machine learning models. [4]

It was found that both HPLC and GC-MS data is very suitable for the prediction of all types of bacterial counts on the meat regardless of the regression models used and the type of packaging. With the HPLC data, the best prediction accuracy of 100% was achieved for LAB by a Random Forest Regression for samples stored under MAP packaging and with the GC-MS data, the best accuracy of 96% was achieved for *Pseudomonads* by kNN Regression also for samples stored under MAP. Although the two methods provide good predictions they are not suitable for use in the production chain and certainly not for the end consumer, as they are very labor-intensive, time-consuming, expensive and due to its size not really portable. [4]

3.3 Multi-Spectral Imaging

In MSI, multi-spectral images in a number of different wavelengths are acquired. In more detail, surface reflections are recorded using a Point Grey Scorpion camera with a standard monochrome charge coupled device chip. The sample is placed into an Ulbricht sphere with a matt interior coating and the camera is mounted on top of the sphere. Special LEDs with narrow-band spectral radiation distribution are placed at the edge of the sphere and due to the coating and curvature, uniform light is present in the sphere. Images are then taken at specific wavelength settings of the LEDs, creating a hyperspectral dimensional cube. [7]

An example of recorded multispectral images can be seen in Figure 2.

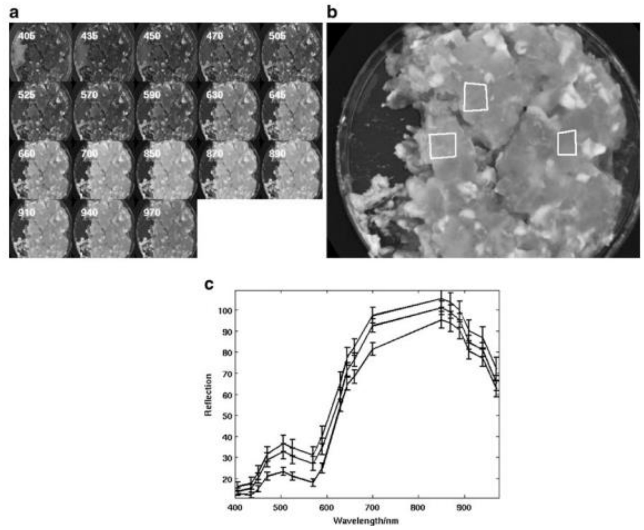

Figure 2: (a) Images at different wavelengths (b) Image recorded at 525 nm wavelength (c) Mean spectra of square areas in (b) [7]

In a study from 2013, spoilage detection of pork meat using MSI was investigated [7]. First, the images acquired were automatically preprocessed and segmented into regions of interest. Next, they applied an unsupervised K-means clustering based on the spectral images. The number of pixels in the images belonging to various clusters correlates with the spoilage degree, thus a regression model on the size of these areas was used to predict the spoilage degree. Based on this output a discrimination between the three classes fresh, semi-fresh and spoiled was made and an overall classification accuracy of 76.1% was reached. In [4] the images acquired through MSI were preprocessed and segmented in a similar way and afterwards the mean and standard deviation of the reflectance spectrum was calculated for each image and used as input for the machine learning models. The results show that data acquired with MSI are very suitable in combination with k-Nearest Neighbors Regression, Support Vector Regression, Random Forests Regression and Partial Least Squares Regression or Principal Component Regression. The best prediction accuracy of 97.5% was achieved for LAB by Random Forest Regression for samples stored under AIR packaging. The use of this method is convenient and could possibly be used in production line, as it is portable, fast and provides accurate results [4].

3.4 Electrical Noses

An electronic nose (eNose) is a device that uses sensors and a suitable pattern recognition system to identify and distinguish odors. The technology is based on the functional principle of the human nose by analyzing the volatile components of the samples with the sensor array, which imitates the olfactory cells of the human nose. [9]

The eNose in [4] resulted in the best prediction accuracy of 94.6% for *Enterobacteriaceae* combined with Random Forest

Regression for samples stored under MAP packaging and also performed quite well overall. In another study from Mohd Ali et al. recent applications of eNoses for the quality measurement of agricultural and food products were investigated and a variety of different statistical models, such as PCA, LDA, ANN, SVM etc. were used for the data analysis in the reviewed studies [10].

3.5 Computer Vision

For the use of computer vision, in addition to the image processing computer hardware, only an ordinary camera for image acquisition and a light source of good quality and uniformity are required. An efficient illumination system that avoids reflections, shadows and noise leads to a better quality and accuracy of the image and saves time during image processing. [6]

In the study of Barbri et al. computer vision was utilized to evaluate beef meat quality and in parallel to the computer vision method a microbiological analysis was used for reference [11]. The bacteriological results of the analyzed samples were used to divide the samples into two classes - spoiled and unspoiled - using a threshold of 10^6 cfu/g for the total viable counts of microbial population. After labeling the samples, a Fuzzy ARTMAP neural network model was developed to classify the samples based on their color attributes and the best accuracy obtained was 95.24%. In Figure 3 the changes in count of bacteria and the changes in color over a certain time period are shown.

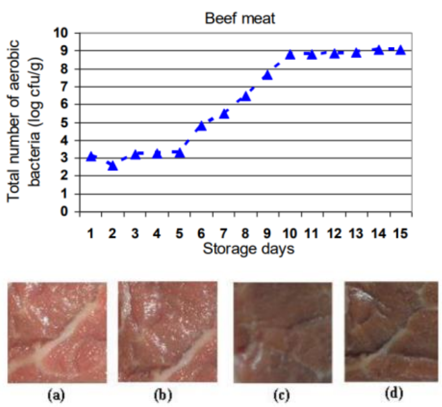

Figure 3: Changes in count of bacteria (top), Evolutionary color change illustration at: (a) day 2, (b) day 6, (c) day 8 and (d) day 10 [11]

In another study of Huang et al. computer vision was combined with near-infrared spectroscopy methods for the freshness classification of fish and the trained Back Propagation Artificial Neural Network model obtained an accuracy of 93.33%. [12]

4 Conclusion

The studies reviewed in this paper show that the use of machine learning models for food spoilage detection is widespread. Either regression models are used to predict the number of bacteria or classification models are used to directly classify food into spoiled and non-spoiled. The application of the methods takes place both on data collected by classical chemical methods, which are often very slow, destructive and inefficient, and on data collected by imaging methods, which rely on visual features. This latter method is nondestructive, faster and more flexible and has the potential for use in production or even by the end consumer. As this area is particularly suitable for the application of machine learning and especially deep learning algorithms, it also offers room for further studies in the future.

REFERENCES

[1] S. Rawat, Food Spoilage: Microorganisms and their prevention, Asian Journal of Plant Science & Research 5 (2015).
[2] S.T. Hammond, J.H. Brown, J.R. Burger, T.P. Flanagan, T.S. Fristoe, N. Mercado-Silva, J.C. Nekola, J.G. Okie, Food Spoilage, Storage, and Transport: Implications for a Sustainable Future, BioScience 65 (2015) 758–768. https://doi.org/10.1093/biosci/biv081.
[3] D.Y.C. Fung, BIOCHEMICAL AND MODERN IDENTIFICATION TECHNIQUES | Introduction, in: Encyclopedia of Food Microbiology, Elsevier, 2014, pp. 223–231.
[4] L. Estelles-Lopez, A. Ropodi, D. Pavlidis, J. Fotopoulou, C. Gkousari, A. Peyrodie, E. Panagou, G.-J. Nychas, F. Mohareb, An automated ranking platform for machine learning regression models for meat spoilage prediction using multi-spectral imaging and metabolic profiling, Food Res. Int. 99 (2017) 206–215. https://doi.org/10.1016/j.foodres.2017.05.013.
[5] A.A. Argyri, E.Z. Panagou, P.A. Tarantilis, M. Polysiou, G.-J.E. Nychas, Rapid qualitative and quantitative detection of beef fillets spoilage based on Fourier transform infrared spectroscopy data and artificial neural networks, Sensors and Actuators B: Chemical 145 (2010) 146–154. https://doi.org/10.1016/j.snb.2009.11.052.
[6] A. Taheri-Garavand, S. Fatahi, M. Omid, Y. Makino, Meat quality evaluation based on computer vision technique: A review, Meat Sci. 156 (2019) 183–195. https://doi.org/10.1016/j.meatsci.2019.06.002.
[7] B.S. Dissing, O.S. Papadopoulou, C. Tassou, B.K. Ersbøll, J.M. Carstensen, E.Z. Panagou, G.-J. Nychas, Using Multispectral Imaging for Spoilage Detection of Pork Meat, Food Bioprocess Technol 6 (2013) 2268–2279. https://doi.org/10.1007/s11947-012-0886-6.
[8] A. Alvarez-Ordóñez, D.J.M. Mouwen, M. López, M. Prieto, Fourier transform infrared spectroscopy as a tool to characterize molecular composition and stress response in foodborne pathogenic bacteria, J. Microbiol. Methods 84 (2011) 369–378. https://doi.org/10.1016/j.mimet.2011.01.009.
[9] A. Sanaeifar, H. ZakiDizaji, A. Jafari, M.d. La Guardia, Early detection of contamination and defect in foodstuffs by electronic nose: A review, TrAC Trends in Analytical Chemistry 97 (2017) 257–271. https://doi.org/10.1016/j.trac.2017.09.014.
[10] M. Mohd Ali, N. Hashim, S. Abd Aziz, O. Lasekan, Principles and recent advances in electronic nose for quality inspection of agricultural and food products, Trends in Food Science & Technology 99 (2020) 1–10. https://doi.org/10.1016/j.tifs.2020.02.028.
[11] N.E. Barbri, A. Halimi, K. Rhofir, A Nondestructive Method Based on an Artificial Vision for Beef Meat Quality Assesement, International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering 02 (2014). https://doi.org/10.17148/IJIREEICE.2014.0210001.
[12] X. Huang, H. Xu, L. Wu, H. Dai, L. Yao, F. Han, A data fusion detection method for fish freshness based on computer vision and near-infrared spectroscopy, Analytical Methods 8 (2016) 2929–2935.