A machine learning approach to microwave sensing for non-invasive alzheimer’s disease early detection

A Machine Learning Approach to Microwave Sensing for Non-invasive Alzheimer's Disease Early Detection

Leonardo Cardinali
Dept. of electronics and telecommunications
Politecnico di Torino
Torino, Italy
leonardo.cardinali@polito.it

Mattia Spano
Dept. of electronics and telecommunications
Politecnico di Torino
Torino, Italy
s292254@studenti.polito.it

Martina Gugliermino
Dept. of electronics and telecommunications
Politecnico di Torino
Torino, Italy
martina.gugliermino@polito.it

David O. Rodriguez-Duarte
Dept. of electronics and telecommunications
Politecnico di Torino
Torino, Italy
david.rodriguez@polito.it

Marco Ricci
Dept. of electronics and telecommunications
Politecnico di Torino
Torino, Italy
marco.ricci@polito.it

Jorge A. Tobon Vasquez
Dept. of electronics and telecommunications
Politecnico di Torino
Torino, Italy
jorge.tobonvasquez@polito.it

Roberta Palmeri
National Research Council of Italy
Institute for the Electromagnetic Sensing of the Environment
Napoli, Italy
palmeri.r@irea.cnr.it

Rosa Scapaticci
National Research Council of Italy
Institute for the Electromagnetic Sensing of the Environment
Napoli, Italy
scapaticci.r@irea.cnr.it

Lorenzo Crocco
National Research Council of Italy
Institute for the Electromagnetic Sensing of the Environment
Napoli, Italy
crocco.l@irea.cnr.it

Francesca Vipiana
Dept. of electronics and telecommunications
Politecnico di Torino
Torino, Italy
francesca.vipiana@polito.it

***Abstract*— Early Alzheimer's disease detection can greatly benefit patients, caregivers, and clinicians. Unfortunately, current diagnostic procedures are invasive, expensive, and not easily portable. To overcome these limitations, microwave sensing is emerging as an alternative non-invasive approach to distinguish between healthy and pathological conditions, based on the variation of permittivity in cerebrospinal fluid in Alzheimer's patients. In this framework, our paper explores the use of machine learning applied to microwave sensing data, by means of a multilayer perceptron classifier. Different architectures have been considered and appraised by relying on experimental data collected with controlled experiments involving a multi-tissue head phantom that can be filled with tissue-mimicking liquids simulating different stages of the pathology. The initial results confirm the potential of the proposed non-invasive approach to early-stage Alzheimer's disease diagnosis.**

Keywords*— *microwave sensing, machine learning, multilayer perceptron, classification, realistic phantom, Alzheimer's disease

I. INTRODUCTION

A. Alzheimer's Disease

The World Health Organization estimates that more than 55 million people worldwide have dementia. The most common form of dementia is Alzheimer's disease (AD) which may contribute to 60-70 % of the cases [1]. AD is a progressive neurodegenerative disease whose causes are yet to be fully understood, the diagnosis is not currently based on a single test, and a lot of effort is put in understanding potential causes of dementia [2]. With the help of other specialists, physicians use several different approaches, including the history of cognitive and behavioral changes of the individual, physiological tests, and the study of family history to produce a diagnosis. Cognitive and physiological symptoms in the early stages of AD include memory loss for recent events, depression, and apathy. In later stages, patients suffer impaired communication, confusion, and behavioral changes, with difficulties walking, speaking, and swallowing [2]. The cognitive symptoms arise when the illness is already in a developed state [3], and early detection of the condition is essential both for slowing the course of the disease and for offering the person living with dementia more opportunities to make decisions about financial and care planning.

The diagnosis is based on the search for AD biomarkers in laboratory tests on blood and cerebrospinal fluid (CSF), and imaging techniques such as computed tomography (CT) scans, positron-emission tomography (PET) and functional magnetic resonance imaging (fMRI) to investigate the presence of beta-amyloid plaques and tau tangles, that are clumps of proteins accumulating in the brain of AD patients [4].

Imaging techniques such as CT scans and PET tomography are considered invasive due to their use of ionizing radiation. Similarly, fMRI is expensive and not portable. Laboratory tests also have the disadvantage of

being invasive since the search of AD biomarkers relies on the CSF extraction through a lumbar puncture [5]. These biomarkers are proteins accumulating in the CSF: amyloid-beta proteins ($A\beta_{1-40}$, $A\beta_{1-42}$), tau and tau-phosphorylated [6]. Several studies investigated the correlation between the concentrations of these proteins in the CSF and AD [7]-[11] and the findings indicate that healthy and pathological conditions can be distinguished by the different concentrations in the CSF. People suffering from AD generally have higher concentrations of $A\beta_{1-40}$, tau, and tau-phosphorylated proteins while the concentration of $A\beta_{1-42}$ is generally lower.

B. Microwave Sensing

Recent studies have shown that concentration variations of proteins used as biomarkers are also related to variations of the dielectric properties (i.e., permittivity and conductivity) of the CSF in the microwave frequency range [12], [13]. In particular, the pathological condition is related to a decrease in permittivity. These findings open the way to the appraisal of AD biomarkers using microwave sensing. Thanks to the capability of penetrating material bodies, microwaves are indeed capable of non-invasively sensing the different electromagnetic properties of hidden regions, by means of a suitable processing of the field backscattered in presence of variations of the tissue properties. Such a concept, which has been for instance used for stroke detection [14]-[17] and inspection of food items [18], [19], is the basis of microwave sensing and imaging. Microwave sensing can be suitable for pathological CSF detection. The aim is to exploit the difference in dielectric properties of human tissues when they are illuminated with low-power electromagnetic (EM) waves at microwave frequencies, radiated by a set of antennas surrounding the tissues. The resulting scattered EM waves are recorded by the same sensors and processed with suitable algorithms to estimate the corresponding dielectric properties. Recent studies investigated the use of microwave imaging for AD detection at various stages [20], [21] using simulated data and measurements and obtaining promising results. They investigated the microwave behavior of tissues emulating the presence of beta-amyloid plaques and tau tangles in the brain [20] or using synthetic data of the physiological changes of the brain in AD patients (i.e., brain atrophy) [21].

C. Machine Learning Applied to CSF Microwave Sensing

In this work, we explore the use of microwaves to sense the changes in the CSF dielectric properties due to amyloid-beta accumulation, since these biomarkers of AD disease in CSF are the first symptom of the illness, allowing early detection. The difference in the scattering phenomena due to the difference in the permittivity can allow the classification of the CSF as healthy or pathological using a binary classifier. Extracting the relevant diagnostic information from a remote measurement of CSF is a challenging task, due to the non-linear and ill-posed nature of the underlying inverse scattering problem. On the other hand, an actual assessment of the EM properties values is not strictly needed for a diagnosis, being rather more important to appraise if the CSF variations sensed by the microwave probing signal is diagnostically relevant or not. For this reason, in this paper we consider the use of a machine-learning based classification approach to perform this task. Hence, we propose to use microwave sensing technology combined with machine learning (ML) to identify these variations of the CSF properties to early detect AD non-invasively and using a low-cost and portable equipment that could allow any hospital frequent acquisitions everywhere with low energy consumption.

II. Methods

A. Measurements Setup

To have real measurements to train a ML algorithm with, we build an anthropomorphic phantom of the human head emulating the dielectric behavior of different tissues [22]. We use mixtures of rubber and graphite in different percentages to mimic the electric behavior of skin, bone, fat, white matter, grey matter, cerebellum, and ventricles. CSF is realized with a liquid mixture made of Triton X-100, water, and salt [23]. To emulate various stages of AD severity, we produce five CSFs mimicking compounds. The dielectric properties of human tissues are retrieved from the IFAC-CNR database [24]. The first CSF has the permittivity and conductivity values of the healthy condition based on the IFAC-CNR database, for each frequency value the difference is less than 1.5 %. Pathological CSFs have lower permittivity than healthy ones, for this reason to simulate different severities of the AD condition we produced compounds with permittivity lowered by 7 %, 11 %, 19 %, and 24 % respect to the nominal –normal– case changing the ingredients quantities accordingly. The liquids permittivity is measured via the coaxial probe method aided by the Keysight software suite [25] and shown in Fig. 1. The CSFs can be easily inserted in the phantom by pouring it in the space between the brain and the skull and the phantom can be emptied using a hole in the bottom that is opened and closed using a clamp. In this way, we can fill and empty the phantom using different liquids without altering the phantom itself, creating a controlled experiment as shown in Fig. 2.

To retrieve information about the scattering parameters we use a Keysight 4-ports vector network analyzer (VNA) obtained using cascaded P9371A and P9375A [26], connected to four antennas as shown in Fig. 2 so that each antenna acts as transmitter and receiver. The antennas we used are circuit-printed low-complexity monopoles, each one has a custom layer of dielectric material on top and another of a different custom material on the bottom, created using a mixture of urethane rubber and graphite powder mixture. The bottom layer, in contact with the head, has the purpose of reducing the mismatch with the phantom, while the top layer is the substrate and supports the ground plane. These antennas are flexible to adapt their shape and have optimal contact with the object under test and have already been used in a previous study [27].

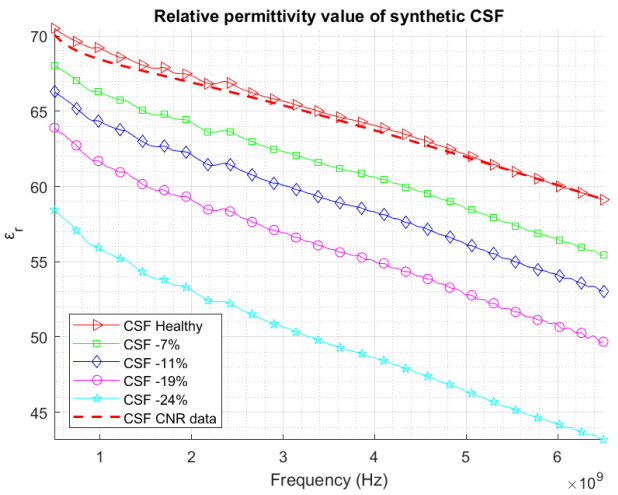

Fig. 1: Measured relative permittivity values of the compounds used to emulate human CSF. Five different compounds are present, one for the healthy and four for the pathological case; the dashed line represents the reference value used for healthy human CSF [16].

B. *Measurements Protocol*

The frequency range chosen for the investigation goes from 500 MHz to 6.5 GHz to have a trade-off between resolution and penetration depth into the head, as penetration is very limited above 6.5 GHz and the subarachnoid region has a very small electrical dimension below 500 MHz. The objective is to collect a dataset suitable for training a binary classification algorithm that distinguishes between healthy and AD condition. The dataset was built taking 28 measurement subsets in four different days. For each measurement subset the phantom is filled with the CSF representing the healthy case, then 10 consecutive measurements of the scattering parameters are performed. Then, the phantom is emptied and filled with a different CSF, measured 10 times, and the procedure is repeated for all the CSFs. Since we have four different mixtures representing AD affected CSF and only one representing healthy CSF, to have a balanced number of healthy and AD representative measurements we also performed additional 29 subsets of the healthy case alone. The final dataset including all the subsets is composed of 570 measurements of the CSF representing the healthy case and 1120 measurements of the CSF representing the AD case (280 measurements for each severity degree).

Once the measurement phase is completed, we create four different training sets to gain insights into how different subsets affect the learning process and to check how different data distributions affect robustness and generalization:

- **Training set 1:** measurement sets widely spread among all 4 days.
- **Training set 2:** measurement sets taken at the beginning and at the end of each day.
- **Training set 3:** measurement sets of days 1, 2, 4.
- **Training set 4:** measurement sets of days 1, 3.

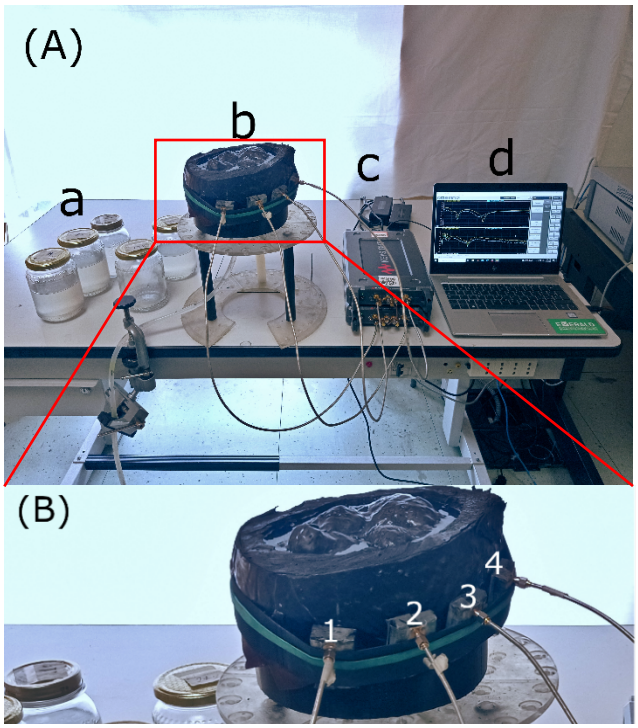

Fig. 2: (A) experimental setup; (a): synthetic CSFs; (b) multi-tissue phantom representing human head, 4 sensors are placed on one side, the phantom can be filled with different compounds and emptied through the tube on its bottom that can be opened and closed through a clamp fixed to the table; (c) 4-ports VNA; (d) central processing unit.
(B) phantom details; it is possible to fill using liquids, the port number that each sensor is connected to is displayed.

For each training set, the measurements subsets not used to train the algorithm are used to test the classification performance. The details of healthy and pathological data belonging to each dataset are summarized in Table 1.

TABLE 1
TRAINING SETS COMPOSITION

	Number of measurements		
	Healthy	***AD***	***Total***
Training set 1	370	480	850
Training set 2	290	600	890
Training set 3	430	800	1230
Training set 4	260	600	860

C. *Machine Learning Implementation*

For the classification problem we chose to use multilayer perceptron (MLP) for its non-linear modeling capabilities and for its robustness [28], [29]. We used three different methods of searching for the optimized network for our classification purpose:

- **Method 1:** We started using a two-layer feed-forward network suited for classification problems using a MATLAB ML toolbox[30]. We used random data division, scaled conjugate gradient as training function

and cross-entropy error as loss function for performance evaluation. We split the training set so that 70 % of the data is used as training for the algorithm, 15 % is used as validation, and the remaining 15 % as testing. Training was performed varying the number of neurons in the hidden layer from 1 to 100, with sigmoid activation function. The second layer is the output layer, consisting of one neuron with softmax activation function. This method is the one with the least number of degrees of freedom and the starting point in our hyperparameters exploration.

- **Method 2:** Using the same toolbox [30] we performed a design space exploration of MLP, varying: the number of hidden layers from 1 to 20, the number of neurons for each layer from 1 to 40 (all hidden layers having the same number of neurons), the maximum number of iterations (or epochs) between 150 and 1000, the learning rate value, the training function responsible for updating the network parameters, and the perform function (or loss function) used to calculate performance during learning. The training functions explored are resilient backpropagation, scaled conjugate gradient, conjugate gradient with Powell/Beale restarts, one step secant, and gradient descent. As perform functions we explored cross-entropy and mean squared error, typically used for binary classification problems [31]. The output layer is the same as method 1. Random data division was always used.

- **Method 3:** Some parameters, as the activation function of neurons in the hidden and output layers, cannot be changed using earlier methods. For this reason, we also used python ML development tools belonging to the scikit-learn library to get the optimized hyperparameters [32]. We implemented the GridSeachCV tool to systematically search through a set of hyperparameters to find the combination with the best performance. The explored parameters are hidden layer sizes and maximum number of iterations as in method 2, activation function for the hidden layer neurons, training function for weight optimization, and alpha parameter (strength of the L2 regularization term). The explored activation functions are logistic sigmoid function, hyperbolic tangent function, and rectified linear unit function. Three training functions were implemented: limited-memory BFGS [33], stochastic gradient descent, and the Adam solver [34]. In the case of stochastic gradient descent, the learning rate types explored were constant learning rate, inverse scaling learning rate that decreases at each time step, and adaptive learning rate that is lowered when training loss does not decrease by a certain amount for two consecutive epochs. The output layer is a neuron with logistic sigmoid activation function.

We applied these three approaches to all training sets, all parameters that have not been reference have been left at their default value. The optimized number of neurons found for each training set using method 1 is shown in Table 2.

TABLE 2
OPTIMIZED NUMBER OF NEURONS FOR METHOD 1

	Training set 1	***Training set 2***	***Training set 3***	***Training set 4***
Number of neurons	20	20	14	17

The optimized combinations of hyperparameters found using method 2 for each training set is shown in Table 3. While the optimized combination of hyperparameters found using method 3 for each training set is shown in Table 4.

TABLE 3
OPTIMIZED PARAMETERS FOR METHOD 2

	Training set 1	***Training set 2***	***Training set 3***	***Training set 4***
Number of hidden layers	2	3	3	1
Neurons per hidden layer	30	19	17	19
Maximum iterations	1000	1000	1000	1000
Learning rate value	3×10^{-4}	4.2×10^{-2}	4×10^{-2}	1×10^{-3}
Training function	One step secant	Scaled conjugate gradient	Scaled conjugate gradient	Scaled conjugate gradient
Peform function	Cross-entropy	Mean squared error	Mean squared error	Cross-entropy

TABLE 4
OPTIMIZED PARAMETERS FOR METHOD 3

	Training set 1	***Training set 2***	***Training set 3***	***Training set 4***
Number of hidden layers	20	10	1	1
Neurons per hidden layer	40	20	19	19
Maximum iterations	1000	1000	1000	1000
Training function	LBFGS[a]	LBFGS[a]	SGD[b]	SGD[b]
Alpha parameter	7.5×10^{-3}	1×10^{-2}	1.5×10^{-4}	1×10^{-4}
Activation function	Hyperbolic tangent	Hyperbolic tangent	Logistic sigmoid	Logistic sigmoid

[a] Limited-memory BFGS [34].
[b] Stochastic gradient descent with adaptive learning rate.

III. Preliminary Results and Discussion

The accuracy scores of all datasets are shown in Table 5. All methods applied to all datasets produce high accuracy scores. The highest accuracy is obtained using training set 4, with all methods reaching over 90 % accuracy on test data. The classification results of training set 4 on test data using the three methods are shown in Table 6. Training set 4, that includes only 2 days out of 4 and has the lowest number of healthy subjects, gave the best results in terms of classification on data that was not part of the training set. Moreover, all data used for testing is from different days than the measurements used for training. The best network architecture for this training set always had a single hidden layer. Method 3 applied to training set 1, however, gives similar results using a different hidden layer configuration and a different activation function for the hidden layer neurons, and this may indicate that the decision boundaries of the problem are relatively smooth.

TABLE 5
CLASSIFICATION RESULTS ON ALL TRAINING SETS

	Accuracy of the best networks on test data		
	Method 1	***Method 2***	***Method 3***
Training set 1	85.1 %	88.0 %	90.2 %
Training set 2	85.1 %	86.3 %	88.1 %
Training set 3	85.9 %	89.6 %	83.9 %
Training set 4	90.8 %	90.6 %	90.1 %

TABLE 6
CLASSIFICATION RESULTS ON TEST DATA USING TRAINING SET 4

		Score			
		Precision	***Recall***	***F1-score***	***Accuracy***
Method 1	**Healthy**	88.5 %	89.2 %	88.8 %	90.8 %
	AD	91.9 %	92.5 %	92.2 %	
Method 2	**Healthy**	88.2 %	86.5 %	87.3 %	90.6 %
	AD	92.0 %	93.1 %	92.5 %	
Method 3	**Healthy**	89.6 %	83.2 %	86.3 %	90.1 %
	AD	90.4 %	94.2 %	92.3 %	

IV. Conclusion and Perspectives

In this paper, we presented a novel approach to early non-invasive AD detection using ML classification of microwave sensing data. In particular, the binary classification (presence of the pathology or not) is based on difference in the measured scattering parameters due to the lower permittivity of pathological CSF. We gathered measurements using a realistic phantom and we trained several classifiers. The best network obtained in terms of accuracy on test data classification was a multilayer perceptron with a single hidden layer, trained using data collected during two of the four days of measurements and it reached a classification accuracy of 90.8 % when given as input all data belonging to the remaining two days.

One of the next steps will be improving the measurements quality. This improvement can be achieved in many ways, for example increasing the number of antennas and increasing the number of performed measurement sets. With a larger data pool, it is also possible to try a multiclass classification to distinguish the AD severity. To improve ML we can implement a different algorithm such as a convolutional neural network and we can also enlarge the search space of hyperparameters during the tuning phase.

Acknowledgments

This work was supported in part by the project PON Research and Innovation "Microwave Imaging and Detection powered by Artificial Intelligence for Medical and Industrial Applications (DM 1062/21)", funded by 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 project "THERAD - Microwave Theranostics for Alzheimer's Disease", funded by Compagnia di San Paolo. It was carried out partially within the Agritech National Research Center, funded by the European Union Next-Generation EU (Piano Nazionale di Ripresa e Resilienza (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022).

References

[1] C. Greenblat, "World Health Organization | Dementia," WHO. https://www.who.int/news-room/fact-sheets/detail/dementia (accessed Apr. 26, 2023)

[2] As Association, "Alzheimer's disease facts and figures," *Alzheimer's & Dementia* vol. 15, n. 3, 2019, pp. 321-387.

[3] K. A. Jellinger, B. Janetzky, J. Attems, E. Kienzl, "Biomarkers for early diagnosis of Alzheimer disease:'ALZheimer ASsociated gene'–a new blood biomarker?" Journal of cellular and molecular medicine, vol. 12, n. 4, 2008, pp. 1094–1117.

[4] C. R. Harrington, "The molecular pathology of Alzheimer's disease," Neuroimaging Clinics North Amer., vol. 22, n. 1, 2012, pp. 11–22.

[5] B. Dubois et al., "Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS–ADRDA criteria," The Lancet Neurology, vol. 6, n. 8, 2007, pp. 734–746.

[6] J. Dumurgier, S. Schraen, A. Gabelle et al., "Cerebrospinal fluid amyloid-β 42/40 ratio in clinical setting of memory centers: a multicentric study," Alzheimer's research & therapy, vol. 7, n. 30, 2015, pp.1-9.

[7] M. Shoji et al., "The levels of cerebrospinal fluid Aβ40 and Aβ42 (43) are regulated age-dependently," Neurobiology of aging, vol. 22, n. 2, 2001, pp. 209-215.

[8] L. M. Shaw et al., "Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects," Annals of neurology, vol. 65, n. 4, 2009, pp. 403-413.

[9] S. Jesse, J. Brettschneider, S. D. Süssmuth et al. "Summary of cerebrospinal fluid routine parameters in neurodegenerative diseases," Journal of neurology, vol. 258 (2011): 1034-1041.

[10] M. Nutu, H. Zetterberg, E. Londos et al., "Evaluation of the cerebrospinal fluid amyloid-β1-42/amyloid-β1-40 ratio measured by alpha-LISA to distinguish Alzheimer's disease from other dementia

disorders," Dementia and geriatric cognitive disorders, vol. 36, n. 1-2, 2013, pp. 99-110.

[11] P. Lewczuk, N. Lelental, P. Spitzer, J. M. Maler, J. Kornhuber, "Amyloid-β 42/40 cerebrospinal fluid concentration ratio in the diagnostics of Alzheimer's disease: validation of two novel assays," Journal of Alzheimer's disease, vol. 43, n. 1, 2015, pp. 183-191.

[12] T. Yoshikawa, Z. Zhang, K. Yamashita, M. Noda, "Biosensing of interaction between phospholipid membrane of liposome as model cell membrane and amyloid-beta protein in human serum by dielectric dispersion analysis," Sensors and Actuators B: Chemical, vol. 236, 2016, pp. 1028–1035.

[13] M. Manoufali, A. T. Mobashsher, B. Mohammed, K. Bialkowski, P. C. Mills, A. Abbosh, "Implantable sensor for detecting changes in the loss tangent of cerebrospinal fluid," IEEE transactions on biomedical circuits and systems, vol. 14, n. 3, 2020, pp. 452-462.

[14] D. O. Rodriguez-Duarte, J. A. Tobon Vasquez, R. Scapaticci, G. Turvani, M. Cavagnaro, M. Casu, L. Crocco, F. Vipiana, "Experimental Validation of a Microwave System for Brain Stroke 3-D Imaging", Diagnostics, vol. 11, n. 7, 2021, pp. 1232.

[15] J. A. Tobon Vasquez, R. Scapaticci, G. Turvani, G. Bellizzi, D. O. Rodriguez-Duarte, N. Joachimowicz, B. Duchêne, E. Tedeschi, M. R. Casu, L. Crocco, F. Vipiana, "A Prototype Microwave System for 3D Brain Stroke Imaging," Sensors, vol. 20, n. 9, 2020, pp. 2607

[16] A. Fedeli, V. Schenone, A. Randazzo, M. Pastorino, T. Henriksson, S. Semenov, "Nonlinear S-parameters inversion for stroke imaging," IEEE Transactions on Microwave Theory and Techniques, vol. 69, n. 3, 2020, pp. 1760-1771.

[17] A. Fhager, S. Candefjord, M. Elam, and M. Persson, "Microwave diagnostics ahead: Saving time and the lives of trauma and stroke patients," IEEE Microw. Mag., vol. 19, no. 3, 2018, pp. 78–90.

[18] J. A. Tobon Vasquez, R. Scapaticci, G. Turvani, M. Ricci, L. Farina, A. Litman, M. R. Casu, L. Crocco, F. Vipiana, "Noninvasive inline food inspection via microwave imaging technology: An application example in the food industry," IEEE Antennas and Propagation Magazine, vol. 62, n. 5, 2020, pp. 18-32.

[19] M. Ricci, B. Stitic, L. Urbinati, G. Di Guglielmo, J. A. Tobon Vasquez, L. P. Carloni, F. Vipiana, M. R. Casu, "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, Sept. 2021, pp. 503 – 514

[20] I. M. Saied, T. Arslan, S. Chandran, C. Smith, T. Spires-Jones, S. Pal, "Non-invasive RF technique for detecting different stages of Alzheimer's disease and imaging beta-amyloid plaques and tau tangles in the brain," IEEE Transactions on Medical Imaging, vol. 39, n. 12, 2020, pp. 4060-4070.

[21] I. M. Saied, T. Arslan, S. Chandran, "Classification of Alzheimer's Disease Using RF Signals and Machine Learning," IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology, vol. 6, n. 1, 2021, pp. 77-85.

[22] C. Origlia, M. Gugliermino, D. O. Rodriguez-duarte, J. A. Tobon Vasquez, F. Vipiana, "Anthropomorphic Multi-tissue Head Phantom for Microwave Imaging Devices Testing," 17th European Conference on Antennas and Propagation (EuCAP), Florence, Mar. 2023.

[23] N. Joachimowicz, B. Duchêne, C. Conessa, O. Meyer, "Anthropomorphic breast and head phantoms for microwave imaging," Diagnostics, vol. 8, n. 4, 2018, pp. 85.

[24] D. Andreuccetti, R. Fossi and C. Petrucci, "An Internet resource for the calculation of the dielectric properties of body tissues in the frequency range 10 Hz - 100 GHz," IFAC-CNR, Florence, 1997. [Online]. Available: http://niremf.ifac.cnr.it/tissprop/ (accessed May 12, 2023).

[25] Keysight Technol., "N1500A Materials Measurement Suite," keysight.com, 2022, [Online]. Available: https://www.keysight.com/it/en/product/N1500A/materials-measurement-suite.html, (accessed May 12, 2023).

[26] Keysight Technol., "Keysight streamline series USB vector network analyzer P937XA 2-port, up to 26.5 GHz," datasheet, 2018, [Online]. Available: https://www.keysight.com/it/en/assets/7018-06033/data-sheets/5992-2765.pdf, (accessed May 12, 2023).

[27] D. O. Rodriguez-Duarte, C. Origlia, J. A. Tobon Vasquez, R. Scapaticci, L. Crocco, F. Vipiana, "Experimental assessment of real-time brain stroke monitoring via a microwave imaging scanner," IEEE Open Journal of Antennas and Propagation, vol. 3, 2022, pp. 824-835.

[28] P. J. Werbos, "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences," Ph.D. dissertation, Harvard University, Cambridge, MA, USA, 1974.

[29] D. E. Rumelhart, G. E. Hinton, R. J. Williams, "Learning representations by back-propagating errors," Nature, vol. 323, no. 6088, 1986, pp. 533-536.

[30] The MathWorks Inc., "Deep Learning Toolbox Documentation," 2023, mathworks.com. [Online]. Available: https://www.mathworks.com/help/deeplearning/ (accessed May 12, 2023).

[31] I. Goodfellow, Y. Bengio, A. Courville, "Deep learning," The MIT press, 2016.

[32] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," The Journal of Machine Learning Research, vol. 12, n. 85, 2011, pp. 2825-2830.

[33] R. H. Byrd, P. Lu, J. Nocedal, C. Zhu, "A limited memory algorithm for bound constrained optimization," SIAM Journal on scientific computing, vol. 16, n. 5, 1995, pp. 1190-1208.

[34] D. P. Kingma, J. Ba, "Adam: A method for stochastic optimization," 3rd International Conference for Learning Representations, San Diego, 2015.