Microwave-Assisted Detection of Physical Intrusions in Commercial Food Packaged Products via Machine Learning

Microwave-Assisted Detection of Physical Intrusions in Commercial Food Packaged Products via Machine Learning

Ali Darwish[#$], Marci Ricci[#], Jorge Alberto Tobon Vasquez[#], Claire Migliaccio [$1], Francesca Vipiana[#2]

[#]DET, Politecnico di Torino, Italy

[$]LEAT, Université Côte d'Azur, France

[1]claire.migliaccio@univ-cotedazur.fr, [2]francesca.vipiana@polito.it

***Abstract*— This study introduces a novel approach using microwaves (MW) combined with a machine learning (ML) classifier to detect physical intrusions within food packaging. Our objective is to validate our previous works [1], [2], and [3], by selecting real commercial water-based and oil-based food packaged products, specifically tomato and pesto sauce. The non-invasive MW sensing system is designed for real-time operation within a food production chain. Experimenting with both the support vector machine (SVM) and the multi-layer perceptron neural network (MLP) algorithms for binary classification, after training on datasets generated from scattering parameters acquired during measurements, resulting in a remarkable precision of 100% accuracy across 200 measurement samples for each food type. This outcome reflects the effectiveness of our previous findings.**

***Keywords*— microwave sensing, machine learning, non-destructive techniques, food safety.**

I. Introduction

The significant increase in the implementation of automation processes within the food industry is contributing to a high probability of physical contaminants generated during the different production processes inside the food packages. Detecting these contaminants through a non-destructive technology within an in-line production chain poses a challenge for companies in the food industry. To address this challenge, many technologies have been used and explored. X-ray technology is widely regarded as the most prevalent and effective method for addressing this issue [4]. However, it has limitations in detecting low-density materials. Metal detector systems [5] are constrained to identifying metals exclusively, making them ineffective in resolving the problem. Near-infrared [6] and terahertz imaging [7] techniques also have limitations in penetration depth, particularly when dealing with lossy media. In addition to the limitations mentioned before, these technologies are complex and costly systems. The exploration of new technologies has garnered significant interest from food companies.

Microwave (MW) sensing technology, assisted by machine learning tools, is making headway in various fields [8], [9] and [10]. The fundamental concept of this approach is manifested in the interaction between the emitted microwaves (MW) and the contaminants within the food jar, thereby altering the scattering behavior of these waves. The greater the contrast in permittivity between the intrusions and the background (i.e., the food materials), the more information we can gather about these intrusions. Additionally, implementing machine learning (ML) tools enables us to conduct real-time, in-line evaluations for packaged food products. Recently, we have used this technology to detect physical contaminants inside food packages that contain water-based [3] or oil-based food products [1] ,[2].

In this study, we move towards commercial food products to assess our previous works and enhance their practicality. We select off the shelf two distinct food items from the market: Pesto sauce, mainly made up of oil and cheese, categorizing it as an oil-based product and non-homogeneous medium, and a pure tomato sauce, identified as a water-based product.

II. Experimental Configurations

A. Sensing System

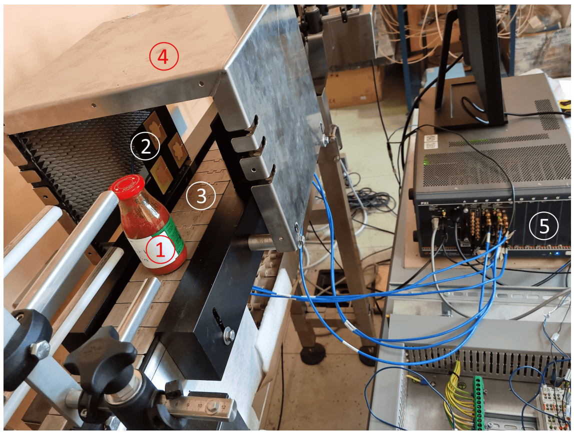

Fig. 1. The MW sensing system: (1) The food jar sample. (2) The antenna array. (3) The conveyor belt. (4) The shielding box. (5) The vector network analyzer (VNA).

The specifics of the MW sensing system (shown in Figure. 1) used in this approach are documented in [11]. However, the chosen commercial products come with a metal cap that prevents us from using the formerly designed antenna arrangement in an arc-shape above the jar, as in [1] and [2]. We have therefore completely redistributed the six antennas on either side of the jar, as shown in Figure 2. The triangular arrangement on each side ensures maximum coverage of

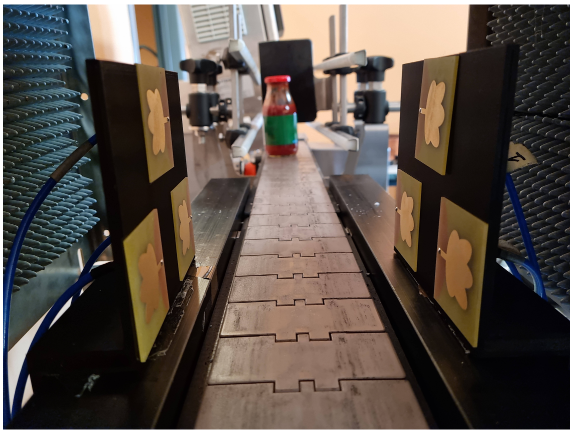

Fig. 2. The antenna array.

the jar. The system comprises six wide-band PCB-printed monopole antennas (inspired from [12]) connected to a six-port vector network analyzer (VNA) triggered by a photocell positioned at the edge of the conveyor belt. The VNA records signals emitted by the radiating antenna within the operating frequency bands. To reduce interference from lower frequencies in the industrial environment, the setup of the antenna array is placed inside a metal shielding box. In addition, the interior of the box includes a layer of MW absorber to minimize signal reflections.

B. Measurements and classifications

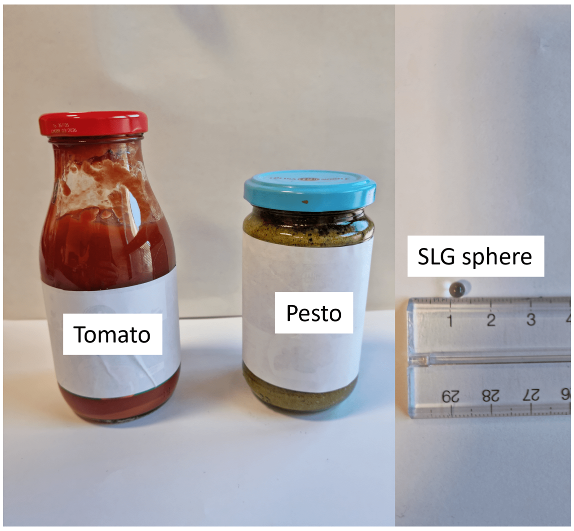

Fig. 3. Off-the-shelf packaged food products and the 4 mm diameter SLG sphere used as a contaminant.

The selection of frequency bands is a trade-off between the size of the region of interest (i.e., the dimensions of the jar), the penetration depth of the emitted MW, and the resolution required for detecting contaminants of a few millimeters in size. Taking into account all these parameters, we select two different frequency bands for conducting our measurements. For tomato sauce samples, which are water-based products with high loss characteristics and low MW penetration, we selected a frequency range of 2 to 4 GHz. While for pesto sauce samples, which are oil-based products with low-loss characteristics and better penetration of the MW within the medium, we choose a higher frequency range of 6 to 10 GHz. We take eleven frequency points equally distributed in each frequency band.

The measurement process is triggered when the jar obstructs the light reaching the photocell. This serves as the initiation point for measurements as the jar approaches the effective radiated region of the antenna array. The conveyor belt moves at a speed of 20 meters per minute, which is standard for the food industry's production chain. We prepare three samples of each type: two are uncontaminated, while the third is contaminated with a soda-lime glass (SLG) sphere measuring 4 mm in diameter. We have selected SLG material for intrusion testing because it is commonly used in packaging food and drink products, similar to the ones being tested. The samples and the intrusion are shown in Figure. 3.

A total of 400 measurements have been conducted for both food types, equally divided between the contaminated and uncontaminated samples. We rotate the sample at a slight angle each time we conduct a test. The purpose behind this practice is to put the jar in different angles of view relative to the antenna array elements during each measurement, thereby enhancing the realism and robustness of our measurements. During each measurement, the VNA captures scattering matrices at 11 frequency points, yielding a 6×6 matrix at each frequency point. In total, we obtain $6 \times 6 \times 11$ scattering parameters for each sample under test. Given the complex numbers nature of these parameters, each sample is represented by a vector of 792 entries in the classifier.

For the classification process, we use two algorithms: the non-linear SVM [13] classifier and the MLP neural network [14]. Employing two algorithms enhances the robustness of the classification process. The SVM, renowned for its robustness and precision among machine learning algorithms, functions as a supervised binary classifier. When dealing with small datasets, its effectiveness becomes particularly evident. The SVM achieves this by identifying support vectors, which are data points closest to the decision boundary (the hyperplane), to apply class separation. There are two primary categories of SVM: linear SVM and non-linear SVM. Linear SVM is preferred for straightforward classification problems where data from both classes can be linearly separated. On the other hand, the non-linear SVM is preferable for scenarios where the data are non-linearly separable and correlated. In both cases, SVM works on identifying the optimal hyperplane to effectively separate the desired classes for classification. The performance of SVM relies on the tuning of its hyperparameters, with one key parameter denoted as "C". This parameter plays a crucial role in adjusting the margin that separates the two classes. Whether linear or non-linear SVM is employed, selecting the appropriate value for "C" is essential. For non-linear SVM, an approach called the "Kernel" function is used to identify the optimal hyperplane. This method involves mapping the data into a higher-dimensional space to find a hyperplane that efficiently separates the data points.

However, employing the Kernel trick requires calculating an additional hyperparameter called "γ". Deciding on the value of "γ" is important because it influences how much weight is given to samples based on their distance from the origin of the training dataset.

Another algorithm we explore in our work is MLP. MLP is an artificial neural network with multiple fully connected layers. It is a powerful and effective tool for addressing classification problems, whether the database is high or low complexity. MLP is a supervised learning algorithm that begins with an input layer containing a vector of input data and aims to predict an output vector. In MLP, it is important to find a trade-off among the different parameters and factors such as the choice of activation function, the selection of solver (optimizer), the number of hidden layers, and the number of neurons.

The two classifiers have been implemented similarly to [15] and [1] respectively. The hyperparameters of both classifiers are the key enablers for the classification success. However, their choice is not straightforward and requires to go through an optimization process. We select the Grey Wolf Optimiser (GWO) [16], which is a bio-inspired method, known for its ability to avoid falling into local minima. For more details on the implementation of the GWO algorithm, we follow the same procedures outlined in [15].

III. Experimental Results

Table 1. Classification results.

Food	Classifier	Dataset Split (Training-Test)	Accuracy (%)
Tomato	Non-linear SVM	(300-100)	100
Tomato	Non-linear SVM	(200-200)	100
Tomato	Non-linear SVM	(150-250)	99
Tomato	Non-linear SVM	(100-300)	99
Tomato	MLP	(300-100)	100
Tomato	MLP	(200-200)	100
Tomato	MLP	(150-250)	99
Tomato	MLP	(100-300)	98
Pesto	Non-linear SVM	(300-100)	100
Pesto	Non-linear SVM	(200-200)	100
Pesto	Non-linear SVM	(150-250)	99
Pesto	Non-linear SVM	(100-300)	99
Pesto	MLP	(300-100)	100
Pesto	MLP	(200-200)	100
Pesto	MLP	(150-250)	99
Pesto	MLP	(100-300)	98

In Figures 4a and 4b, we show the projection of principal component analysis (PCA) on the three most significant eigenvectors for both the tomato and pesto datasets. These three main components exhibit a high variance: 85% for the pesto dataset and 95% for the tomato dataset. This high variance indicates that these components capture a significant amount of variation in the data, thus contributing significantly to our overall understanding of the dataset's structure. The structure of the datasets reveals a highly nonlinear relationship and correlated data in the case of pesto, while the data for tomatoes exhibit a less nonlinear relationship and correlation. This difference in the data structure may be attributed to the significant contrast in permittivity values between the tomato sauce and the glass sphere (the intrusion). This contrast arises from the water content in tomatoes, which results in a higher permittivity compared to pesto, an oil-based food. The strong correlation between the data of the two classes in both datasets supports the decision to implement nonlinear algorithms for classification, like the non-linear SVM and MLP.

[a]
[b]

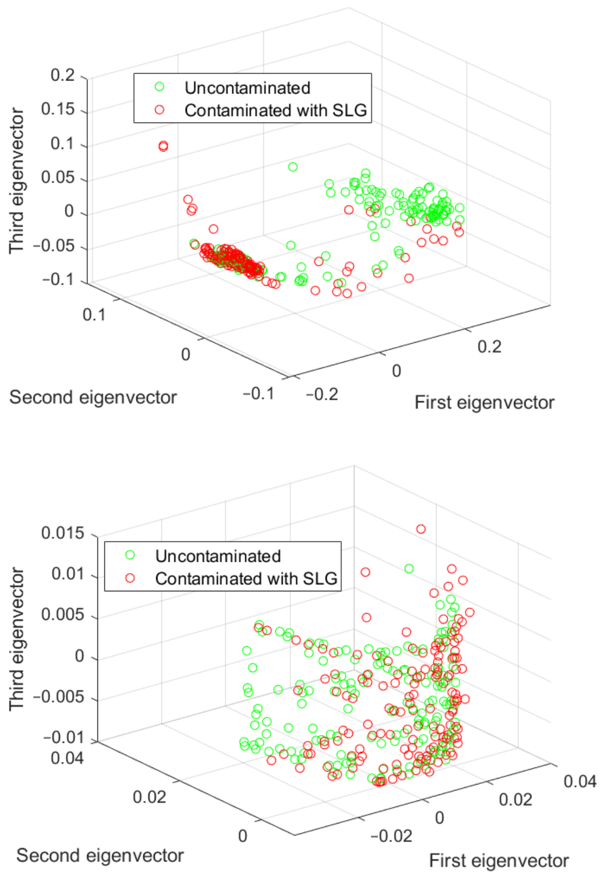

Fig. 4. Projection of PCA results on the three most significant eigenvectors. [a] Tomato, [b] Pesto.

Table 1 presents the outcomes of the classification process. We split our datasets into two segments: the training set and the test set as shown in Table 1. These results represent the average of 100 iterations of applying the classifiers, including both the training and testing phases. At each iteration, the algorithm randomly selects samples for training and testing after shuffling the datasets. The results show perfect accuracy of 100% achieved by employing both classifiers on the two different datasets.

In the context of industrial applications, we try to decrease the number of samples used for training while increasing the test segment with these samples. This interest arises from the fact that in the industry they may use a specific number of samples for the training phase, while testing may involve a significantly larger number of samples. Given this, we start to reduce the number of samples for training and observe the performance of the two classifiers with the updated dataset splits. The accuracy remains consistently perfect at 100% across all dataset splits until we reach the range of (100-300) for the training-test split. At this point, we begin to observe errors in some samples classified as uncontaminated while they are contaminated. However, the accuracy remains high, approximately at 98% for both algorithms as shown in Table 1. The slight decrease in accuracy performance across different dataset splits is normal and expected, influenced by various factors such as dataset complexity, data quality, and the selected algorithm.

IV. Conclusion

In summary, this paper introduced an innovative technique employing MW sensing technology assisted with non-linear SVM and MLP classifiers for the real-time detection of contaminants within commercially packaged tomato and pesto sauces, selected directly from store shelves. The two algorithms attained 100% precision on the two food products and with different dataset splits, ranging from 100 samples for testing to up to 250 samples out of 400. With this study, we have bridged the gap between the previous system capable of handling the detection of inclusions in a homogeneous medium (oil or water) and the detection of inclusions in inhomogeneous media, i.e. commercial products. As a part of our future work, we look to expand our research by incorporating more contaminants, implementing multiclass classifications, and extending the application of the technique to other commercial food products to enhance the generalization of our methodology.

Acknowledgment

The work is supported by Université Franco-Italienne under the doctoral grant VINCI 2021, C3-1978. In addition, the work has been supported in part by PRIN project “BEST-Food - Broadband Electromagnetic Sensing Technologies for Food quality and security assessment” (grant no. 20179FLH4A), funded by MUR, in part by the PON Research and Innovation project “Microwave Imaging and Detection powered by Artificial Intelligence for Medical and Industrial Applications (DM 1062/21)”, funded by the MUR, in part by the project “INSIGHT – An innovative microwave sensing system for the evaluation and monitoring of food quality and safety”, funded by MAECI, and in part by the Agritech National Research Center, funded by the European Union Next-GenerationEU PNRR – MIS- SIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022).

References

[1] M. Ricci, B. Štitić, L. Urbinati, *et al.*, “Machine-learning-based microwave sensing: A case study for the food industry,” *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 11, no. 3, pp. 503–514, 2021.

[2] A. Darwish, M. Ricci, F. Zidane, *et al.*, “Physical contamination detection in food industry using microwave and machine learning,” *Electronics*, vol. 11, no. 19, p. 3115, 2022.

[3] A. Darwish, M. Ricci, F. Zidane, *et al.*, “In-line microwave nondestructive evaluation of packaged food products via the support vector machine algorithm" https://2023. apsursi. org/registration. asp,” in *IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, 2023.

[4] R. P. Haff and N. Toyofuku, “X-ray detection of defects and contaminants in the food industry,” *Sensing and Instrumentation for Food Quality and Safety*, vol. 2, pp. 262–273, 2008.

[5] B. Liu and W. Zhou, “The research of metal detectors using in food industry,” in *Proceedings of 2011 International Conference on Electronics and Optoelectronics*, IEEE, vol. 4, 2011, pp. V4–43.

[6] J. U. Porep, D. R. Kammerer, and R. Carle, “On-line application of near infrared (nir) spectroscopy in food production,” *Trends in Food Science & Technology*, vol. 46, no. 2, pp. 211–230, 2015.

[7] L. Afsah-Hejri, P. Hajeb, P. Ara, and R. J. Ehsani, “A comprehensive review on food applications of terahertz spectroscopy and imaging,” *Comprehensive Reviews in Food Science and Food Safety*, vol. 18, no. 5, pp. 1563–1621, 2019.

[8] S. Caorsi and G. Cevini, “An electromagnetic approach based on neural networks for the gpr investigation of buried cylinders,” *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 1, pp. 3–7, 2005.

[9] I. T. Rekanos, “Neural-network-based inverse-scattering technique for online microwave medical imaging,” *IEEE transactions on magnetics*, vol. 38, no. 2, pp. 1061–1064, 2002.

[10] S. A. AlShehri and S. Khatun, “Uwb imaging for breast cancer detection using neural network,” *Progress In Electromagnetics Research C*, vol. 7, pp. 79–93, 2009.

[11] M. Ricci, J. A. T. Vasquez, R. Scapaticci, L. Crocco, and F. Vipiana, “Multi-antenna system for in-line food imaging at microwave frequencies,” *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 8, pp. 7094–7105, 2022.

[12] S. K. Palaniswamy, M. Kanagasabai, S. A. Kumar, M. G. N. Alsath, S. Velan, and J. K. Pakkathillam, “Super wideband printed monopole antenna for ultra wideband applications,” *International Journal of Microwave and Wireless Technologies*, vol. 9, no. 1, pp. 133–141, 2017.

[13] X. Ding, J. Liu, F. Yang, and J. Cao, “Random radial basis function kernel-based support vector machine,” *Journal of the Franklin Institute*, vol. 358, no. 18, pp. 10 121–10 140, 2021.

[14] M. W. Gardner and S. Dorling, “Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences,” *Atmospheric environment*, vol. 32, no. 14-15, pp. 2627–2636, 1998.

[15] F. Zidane, V. L. Coli, J. Lanteri, *et al.*, “Artificial intelligence-based low-terahertz imaging for archaeological shards’ classification,” *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 8, pp. 6300–6312, 2022.

[16] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey wolf optimizer,” *Advances in engineering software*, vol. 69, pp. 46–61, 2014.