A Machine Learning-Based Classification Method for Monitoring Alzheimer's Disease Using Electromagnetic Radar Data

A Machine Learning-Based Classification Method for Monitoring Alzheimer's Disease Using Electromagnetic Radar Data

Rahmat Ullah, Yinhuan Dong, *Graduate Student Member, IEEE,* Tughrul Arslan, *Senior Member, IEEE,* and Siddharthan Chandran

***Abstract*—Alzheimer's and Parkinson's disease are two neurodegenerative brain disorders affecting more than 50 million people globally. Early diagnosis and appropriate assessment of disease progression are critical for treatment and improving patient's health. Currently, the diagnosis of these neurodegenerative diseases is based primarily on mental status exams and neuroimaging scans, which are costly, time-consuming, and sometimes erroneous. A novel, cost-effective, and precise diagnostic tools and techniques are thus urgently required, particularly for early detection and prediction. In the recent decade, electromagnetic imaging has evolved as a cost-effective and non-invasive alternative approach for studying brain diseases. These studies focus on wearable and portable devices and imaging algorithms. However, microwave imaging can not detect minimal changes in the brain at early stages accurately due to its lower resolution. This paper investigates machine learning techniques for the early diagnosis of acute neurological diseases, especially Alzheimer's disease. A machine-learning-based classification method is proposed. Simulations are performed on realistic numerical brain phantoms using the CST studio suite to get the scattered signals. A novel data augmentation method is proposed to generate synthetic data required for machine learning algorithms. A deep neural network-based autoencoder extract features to train various machine learning algorithms. The classification results are compared with raw data and manual feature extraction. The study shows that the proposed machine learning-based method could be used to monitor Alzheimer's disease at its early stages.**

***Index Terms*—Alzheimer's disease (AD), classification, data augmentation, microwave sensing, machine learning, radar data.**

I. Introduction

ALZHEIMER'S disease (AD) induces physiological and pathological changes in the human brain, including brain atrophy, lateral ventricle enlargement, increase in the Cerebrospinal fluid (CSF), and the accumulation of plaques and tangles. Current approaches for monitoring neurodegenerative diseases, such as magnetic resonance imaging (MRI) and positron emission tomography (PET), are bulky, costly, and uncomfortable for elderly patients [1]. Microwave-based image reconstruction techniques are getting attention to be used as an alternative imaging modality. The advantages over traditional imaging modalities include their portable and wearable nature. Microwave-based imaging techniques are being investigated for breast cancer and brain stroke detection [2]. These techniques have recently been investigated to detect AD-related pathological and physiological changes in the brain [3]. The imaging-based studies can detect significant pathological and physiological changes in the brain. However, it is essential to detect the disease at its early stages. Due to their low spatial resolution, small pathological and physiological changes can not be detected using microwave imaging techniques.

AD can be detected at the early stages by observing CSF and brain atrophy changes even before the onset of symptoms. Current machine and deep learning methods use images from different imaging modalities such as MRI and PET scans. The algorithms are trained using various features such as brain volume, grey matter densities, and cerebral amyloid accumulation in the hippocampus [4]. All these approaches need images from imaging modalities. Recently, a novel, lightweight, non-invasive microwave imaging device for diagnosing neurodegenerative diseases has been proposed. [5] However, the device has low Noise to Signal Ratio, which leads to low-resolution images. This is due to the radiation pattern of a wideband antenna that leads to spatial blurring and signal scattering in tissues. More sophisticated signal processing techniques are needed for accurate monitoring or detection of the early stages of AD.

Machine learning (ML) has already shown potential in neuroimaging data analysis for various diseases [6] and applications for classifying stroke types[7] [8]. Different classification methods are proposed to distinguish intracerebral haemorrhage (ICH) from ischemic stroke (IS) and healthy volunteers by classifying high-frequency microwave data [9][8]. However, the potential of ML in predicting AD using microwave data has yet to be thoroughly exploited. The data obtained from pathological phantoms using simulations were used to classify different stages of AD [10]. The paper utilises ML to analyse the RF data (reflection coefficients) simulated in CST to train the model for AD progression tracking. The reflection coefficients from all antennas were integrated and treated as one sample, which increases pre-processing overhead. Our previous work used the raw data from both physiological

This paper is an expanded paper from "Big Data-Machine Learning Processing of Recorded Radiofrequency Physiological and Pathological Measurements to Predict the Progression of Alzheimer's Disease", published in the 2021 IEEE Asia-Pacific Microwave Conference (APMC) held on 28 Nov.-1 Dec. 2021 in Brisbane, Australia.

Rahmat Ullah, Yinhuan Dong and Tugrul Arslan are with the School of Engineering, University of Edinburgh, Edinburgh EH8 9YL, U.K. (e-mail: rahmat.orakzai@ed.ac.uk; yinhuan.dong@ed.ac.uk; tughrul.arslan@ed.ac.uk;).

Siddharthan Chandran is with the Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh EH9 3JW, Scotland (e-mail:siddharthan.chandran@ed.ac.uk)

and pathological phantoms to train various ML algorithms [11]. The presented results show a reasonable accuracy of the proposed method. However, due to the limited and unrealistic data used in training, the accuracy is questionable for a more accurate prediction. In this work, first, data from both pathological and physiological changes are used. In our previous work, multiple phantoms were created based on the dielectric properties obtained from postmortem tissues [12]. Multiple compositions of mesh cells for each phantom are used to accommodate various head sizes. Furthermore, a new data augmentation technique is proposed to overcome the scarcity of data. A more realistic form of augmented data is generated, and an integrated data set is created. Furthermore, previous studies use raw data directly without feature extraction, which is susceptible to redundant data. This paper uses a neural network-based feature extraction technique to extract various features from the integrated data set.

Conversely, a number of works based on deep learning have already been introduced intensively in inverse scattering[13] [14] [15]. For example, a deep learning network-based system for mapping a scattered field of objects to high-resolution images is presented in [16]. A two-stage training strategy is used to lower the training difficulty. In the first step, an auto-encoder consisting of an encoder and a decoder is constructed to identify compact representations for high-resolution images. In the second step, a neural network maps microwave signals to the extracted feature. Similarly, a deep learning approach based on the deep convolutional encoder-decoder structure is proposed to solve the electromagnetic inverse scattering problem [17]. The model is built using a complex-valued deep convolutional neural network consisting of an encoder and decoder network. The proposed network rebuilds the permittivity of dielectric objects based on scattering data. A direct inversion scheme based on a neural network is proposed for the quantitative imaging of highly nonlinear profiles [18]. A neural network with complete connectivity is deployed in the training phase to estimate the object's complex permittivity profile under test. Similarly, deep learning is proposed to classify breast tomographic microwave and ultrasound tissue types. The method employs a CNN with U-net architecture to assess pixel categorization uncertainty. The training set consists of quantitative tomographic reconstructions of dielectric properties and ultrasonic properties, along with tissue-type classification. The proposed deep learning-based methods provide promising results. However, these techniques are computationally expensive [19].

In microwave imaging, learning approaches have also been used to solve the inverse scattering problem. Artificial neural networks [14] and support vector machines [20] were previously used to extract different parameters from microwave scattering signals. These deep learning methods are divided into three categories; the direct learning method, the learning-assisted objective function approach, and the physics-assisted learning approach. The direct inversion scheme presented in [21] is part of the direct learning strategy that uses CNN as a "black box" and directly calculates the relative permittivity from scattered signals. The direct learning method is straightforward but typically results in low image quality. The learning-assisted objective-function approach uses networks to learn some components of conventional iterative solvers, such as the descent learning methodology [22], the learning-assisted multimodality method [23], and the deep learning embedding method [24]. Finally, physics-assisted learning methodologies, such as ICLM [25] DeepNIS [15], physics embedded deep learning method [26], incorporate physical knowledge into either the inputs or internal network structures.

A machine learning-based approach was proposed for breast lesion detection based on clinical data in [27]. The microwave scattering signal obtained using a moving transmitting and receiving antenna are used. Traditional radiologist research on the same participants is used to pre-process microwave scattering data. The labelled data are used to train and analyse supervised machine learning methods, including nearest neighbour, multi-layer neural network, and support vector machine (SVM). According to statistical analyses, the SVM can recognise breast data with 98% accuracy. In the study, a total of 36 frequency samples are used, which leads to overfitting. Detecting different stages of AD is more complex than breast cancer and brain stroke detection due to the complicated brain structure, and slight differences in the dielectric properties of adjacent brain tissue layers [28]. Microwave imaging methods that are successful in breast cancer detection can not be used for brain imaging due to the heterogeneous nature of brain tissues [8]. Therefore, the microwave detection method based on feature extraction and classification can potentially be a promising alternative solution for monitoring AD.

There are several challenges in classification methods for microwave-based imaging. The main challenge is the scarcity of data, as these systems are still being investigated, and only a few systems are clinically used. Most studies rely on simulations or experimental data obtained from numerical or fabricated phantoms. In a recent study, the scarcity of data was addressed by employing a graph-based method [7]. A total of 256 reflected and received signals are collected from each phantom. Each collected signal is converted to a graph, and the relationship between each pair of graphs is established by calculating mutual information. The results of mutual information between graphs are sent to a support vector machine to determine the stroke subtype. The results demonstrated that the accuracy of detecting ICH from IS and vice versa is 84%; however, the proposed method is compute-intensive. Second, microwave imaging systems require an antenna array usually erected around the brain to acquire scattered signals. A switching circuit is used to change the signal source. Having dozens of antennas and cables adds weight to the system and generates noise due to antenna couplings, cable movement, and antenna fabrication inconsistencies. Third, microwave signals contain many dimensions, leading to issues such as redundant data unrelated to the intrinsic characteristics and data singularity [29]. All these challenges contribute to the non-effectiveness of ML-based classification methods. The scattered signal from the target is submerged in the noise, making it difficult to detect due to the subtle difference between the object under test and surrounding tissues. Therefore extensive and integrated scattered data, using less number of antennas and effective feature learning, is required for more accurate classification.

This paper investigates ML-based methods to detect these small changes that will help monitor different stages of AD using radar-based electromagnetic sensor data. The first and most crucial step for any ML-based method is to have a reliable data source. Since ML techniques are being investigated, getting a huge amount of data that could be used to train ML models is challenging. A novel data augmentation method is proposed to generate synthetic data that solve the data scarcity problem. The proposed ML-based classification method requires fewer antennas to alleviate the challenge of constructing an ultra-wideband (UWB) antenna array used by microwave imaging systems [8]. The paper has the following main contributions.

- More realistic brain phantoms having multiple layers such as skin, skull, CSF, grey matter and white matter are created. Multiple phantoms with different compositions of mesh cells are created in the CST studio suite to accommodate different head sizes. Simulations are performed to generate data from numerical phantoms mimicking different stages of AD. These phantoms mimic the brain's physiological and pathological changes associated with various AD stages.
- A novel data augmentation method based on statistical measurements is proposed to generate synthetic data for both physiological and pathological phantoms, producing an integrated data set.
- Two feature extraction strategies have been stipulated and compared. This includes the statistical features (e.g. mean, standard deviation) manually extracted from the samples and the latent features automatically extracted by the developed deep neural network-based Stacked Auto-encoder (SAE).
- Multiple ML techniques are trained using raw data manually extracted features, and automatically extracted features. The results are evaluated for accuracy and compared with raw data and manually extracted features.

The results indicated that the proposed method significantly increases accuracy compared to raw data and manually extracted features from the RF scattered data. Moreover, unlike the previous studies, [11][10], the proposed method eliminates the need for integrating data from multiple antennas for each case. Instead, the data obtained from a single antenna could be used to monitor different stages of AD.

The paper is organised as follows. Section II contains the proposed ML-based classification method. The Imaging system setup used to collect data, data processing, including data augmentation and prepossessing techniques, feature extraction and classification algorithms, is described. Section III contains results and analysis, including the dataset, feature extraction and classification accuracy. Section IV includes a detailed discussion and comparisons of different algorithms and their accuracy and suitability for monitoring AD. Finally, a summary and future work are presented in Section V.

II. Proposed ML-based Classification Method

The proposed ML-based classification method consists of four major steps, as shown in Figure 1. Each of these steps has sub-steps, from data acquisition to classification. This section describes these steps in detail.

A. Data Acquisition

Data collection is a significant problem for any ML problem. Since machine and deep learning algorithms require a large amount of data for better results in terms of accuracy, therefore using PET and CT images to train models is easy to find. However, getting microwave-scattered data is challenging since the technology is still in the testing and evaluation phase. In this paper, simulations are performed using realistic human brain phantoms that mimic the brain's physiological and pathological changes due to AD. CST Microwave Studio Suite was used to develop these simulation models. CST Studio Suite is a specialised tool for three-dimensional electromagnetic simulation. Various brain phantoms are created to mimic different stages of AD, mainly physiological (brain atrophy) and pathological changes (caused by beta-amyloid plaques and tau tangles) in the human brain. Multiple variations of mesh cells are used to accommodate different head sizes.

1) Phantoms Used: Phantoms are often created for human tissue imaging, such as the breast and the brain, using substances that mimic the surrounding tissue's properties and the object of interest, such as a tumour or stroke. It provides more consistent results, avoids exposure towards a living human, and is mainly used for experimental studies and ex-vivo validation of antennas and imaging algorithm results. In this work, multiple numerical phantoms are created to mimic physiological and pathological changes in the brain caused by AD. CST microwave studio is used to create these realistic human head phantoms. Phantoms with different compositions of mesh cells were first developed to accommodate various head sizes based on a realistic human head voxel model presented in [30]. These head models contained different layers of tissue, such as the skull, skin, blood, white matter, and grey matter. These layers have different dielectric properties that change with frequency. The frequency-dependent dielectric properties (conductivity and relative permittivity) of grey matter, white matter, and CSF can be found online [31]. Fig. 2 shows the conductivity and relative permittivity of grey matter, white matter and CSF from 0.5 GHz to 5 GHz.

Two different changes are reflected in these phantoms, i.e. physiological and pathological. For physiological models, the size of grey matter and white matter were uniformly reduced in the original head model to simulate brain atrophy. As a result of shrinking the size of these objects in the model, there is a resulting gap [32]. In reality, when brain atrophy occurs, the resulting gap due to the shrinkage of brain tissues is filled with CSF [33], [34], [35]. As a result, an additional CSF layer was added to the brain models to emulate the characteristics of brain atrophy more realistically. The other layers, such as the skin and skull, were kept constant, and changes were made only to the grey matter, white matter and CSF. The mild and severe case phantoms mimic different stages of AD in terms of brain atrophy levels and increase in CSF thickness. To imitate the mild case of the AD, a decrease in the white matter and grey matter from 2.23×10^6 mm^3 by 20% to 1.78×10^6 mm^3 was made. As mentioned earlier, the brain's area after neurodegeneration is accumulated with CSF. Likewise, the CSF layer volume was increased by 20%, from 1.79×10^5

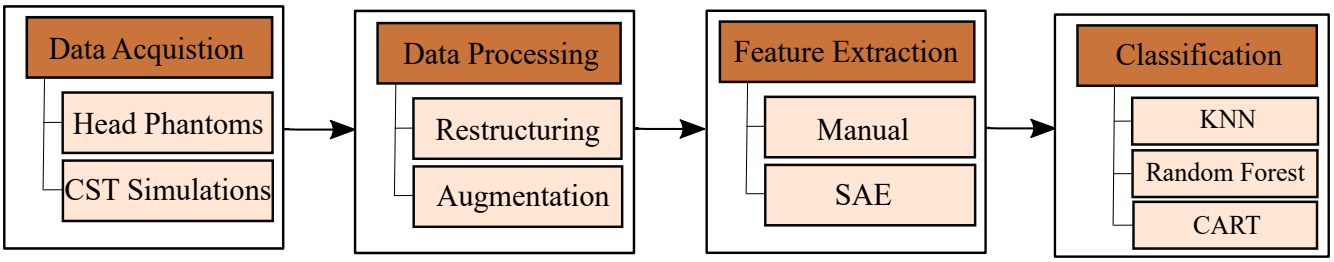


Fig. 1. The proposed classification method for monitoring different stages of AD.

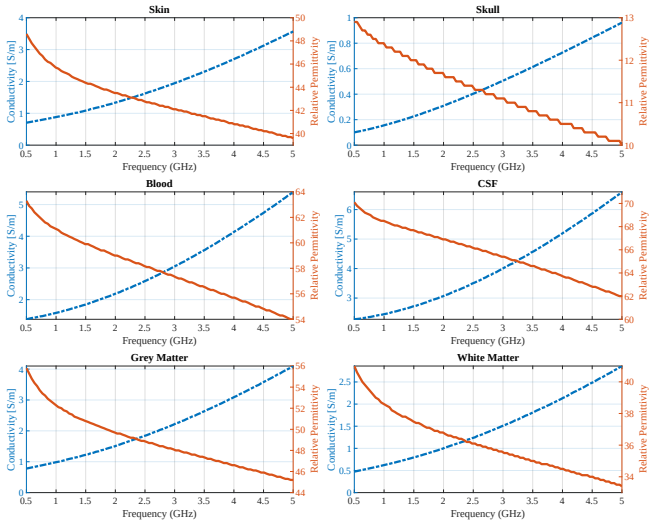


Fig. 2. Dielectric properties of major brain tissues model used for simulations.

mm^3 to 1.97×10^5 mm^3.

For pathological models, pathological changes need to be mimicked in the numerical phantoms. The dielectric characteristics of biological tissues influence how electromagnetic signals propagate, reflect, and are absorbed. Our previous study on postmortem tissue found a significant change in the dielectric properties in healthy and AD-affected brains [12]. The study measured the relative permittivity and conductivity of brain tissue samples with the severe form of AD and compared them to those of healthy brain tissue. Two main categories of tissue samples were considered: the brain's grey matter and white matter. The experiments were performed at frequencies from 20 MHz to 3 GHz on defrosted samples. The results indicated that the relative permittivity for the grey matter at 770 MHz in a healthy and AD-affected brain is 53.3 and 45.3, respectively. It showed that the relative permittivity of the AD-affected brain had decreased by 15%. Similarly, at 770 MHz, the conductivity of grey matter is 0.8906 S/m in the healthy brain and 1.6253 S/m in the AD-infected brain, indicating an increase of 82.5 percent. The impact of plaque and tangle on the white matter has similar outcomes. For example, at 770 MHz, the relative permittivity for white matter in healthy and AD-infected brains is 39.35 and 32.57, respectively, indicating a decrease of about 17.61%. Similarly, at 770 MHz, white matter conductivity is 0.554 S/m in the healthy brain and 0.702 S/m in the AD brain, showing a 26.72 percent increase in conductivity.

These findings show that monitoring changes in plaque and tangle in the early stages of AD may be a way to detect the disease. The accumulation of beta-amyloid plaques and tau tangles, which impede the transfer of nutrients within brain cells, is one of the significant changes in the brain during AD [36]. This leads to pathological changes in the brain tissue composition with the progression of AD. Different voxel models are developed based on the evolution of plaques and tangles in the brain to appropriately depict the brain regions affected by the spread of AD pathology.

For this study, three different stages are considered: 1) Healthy, 2) Mild AD, and 3) Severe AD. Fig. 3 shows the CST head models used to represent the different AD cases for both physiological and pathological changes in the brain [11]. The first row shows the models that reflect the brain's physiological changes, i.e. brain atrophy and CSF thickness. The second row shows models reflecting the pathological changes caused by the accumulation of plaques and tangles in the brain as AD progresses. A healthy human brain model was created in CST to compare the results with AD (mild and severe) brain models. This detailed phantom creation, while time-consuming, enables the investigation of both physiological and pathological changes associated with different stages of AD.

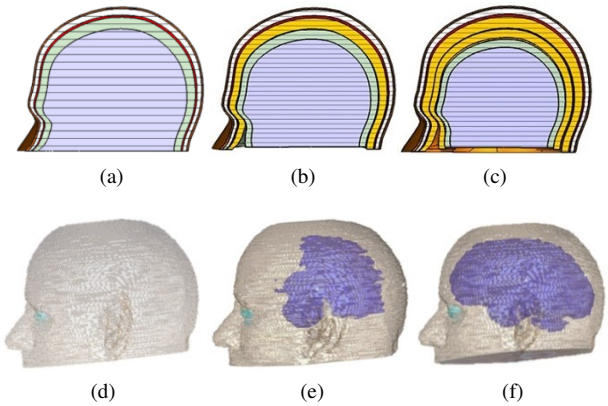


Fig. 3. CST head models used to simulate the physiological and pathological changes in the brain [11] (a) Healthy, (b) Mild AD (10% brain atrophy), (c) Severe AD (25% brain atrophy), (d) Healthy model, (e) Mild model for pathological changes and (f) Severe model for pathological changes.

2) Antenna Used: An array of six monopole directional antennas is used. These antennas are evenly spaced around

the models to cover the head. The antenna design is based on a rectangular planar monopole structure. Far-field simulations were used to calculate the front-to-back (FTB) ratio, ensuring the antenna's directionality. The simulated FTB ratio of the utilised antenna is 6 dB at its central frequency of 2.5 GHz, which validates its directional nature. The lower frequency range allows more penetration into the bio phantom with the loss of resolution. While at higher frequencies, better resolution can be achieved with less penetration of the signals into the phantom. The UWB performance of the antenna allows good penetration and resolution for the radar-based imaging application. Furthermore, the E-field distribution reveals a good penetration depth into the head. This is necessary for the antenna to probe into the head and detect minor changes. Further information about the antennas used can be found in [37].

Several simulation cases are performed to investigate the effects of the head phantom on the antenna performance. Fig. 4 shows the S_{11} simulation results of four cases: the absence of the head phantom, with a head phantom at a distance of 0 mm, 5 mm, and 10 mm from the antenna, respectively. The simulation result indicates a shift in operating frequency when a head phantom is added. The distance between the antenna and the bio phantom is chosen at 0 mm to allow better signal penetration at a lower frequency range.

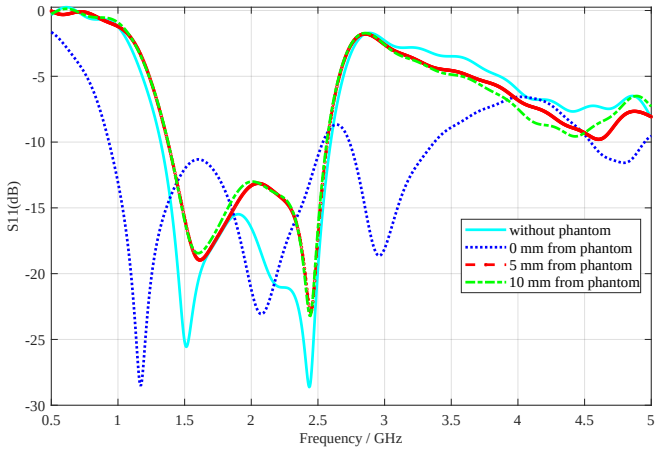


Fig. 4. The S_{11} plot for the designed antenna for different separation distances between the antenna and skull model with and without phantom. Smaller separation distance leads to better penetration.

3) Data Collection: These phantoms were used in electromagnetic simulations. For each case, simulation models were run with multiple compositions of mesh cells to accommodate various head sizes. The simulations are performed in a multi-static manner, where the microwave signal is sent by each antenna, one at a time, and all other antennas capture the scattered signals. This procedure is carried out for each antenna until the required signal samples for all antennas are collected. The reflection and transmission coefficient data were stored. However, only the reflection coefficient data is used in the proposed method. A model along with six surrounding antennas can be seen in Fig. 5.

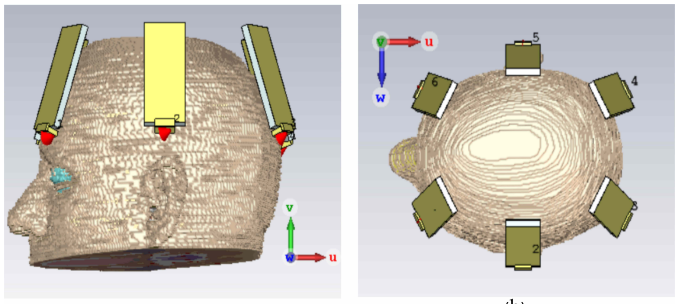


Fig. 5. (a) Front view and (b) Top View of Brain phantom and six surrounding antennas.

B. Data Processing

Simulations were performed to capture the reflected signals (S_{11}) for all antennas and all cases. These reflected signals are the base for the proposed method to differentiate and classify stages of AD. The raw data for healthy, mild and severe cases of both physiological and pathological changes can be seen in Fig. 6. The figure illustrates the S-parameters of a single antenna when measuring a specific phantom that has been designed to mimic three different stages of AD; healthy, mild, and severe. It is important to note that the size of the phantom remains constant across all three versions. However, the internal composition of the different tissues within the phantom is modified to mimic the different stages of AD. Additionally, the S-parameters are dependent on the location of the antenna within the phantom. Thus, the same antenna location is used for all three versions of the phantom. This allows for a fair comparison of the S-parameters between the different stages of AD at that specific location within the phantom. To mimic the pathological changes in brain tissue, a different set of phantoms are used. It is important to note that the S-parameters are dependent on the internal composition of the tissues and not just the size of the phantom.

The figures indicate a clear difference between different stages for both physiological and pathological phantoms. For Physiological changes, a shift in signal and drastic drop in the magnitude from healthy to mild and severe cases is observed and can be seen in the figure. This showed that there is a target with different sizes in the skull model. For Pathological changes, It was found that between 3 to 5 GHz, the electromagnetic wave reaches the region associated with plaques and tangles in the brain. Furthermore, significant variations among the reflected signals can be seen due to the change in the dielectric properties. This raw data obtained from simulations of multiple phantoms were pre-processed before data augmentation.

1) Data Restructuring: A total number of 201 frequency points were stored in complex form for each antenna. This data was first converted to absolute form. Each sample consists of 1206 discrete points (201 for each antenna), as shown in Table I. Nine different compositions of mesh cell phantoms for each stage (healthy, mild and severe) are used to accommodate multiple head sizes and dimensions. This gives 9 samples for each of the three stages. The final data for each stage

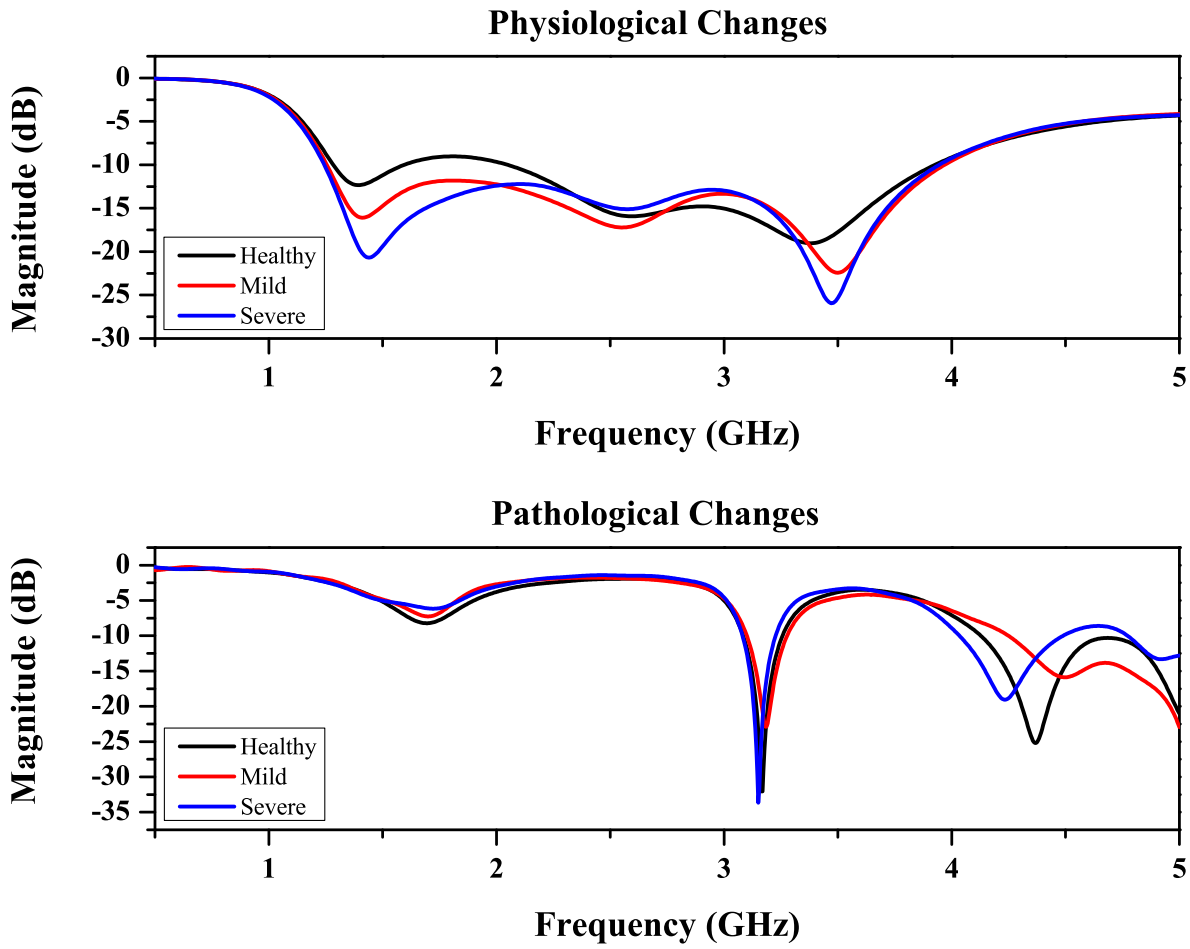


Fig. 6. Comparison of data for healthy, mild and severe stages of data for both physiological and pathological changes.

consists of 9×1206 points. Based on the antenna's location, this data was restructured to 54×201 points for each stage. This data is merged together, and the final data consists of 162×201 data points. It is essential to restructure the data before data augmentation to ensure that augmented values are created based on antenna locations around the head. The data for pathological models are restructured in the same way.

TABLE I
RAW DATA RESTRUCTURED FOR DATA AUGMENTATION

	Raw Data	Composition of mesh cells	Restructured Data	Final Data
Healthy	1×1206	9×1206	54×201	162×201
Mild	1×1206	9×1206	54×201	
Severe	1×1206	9×1206	54×201	

2) Data Augmentation: Since ML algorithms require more data for better results in terms of accuracy, therefore using images to train models is easy to find. However, getting microwave-scattered data is challenging since the technology is still in the testing and evaluation phase. The experiments required phantoms and fabricated antennas. However, these phantoms could not be used for longer because their dielectric properties change over time. On the other hand, simulations are time-consuming, especially for multi-layer phantoms that contain a huge amount of mesh cells for more accurate brain representation.

Therefore, synthetic data was used to train different ML algorithms to classify AD stages based on pathological and physiological changes in the brain. Data augmentation is a technique for artificially increasing the size of a small training set by generating new samples out of the ones present on the actual training set [38]. This technique has been widely used for SAR imaging [39][40]. A new data augmentation method based on statistical learning is employed to create synthetic data for a hundred cases of each AD stage. The values of the reflected signal for individual antennas were statistically analysed to find the approximate difference between different levels of atrophy for physiological data. A similar technique was employed for pathological data.

The proposed data augmentation method consists of multiple steps, as shown in Fig. 7. First, the absolute values of obtained S parameters for each stage, i.e. healthy, mild and severe, and multiple variations of these cases are arranged into vectors. The second step is to restructure the data based on the location of antennas due to the homogeneous nature of brain tissue in all cases. In the third step, the data is divided into multiple subsets in the step of 5 values for each antenna

similar to [41]. For each subset of S parameters, maximum and minimum values among themselves are calculated for all antennas. A random number with a 10% difference is generated between this calculated maximum and minimum value. This new number is added as a new data point in the same column of the same output. The new vectors are reused to calculate the successive vectors, making them more comprehensive. This process is repeated for all cases and antennas to generate augmented data for a hundred cases.

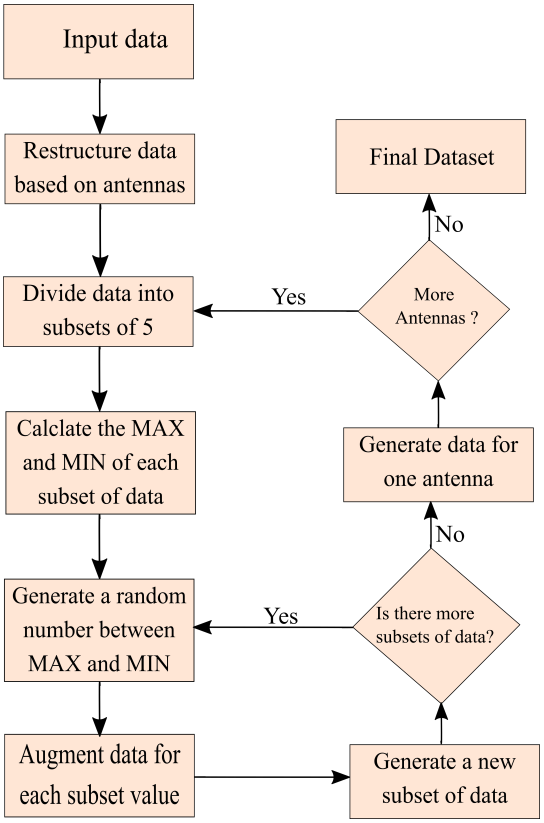


Fig. 7. The proposed data augmentation method to generate synthetic data for one case. The same method is used to generate data for 100 cases.

These steps augment data for hundred cases for each stage and feed it to the ML algorithms for training and testing. For the training and the performance assessment, 600 cases are generated. 80% were used as the training set and 20% as the test set.

C. Feature Extraction

1) Manual Feature Extraction: The manually extracted features are calculated from the statistical characteristics of the raw data (features). The mean, standard deviation, quartile deviation, range, skewness and kurtosis are calculated as the manually extracted features for each sample. The details can be found in Table II.

2) Automatic Feature Extraction: Auto-encoders are utilised to reduce the dimensionality of data when a non-linear function determines the relationship between dependent and independent features. Auto-encoders are a form of unsupervised artificial neural networks. Auto-encoders are employed to extract data features automatically. It is one of the most promising feature extraction technologies used for similar applications, such as human gesture detection [42]. The stacked auto-encoders are multiple encoders placed on top of one another. The microwave scattering dataset has very complex patterns; thus, a single auto-encoder cannot reduce the dimensions of the data. Therefore, a stacked auto-encoder is used. The stacked auto-encoder (SAE) is DNN-based feature extraction and dimension reduction method. It has been proven to be more powerful and effective than the traditional algorithm, such as Principal Component Analysis (PCA). The two parts compose its symmetrical structure: encoder and decoder, as shown in Figure 8.

TABLE II
INTRODUCTION OF THE MANUALLY EXTRACTED FEATURES

Symbol	Feature
μ	Mean
σ	Standard deviation
Q	Quartile deviation
R	Range
$\mathcal{S}$	Skewness
$\mathcal{K}$	Kurtosis

Fig. 8. The components and data flow of Stacked auto-encoder

The encoder aims to learn and extract features from the input data, while the decoder is required to reconstruct the encoder's output to be the same as the original input data. The final output of the decoder $\hat{x}$ can be expressed by:

$$\hat{\mathbf{x}} = g(f(\mathbf{x})) \tag{1}$$

where $f(\cdot)$ and $g(\cdot)$ denote the operation in encoder and decoder, respectively. To evaluate the difference between the output $\hat{x}$ and the input x, mean squared error (MSE) is applied as the cost function:

$$J = \frac{1}{n}\sum_{i=1}^{n}(\mathrm{x} - \hat{\mathrm{x}})^2 \tag{2}$$

The parameters in this network are then updated during the training session. A well-trained SAE is able to extract the features $\mathbf{y}$ from the input using the encoder:

$$\mathbf{y} = f_{\text{new}}(\mathbf{x}) \tag{3}$$

where $f_{\text{new}}(\cdot)$ is the operation of the well-trained encoder.

In this paper, we designed three SAEs with different structures for comparison. All SAEs are constructed of 5 dense layers (fully connected layers). As the structure of the SAE is shown in Figure9, from left to right, the structures of the SAEs are 128-64-32-64-128, 128-64-16-64-128 and 128-64-128, respectively. The original features of x are fed to the network as input. Once the training process finishes, the well-trained encoder will be disconnected from the network, and the automatically extracted four features will be used for

classification in the next step. More details about the network can be found in Table III.

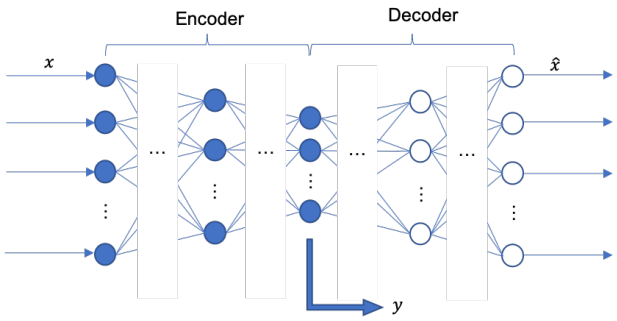


Fig. 9. Structure of the SAE for feature extraction.

TABLE III
DETAILS OF THE NEURAL NETWORK

Activation function	tanh
Batch size	4
Learning rate	0.0001
Epoch†	100
Early stopping patience	10
Optimizer	Adam
Loss function	mean squared error

† This is a pre-set epoch number. As an early stopping strategy is applied, the network usually stops training earlier.

D. Machine Learning-based Classification Algorithms

The choice of ML algorithm for classification is challenging because there are many ML-based classifiers, each with its learning method. There is no one-size-fits-all solution. These algorithms need to be tested before they can be used to solve a classification problem. However, algorithm selection is influenced by the size and complexity of the problem, the type of data, the expected results and how the results will be used.

For the classification of different stages of AD, based on microwave scattering data, the following algorithms are exploited. The multi-class classification will require ML algorithms that classify one or more class labels. In multi-classification problems, each training point is classified into one of the N classes. The goal is to create a function that will accurately forecast the class to which a new data point belongs when given a new data point.

1) KNN (K-nearest neighbours): The KNN algorithm is a supervised ML technique that can address classification and regression problems. It is one of the most basic yet effective algorithms. It memorises the training data rather than learning a discriminative function from the input.

The K-nearest neighbour approach uses a distance measure between the two data points to define them as similar and then produces a majority vote between the K most similar instances. The most popular choice is Euclidean distance, which is expressed as:

$$d(u,v) = \sqrt{\left((u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2\right)} \tag{4}$$

The hyperparameter K in KNN needs to be chosen carefully to acquire the best fit for the dataset. For the minimum value for K, i.e. K=1, the model will have low bias but significant variance due to overfitting. A higher value of K, such as K=10, will undoubtedly smooth the decision border, resulting in low variance but significant bias. As a result, the bias-variance trade-off, always exists.

KNN is a nonparametric classifier since it makes no assumptions about the distribution of classes. It is one of the multi-class classification algorithms that can be widely used.

2) Random Forest: A random forest comprises numerous separate decision trees that act as an ensemble. The random forest generates a class prediction for each tree, and the class with the most votes becomes the model's prediction. It produces a "forest" from a collection of decision trees that are often trained using the "bagging" method. The main idea of the bagging method is that combining different learning models enhances overall output. The random forest combines boosting and bagging, resulting in a neither overfitted nor inefficient model. A random forest's hyperparameters are similar to a decision tree or a bagging classifier. The random forest adds more randomness to the model; as the trees grow. Instead of looking for the important feature when dividing a node, it seeks the best feature from a random group of features. As a result, the approach only considers a random subset of the features when dividing a node in a random forest.

The random sampling strategy is used for feature selection for classifying the RF signals. Assuming the training set comprises total F features, the potential split feature set is $f = log_2 F + 1$. The Gini index [43] is a widely used metric for determining the best split feature. The Gini index aims to allocate balanced examples to the same class within each partition. Let D stand for the training set with m different labels $C_i (i = 1, 2, ..., m)$, $|D|$ for the data set size, and $|C_i, D|$ for the number of labels C_i in D. Eq 5 may then be used to calculate the Gini index.

$$\text{Gini}(D) = 1 - \sum_{i=1}^{m} p_i^2 \tag{5}$$

Where p_i is the probability that instances in D belong to class C_i is . If D is divided into D_1 and D_2 by a binary partition divided by feature A, the relevant Gini index is obtained using the following eq:

$$\text{Gini}_A(D) = \frac{|D_1|}{|D|}\text{Gini}(D_1) + \frac{|D_2|}{|D|}\text{Gini}(D_2) \tag{6}$$

As a result, selecting the maximum Gini factor is the best way to identify the best partition features. The majority win was used. All the classifier's votes were added to the basic random forest classifier, and the highest vote signifies the highest chance that the test data sample belongs to this class.

3) Classification And Regression Tree (CART): CART is a nonparametric statistics approach that assumes the decision tree as a binary tree. The anticipated values are determined by dividing the input feature space into finite units, i.e. the predicted output values are mapped to the provided circumstances. The CART contains three primary steps. A decision tree is created using the training data set in the first step. In the second step, the decision tree is pruned according to some constraints, such as the maximum depth of the tree, minimum sample number of the leaf node, and minimum node impurity for optimisation. In the last step, the data is predicted in the test set. The advantages of CART include its ability to deal with both continuous and categorical data simultaneously and handle multi-class classification problems.

Several methods can be used in CART to identify the best splits. The Gini impurity method, similar to the random forest, is used. The weighted sum of Gini Impurity is calculated for both child nodes to get the optimal split. This is repeated for all potential divides, with the best split being the one with the lowest Gini Impurity.

III. Results and Analysis

The ML algorithms were performed on the dataset using a Python environment. SciPy, which is an ecosystem of Python libraries such as NumPy (work with data in arrays), Matplotlib (generate 2D graphs), and Pandas (tools to organise and analyse data), were used. In addition, the sci-kit-learn library was declared for ML model evaluation. The following ML algorithms were employed: 1) K-Nearest Neighbours (KNN), 2) Random forest, and 3) Classification and Regression Trees.

The final RF dataset was used to test each ML method. Some data from the dataset was kept back to ensure that the algorithms were not biased during their training phase. Later, statistical methods were used to estimate the models' accuracy on data that had never been seen before (or validation data). This method provides an independent estimate of the accuracy of each model.

The final RF dataset was divided into two sets; roughly 80% of the dataset was used to train the models (e.g. training dataset). The remaining 20% was used to create a validation dataset. The dataset was split using a random number generator in Python. A test harness was developed to measure the accuracy of the models on the training dataset using k-fold cross-validation. The model's skill is summarised with the help of an evaluation score. The resulting accuracy metric is then used to evaluate the models. The next stage is to test the best ML algorithm's performance on the validation dataset set aside to evaluate how well the chosen model can classify AD based on previously unknown data.

A. Dataset

The proposed data augmentation method generates realistic synthetic data for each AD stage. A comparison of the original and augmented data for a healthy brain (all six antennas) is shown in Fig. 10. It can be observed that the augmented data follows the distribution of original data for all antennas. This is due to the small subsets of data utilised during data augmentation. The graph for original data shows a few outliers due to the location of the antenna and different compositions of brain tissue on each side, for example, the front and back of the head. In contrast, the augmented data have more outliers due to the 10% difference in generating each new value within the subset. The value of 10% is chosen to accommodate more realistic scenarios as the current data is based on simulation. The value range in augmented data is slightly increased. This accommodates realistic scenarios where the experimental data may contain noise from other sources.

In order to ensure that the statistical properties of the original data set are captured and transferred to the augmented dataset, statistical measurements are performed. The statistical comparison based on the mean and standard deviation for each antenna of one case (healthy brain) can be found in Fig. 11. It can be seen that there is a subtle difference in both the mean and standard deviation in original and augmented data. As mentioned earlier, this difference is based on the 10% different in each subset of 5 values used for data augmentation. This comparison is useful to ensure that each antenna's augmented data follows the original data's underlying structure.

The Grubb's test is performed to ensure that the proposed method is not generating outliers. The Grubb's test statistics is defined as [44]

$$G = \frac{\max \left| Y_i - \bar{Y} \right|}{s} \tag{7}$$

with $\bar{Y}$ and s denote the sample mean and standard deviation, respectively. The Grubbs' test statistic measures the largest absolute deviation from the sample mean in units of the sample standard deviation. The hypothesis of no outliers is rejected if [45]

$$G > \frac{(N-1)}{\sqrt{N}} \sqrt{\frac{\left(t_{\alpha/(2N),N-2}\right)^2}{N-2+\left(t_{\alpha/(2N),N-2}\right)^2}} \tag{8}$$

with $t_{\alpha/(2N),N-2}$ denoting the critical value of the t distribution with $(N-2)$ degrees of freedom. The Origin software [46] was used to conduct the test on multiple samples of original and augmented data. The resulting value of G is 2.18, with a critical value of 3.61. These test results show that there are no significant outliers at the 0.05 level. Furthermore, the paired t-test is performed to test the mean difference between the original and augmented data pair. Calculating a t-test requires three fundamental data values that include the mean difference, the standard deviation of each group, and the number of data values of each group. The standard error of mean (SEM) measures the discrepancy between a sample mean to the population mean. A large t-score, or t-value, indicates that the groups are different, while a small t-score indicates that the groups are similar. A sample of the original and augmented can be found in Fig. 12. The average t-score based on multiple samples of original and augmented data 0.90, indicating that the difference of the population means is not significantly different. These tests ensure the robustness of the proposed data augmentation method.

The final data for all antennas and 100 cases are organised as a matrix and labelled according to known changes made to

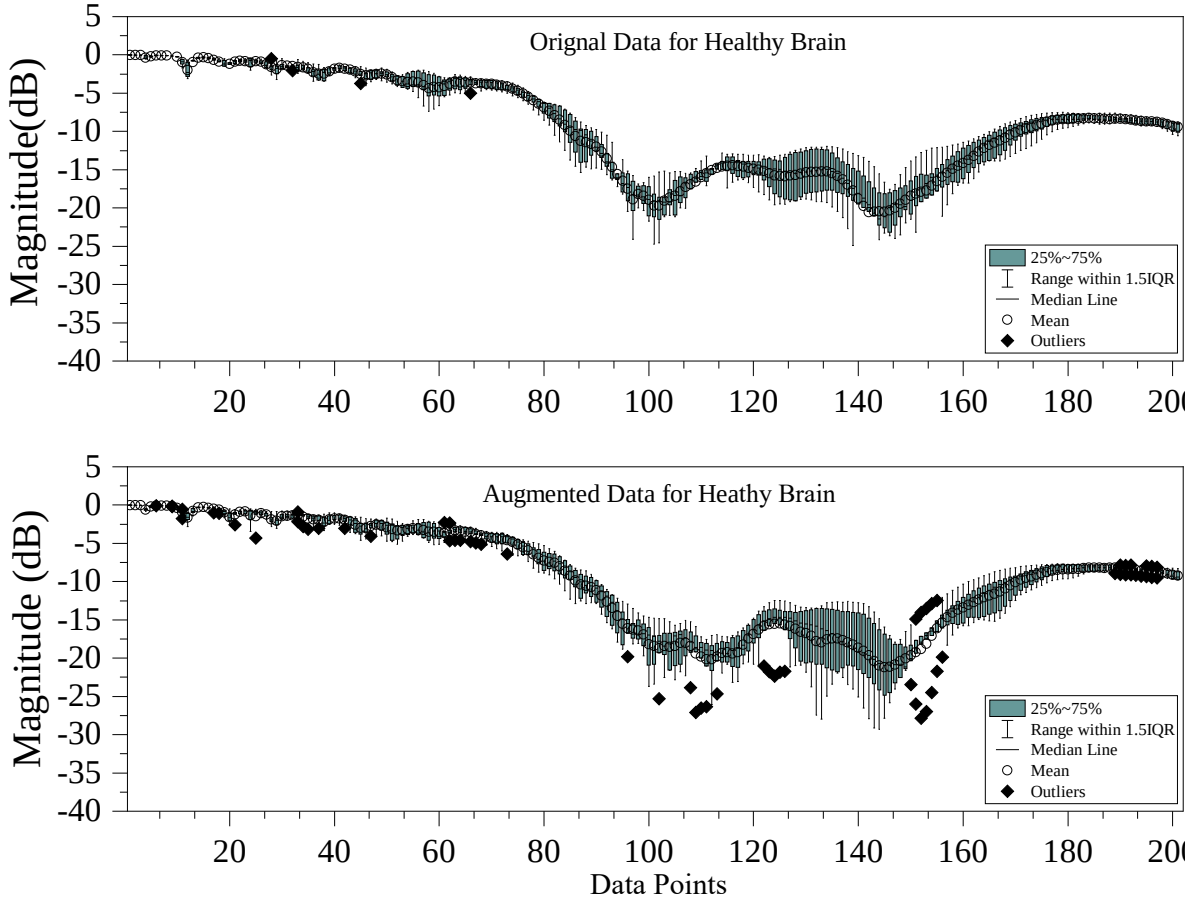


Fig. 10. Comparison of original and augmented data for a healthy brain.

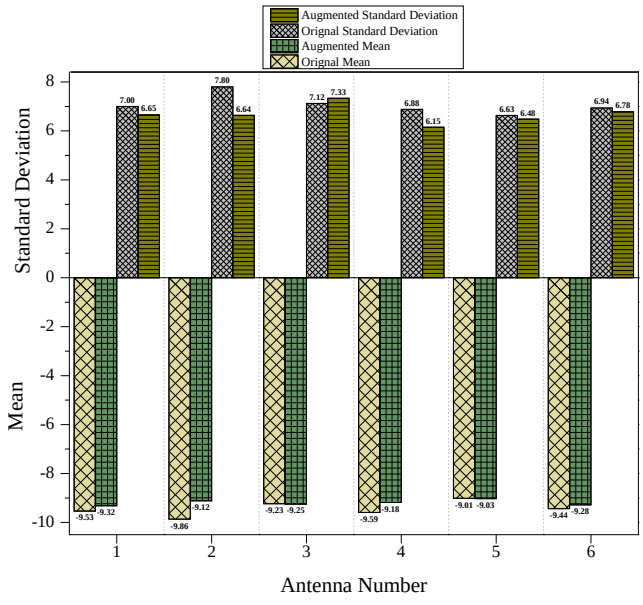


Fig. 11. Mean and standard deviation of original and augmented data for a healthy brain.

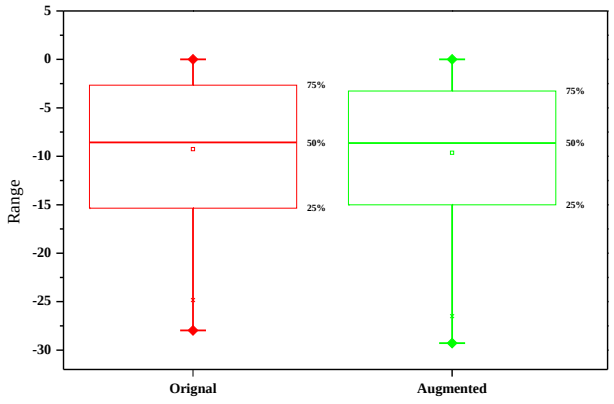


Fig. 12. A sample of the original and augmented data used for statistical analysis.

phantoms. The number of discrete frequency points for each antenna is 201. The total number of original samples for all stages of physiological models are 162×201. The data for pathological models contains the same number of samples. One data point from the beginning was removed to create subsets of 5 values for data augmentation. This does not affect the signal as the signal, in the beginning, are skull and antenna reflections. In clutter rejection techniques, these values, in the beginning, are filtered out or subtracted to eliminate clutter effects [47]. The purpose of removing one discrete point was

not to remove the skull reflections but rather to make the subset of equal size (5 values) required for the proposed augmentation method. Each of these samples contains six antenna data (200 each). The restructured data contain 162×200 points, as shown in Table IV. These data were used to generate data for 100 cases for all stages. The augmented data for all three stages of physiological models contain 1800×200. The data for each antenna in the antenna array is treated as a separate sample. Therefore the final matrix contains 3600×200 sets of data for all cases of both physiological and pathological models. The data is then labelled for three cases and all six antennas. This multi-label data is then used to train various ML algorithms.

TABLE IV
DATA USED TO TRAIN ML ALGORITHMS

	Restructured Data	Augmented Data	Final Data
Physiological Data	162×200	1800×200	3600×200
Pathological Data	162×200	1800×200	

B. Feature extraction with SAE

This paper uses three SAEs to extract key features from the raw RF measurements. To select the optimum SAE, we supervise the training loss curves and the final loss of the network. It can be seen from the Figure 13 that both training loss and validation loss of all three SAEs drop very fast and turn to convergence after about 30 epochs. In detail, it can be seen from Table V that the 128-64-16-64-128 SAE shows the largest training loss among the three SAEs. This is because such a network loses much information as it extracts only 16 features from the original data, which makes the network unable to recover the 16 extracted features back to the original features during the training process. The 128-64-32-64-128 SAE extracts more features and hence provides lower training loss; however, the network shows even a bit higher validation loss. Therefore, the 128-64-128 SAE is designed to extract more features (64) with fewer hidden nodes (less computation complexity). Such SAE provides both the lowest training loss and validation loss among the three SAEs; hence, the 128-64-128 structure is adopted for feature extraction in this study.

TABLE V
TRAINING PERFORMANCE OF THE THREE SAES

Structure	Training loss	Validation loss	Epochs number
128-64-32-64-128	9.87×10^{-5}	4.60×10^{-4}	86
128-64-16-64-128	1.42×10^{-4}	4.34×10^{-4}	96
128-64-128	7.11×10^{-5}	3.40×10^{-4}	97

After applying the well-trained encoder to the raw RF measurements, 64 core features are extracted from the raw 200 features. As observed from Figure 14a, the value of most of the original features is around -20 to 0 dB. Such a sparse distribution of the features could be detrimental to ML models' learning/training process to learn the connection between the features and labels. On the contrary, it can be seen from Figure 14b that the features extracted from the SAE are uniformly distributed between -1 and 1. It is expected that the ML model could have better learning results with such extracted core features.

C. Classification Accuracy

In this subsection, the classification accuracy is calculated to evaluate the performance of the proposed method. The proposed method is compared to two other methods for analysis, including:

- Raw features: applying the ML algorithms to the raw data (features) for classification.
- Manually extracted features: applying the ML algorithms to the manually extracted features for classification.

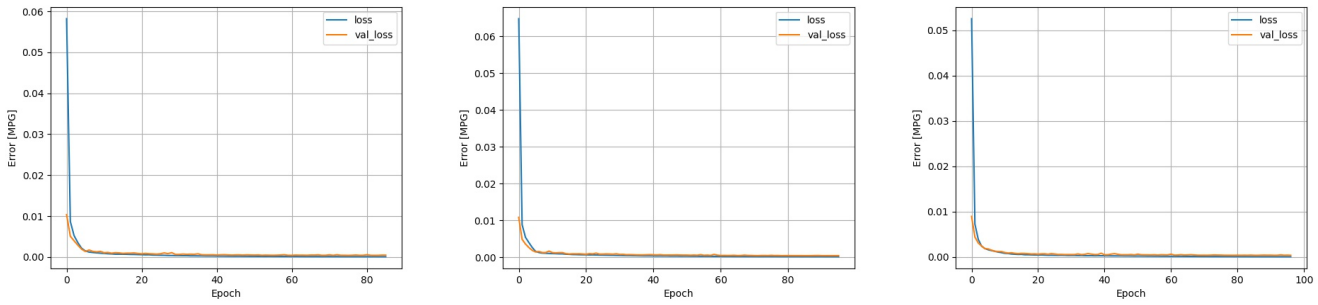
The ML algorithms were trained on raw data, manually extracted features and automatic feature extraction using SAE. The comparison results can be found in Table VI and Figure 15. It can be noted that raw data gives reasonable accuracy, with the random forest having the highest accuracy of 76.8%. The KNN and CART classify these stages with similar accuracy. In comparison with the manual feature extraction, the overall accuracy reduces. However, the accuracy of CART increased by 3%. This is due to the ability of Cart to segment different variables thoroughly to come up with a decision model that it can rely on for classification.

TABLE VI
CLASSIFICATION ACCURACY OF DIFFERENT ALGORITHMS AND FEATURES

	Raw	Manual	Automatic (SAE)
KNN	71.8%	71.4%	79.0%
Random Forest	76.8%	74.7%	83.2%
CART	71.1%	73.1%	81.0%

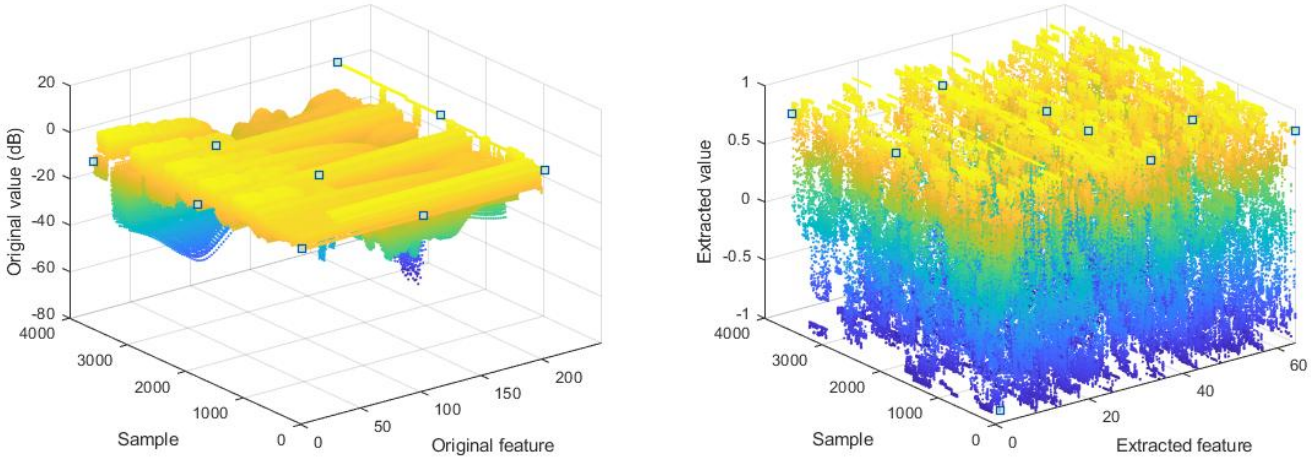
On the other hand, the extracted feature using the SAE significantly improves the overall accuracy compared to the raw data and manually extracted features. The random forest has the highest accuracy of approximately 83.2%, while CART follows up with an accuracy of 81%. These results indicated that the proposed SAE method is best suited to extract features and train ML algorithms. The classification results show that all these algorithms perform well on the multi-class nature of the data. The overall accuracy of 81% is promising for using ML on the RF dataset to monitor early-stage AD detection.

To further evaluate the robustness of the proposed method with Random Forest, which shows the highest classification accuracy in the previous test, we added different levels of Gaussian noise to the data and re-conducted the evaluation process. We choose to evaluate the accuracy by adding different levels of Gaussian noise based on the results reported in [48]. As the results are listed in Table VII, the classification accuracy decreases when the data is noisy in all three situations. Especially, the noise varies the data distribution and shows a significant impact on the manually extracted features. However, the automatically extracted features from SAE perform the best in all cases. Although there is some expected degradation as the noise level rises, the automatically



(a) (b) (c)

Fig. 13. Training loss of three SAEs: (a) 128-64-32-64-128; (b) 128-64-16-64-128; (c) 128-64-128.



(a) (b)

Fig. 14. Comparison between the original and extracted features: (a) original features; (b) extracted features.

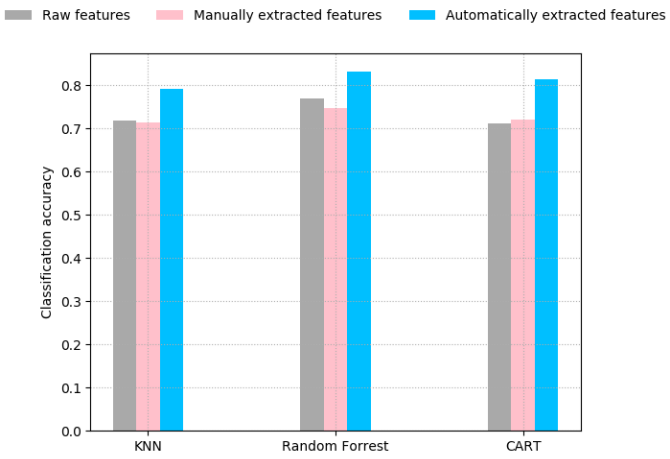


Fig. 15. Comparison of the classification accuracy using different algorithms and features. The auto-encoder based automatic feature extraction method significantly increases the accuracy of all algorithms.

extracted features can always provide 3% to 20% higher accuracy than either using the raw data or manually extracted features.

TABLE VII
CLASSIFICATION ACCURACY FOR RANDOM FOREST ALGORITHM WITH DIFFERENT LEVELS OF NOISE

Noise Level	Raw	Manual	Automatic (SAE)
Noise level 0 (no noise)	76.8%	74.7%	83.2%
Noise level 1 ($\mu = 0, \sigma = 0.02$)	71.8%	66.0%	75.4%
Noise level 2 ($\mu = 0, \sigma = 0.04$)	68.5%	55.4%	70.0%
Noise level 3 ($\mu = 0.2, \sigma = 0.04$)	69.4%	52.8%	72.6%

IV. DISCUSSION AND COMPARISON

Neurodegenerative disorders are among the leading causes of death. These disorders, which impair cognitive abilities and are caused by the progressive loss of neural connections in the brain, affect more than 50 million people worldwide. As the world's population ages, the number of persons affected by such diseases will rise. Alzheimer's disease, in particular, is the most common kind of dementia, with a life

expectancy of three to nine years for those affected. The main physiological symptom of AD is brain atrophy, which causes the brain volume to shrink over time. The pathological changes include the accumulation of plaques and tangles in the brain. Alzheimer's disease starts as a modest preclinical stage and progresses to mild cognitive impairment. The patient's cognitive performance deteriorates as the disease progresses. The patient's cognitive function is significantly impacted in the advanced stages, and the patient cannot live independently. As a result, caring for these individuals is incredibly difficult, and AD is one of the most costly diseases. Early diagnosis and effective patient monitoring can considerably enhance treatment and lower expenses.

Microwave-based medical devices and imaging techniques have been widely investigated during the last decades due to their wearable and portable nature. Many research studies aim to detect abnormal tissues, such as breast cancer and stroke[49] [2]. Recently it has been discovered that there is a significant difference between the dielectric properties of healthy and AD-affected tissue [12]. These changes can be monitored using the microwave [50]. Microwave-based medical imaging has recently been investigated to detect these changes in the brain [3]. It provides promising results; however, small brain changes at early stages can not be detected due to their low spatial resolution.

ML techniques have been widely used in healthcare. Research interest in ML for classifying AD has increased using MRI and CT scan images over the past decade. There has not been much work done on raw signals or dataset(s), such as RF signals. This study explores the use of ML algorithms to detect different stages of AD using microwave-scattered data. The measured RF data for physiological and pathological changes in the brain are scaled to generate synthetic data using a new data augmentation method. Features were extracted using different statistical techniques as well as an SAE. These data were then processed in three ML algorithms to validate how well each algorithm could classify the different stages of AD. Results showed that classical ML algorithms could accurately classify the different stages of AD using the extended RF dataset.

Based on the results, it is shown that the random forest algorithm has the highest average accuracy of approximately 77%. CART follows up with an average accuracy of 75.3% for all three types of input data. Given the large dataset and multi-class nature of the data, random forest and CART looks to be the obvious choice in classifying AD from the dataset. The random forest algorithm is applied to the validation dataset to determine how well it can classify the cases. After running the random forest model on the validation dataset, the confusion matrix and classification report is provided in Table.VIII and Table.IX, respectively.

It was found that random forest had an overall accuracy of 83.2% on the validation dataset. Precision, which is defined as the percent of correct predictions, was less than 90% for all cases. Additionally, it can be noted that there were more misclassifications for the Healthy case and Mild AD as opposed to the severe AD case. This is due to the minimal physiological and pathological changes made to the phantoms for healthy and Mild cases compared to the severe case. While the classification report shows that the random forest performance on the validation dataset has some limitations, the overall accuracy of 83.2% is promising. It gives them the incentive to process the data further to differentiate features more quickly. In addition, the precision of Mild AD detection is relatively high, which is promising as our goal is to use ML on the RF dataset to determine early-stage detection of AD.

TABLE VIII
CONFUSION MATRIX FOR RANDOM FOREST ALGORITHM

	Healthy	Mild AD	Severe AD
Healthy	180	24	12
Mild AD	14	200	26
Severe AD	17	28	219

TABLE IX
CLASSIFICATION REPORT FOR RANDOM FOREST ALGORITHM

	Precision (%)	**Recall (%)**	**F1 Scores (%)**	**Support (#)**
Healthy	83	83	83	240
Mild AD	81	83	82	216
Severe AD	85	83	84	264

Compared to our previous work, [11][10], all these algorithms on raw data do not perform well. This is due to the type of data set used to train these algorithms. In the previous work, the data was generated based on phantoms having similar mesh cells for each case. This work uses nine different compositions of mesh cells to create phantoms for each case to accommodate various head sizes. As a result, a more realistic data set is generated. Furthermore, in the previous work, all antenna data was used as one sample for each case, and 600×1200 sets of data were used. In this work, each antenna data is treated as a separate sample making it 3600×200 sets of data. The classification algorithms do not need all antenna data as one feature to classify different stages of AD. The accuracy of automatic feature extraction significantly improves the accuracy of KNN; however, the accuracy of the CART algorithm remains the same.

It is important to ensure that the measurements are taken with similar configurations and locations of antennas. The current study used simulation data that does not pose any such challenges. However, in real scenarios, follow-up measurements for monitoring different stages may be performed with slightly different positions of the antennas. This would give a natural variation in the data and could overshadow the effect of the minute variations in the dielectric properties of AD-affected and surrounding tissues. In order to avoid these natural variations, different positions on the wearable device should be marked and aligned with the patient's head. Using more antennas to cover the head's surface area will also reduce the risk of variation from a position point of view. Additionally, surface estimations techniques could also be used to allow for adjustable sensor positioning [51]. The surface estimation techniques, such as those utilising laser data, have previously been used for breast cancer detection [52]. Similar techniques could also be used for brain imaging. These algorithms would

enable the evaluation of similarity in positioning and orientation of sensors for repeat visits of a patient. Therefore, the patient-specific surface estimate is a useful tool for evaluating the repeatability in microwave imaging. Furthermore, instead of focusing on readings of one antenna, average reflection or transmission coefficients from all antenna could be used. These average values could be compared with calibrated values to determine progression (even at minute stages). Using the average reflection and transmission coefficients across all frequencies will aid in the comprehension of each brain region and determine the difference in the dielectric values and hence the disease progression.

V. Conclusion and Future Work

An ML-based classification method for classifying different stages of AD is proposed. The data were obtained from numerical phantoms mimicking major stages of AD. Multiple phantoms with different mesh cells compositions were used to accommodate various head sizes. A novel data augmentation method is proposed to create synthetic data for 100 cases for each AD stage. Multiple statistical features are extracted and used to train the ML algorithms. An automatic feature extraction method using a deep neural network-based autoencoder is designed to extract the core latent features from the samples for better classification. Three different classification algorithms were trained on these data and the extracted features. The results were compared with the raw data and manually extracted features. The results indicated that SAE outperformed raw and manually extracted features. The overall accuracy of these algorithms using features extracted through the SAE is 81%, with random forests having the highest accuracy of 83.2%. The proposed ML-based classification method, with microwave imaging algorithms, could be used to monitor AD at its early stages. The proposed method will be evaluated on more realistic data obtained from AD patients in the future.

References

[1] R. Ullah, I. Saied, and T. Arslan, "Measurement of whole-brain atrophy progression using microwave signal analysis," *Biomed. Signal Process. Control*, vol. 71, p. 103083, 2022.

[2] D. O'Loughlin, M. O'Halloran, B. M. Moloney, M. Glavin, E. Jones, and M. A. Elahi, "Microwave breast imaging: Clinical advances and remaining challenges," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 11, pp. 2580–2590, 2018.

[3] I. Saied, T. Arslan, S. Chandran, C. Smith, T. Spires-Jones, and 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 Trans. Med. Imaging*, vol. 39, no. 12, pp. 4060–4070, 2020.

[4] T. Jo, K. Nho, and A. J. Saykin, "Deep learning in alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data," *Frontiers in aging neuroscience*, vol. 11, p. 220, 2019.

[5] I. M. Saied and T. Arslan, "Noninvasive wearable rf device towards monitoring brain atrophy and lateral ventricle enlargement," *IEEE Journal of Electromagnetics, RF and Microw. in Med. and Biol.*, vol. 4, no. 1, pp. 61–68, 2019.

[6] G. Tsang, X. Xie, and S.-M. Zhou, "Harnessing the power of machine learning in dementia informatics research: Issues, opportunities, and challenges," *IEEE reviews in Biomed. Eng.*, vol. 13, pp. 113–129, 2019.

[7] G. Zhu, A. Bialkowski, L. Guo, B. Mohammed, and A. Abbosh, "Stroke classification in simulated electromagnetic imaging using graph approaches," *IEEE Journal of Electromagnetics, RF and Microw. in Med. and Biol.*, vol. 5, no. 1, pp. 46–53, 2020.

[8] M. Persson, A. Fhager, H. D. Trefná, Y. Yu, T. McKelvey, G. Pegenius, J.-E. Karlsson, and M. Elam, "Microwave-based stroke diagnosis making global prehospital thrombolytic treatment possible," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 11, pp. 2806–2817, 2014.

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

[10] I. Saied, T. Arslan, and S. Chandran, "Classification of alzheimers disease using rf signals and machine learning," *IEEE Journal of Electromagnetics, RF and Microw. in Med. and Biol.*, 2021.

[11] R. Ullah, I. Saied, and T. Arslan, "Big data-machine learning processing of recorded radiofrequency physiological and pathological measurements to predict the progression of alzheimer's disease," in *2021 IEEE Asia-Pacific Microwave Conference (APMC)*. IEEE, 2021, pp. 223–225.

[12] I. Saied, M. Bashri, T. Arslan, C. Smith, and S. Chandran, "Dielectric measurements of brain tissues with alzheimer's disease pathology in the microwave region," in *2019 IEEE Int. Symp. on Med. Measurements and Appl. (MeMeA)*. IEEE, 2019, pp. 1–6.

[13] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Int. Conf. on Medic. image Comput. and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[14] I. T. Rekanos, "Neural-network-based inverse-scattering technique for online microwave medical imaging," *IEEE Trans. on Magn.*, vol. 38, no. 2, pp. 1061–1064, 2002.

[15] L. Li, L. G. Wang, F. L. Teixeira, C. Liu, A. Nehorai, and T. J. Cui, "Deepnis: Deep neural network for nonlinear electromagnetic inverse scattering," *IEEE Trans. Antennas Propag.*, vol. 67, no. 3, pp. 1819–1825, 2018.

[16] W. Shao and Y. Du, "Microwave imaging by deep learning network: Feasibility and training method," *IEEE Trams. on antennas and Propag.*, vol. 68, no. 7, pp. 5626–5635, 2020.

[17] H. M. Yao, L. Jiang, and E. Wei, "Enhanced deep learning approach based on the deep convolutional encoder–decoder architecture for electromagnetic inverse scattering problems," *IEEE Antennas Wirel. Propag. Lett.*, vol. 19, no. 7, pp. 1211–1215, 2020.

[18] M. Ambrosanio, S. Franceschini, F. Baselice, and V. Pascazio, "Machine learning for microwave imaging," in *2020 14th European Conf. on Antennas and Propag. (EuCAP)*. IEEE, 2020, pp. 1–4.

[19] X. Chen, Z. Wei, M. Li, and P. Rocca, "A review of deep learning approaches for inverse scattering problems (invited review)," *Progress In Electromagnetics Research*, vol. 167, pp. 67–81, 2020.

[20] A. Kerhet, M. Raffetto, A. Boni, and A. Massa, "A svm-based approach to microwave breast cancer detection," *Eng. Appl. Artif. Intell.*, vol. 19, no. 7, pp. 807–818, 2006.

[21] Z. Wei and X. Chen, "Deep-learning schemes for full-wave nonlinear inverse scattering problems," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1849–1860, 2018.

[22] R. Guo, X. Song, M. Li, F. Yang, S. Xu, and A. Abubakar, "Supervised descent learning technique for 2-d microwave imaging," *IEEE Trans. Antennas Propag.*, vol. 67, no. 5, pp. 3550–3554, 2019.

[23] G. Chen, P. Shah, J. Stang, and M. Moghaddam, "Learning-assisted multimodality dielectric imaging," *IEEE Trans. Antennas Propag.*, vol. 68, no. 3, pp. 2356–2369, 2019.

[24] Y. Sanghvi, Y. Kalepu, and U. K. Khankhoje, "Embedding deep learning in inverse scattering problems," *IEEE Trans. Comput. Imaging*, vol. 6, pp. 46–56, 2019.

[25] Z. Wei and X. Chen, "Physics-inspired convolutional neural network for solving full-wave inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 67, no. 9, pp. 6138–6148, 2019.

[26] R. Guo, Z. Lin, T. Shan, X. Song, M. Li, F. Yang, S. Xu, and A. Abubakar, "Physics embedded deep neural network for solving full-wave inverse scattering problems," *IEEE Trans. on antennas and Propag.*, 2021.

[27] S. P. Rana, M. Dey, G. Tiberi, L. Sani, A. Vispa, G. Raspa, M. Duranti, M. Ghavami, and S. Dudley, "Machine learning approaches for automated lesion detection in microwave breast imaging clinical data," *Scientific reports*, vol. 9, no. 1, pp. 1–12, 2019.

[28] R. Ullah and T. Arslan, "Pyspark-based optimization of microwave image reconstruction algorithm for head imaging big data on high-performance computing and google cloud platform," *Appl. Sciences*, vol. 10, no. 10, p. 3382, 2020.

[29] Y. Wu, B. Liu, and M. Zhu, "A single-pair antenna microwave medical detection system based on unsupervised feature learning," in *Int. Conf. on Computational Social Networks*. Springer, 2018, pp. 404–414.

[30] A. Arayeshnia, A. Keshtkar, and S. Amiri, "Realistic human head voxel model for brain microwave imaging," in *2017 Iranian Conf. on Elec. Eng. (ICEE)*. IEEE, 2017, pp. 1660–1663.

[31] "Dielectric properties of body tissues," http://niremf.ifac.cnr.it/tissprop/htmlclie/htmlclie.php, (Accessed on 10/11/2021).

[32] D. Ancora, L. Qiu, G. Zacharakis, L. Spinelli, A. Torricelli, and A. Pifferi, "Noninvasive optical estimation of csf thickness for brain-atrophy monitoring," *Biomedical optics express*, vol. 9, no. 9, pp. 4094–4112, 2018.

[33] J. B. Pereira, E. Westman, O. Hansson, A. D. N. Initiative, *et al.*, "Association between cerebrospinal fluid and plasma neurodegeneration biomarkers with brain atrophy in alzheimer's disease," *Neurobiology of aging*, vol. 58, pp. 14–29, 2017.

[34] Y. Hidaka, M. Hashimoto, T. Suehiro, R. Fukuhara, T. Ishikawa, N. Tsunoda, A. Koyama, K. Honda, Y. Miyagawa, K. Yoshiura, *et al.*, "Impact of age on the cerebrospinal fluid spaces: high-convexity and medial subarachnoid spaces decrease with age," *Fluids and Barriers of the CNS*, vol. 19, no. 1, p. 82, 2022.

[35] Á. Bernabéu-Sanz, J. V. Mollá-Torró, S. López-Celada, P. Moreno López, and E. Fernández-Jover, "Mri evidence of brain atrophy, white matter damage, and functional adaptive changes in patients with cervical spondylosis and prolonged spinal cord compression," *European radiology*, vol. 30, no. 1, pp. 357–369, 2020.

[36] H. Braak and E. Braak, "Neuropathological stageing of alzheimer-related changes," *Acta Neuropathol. (Berl.)*, vol. 82, no. 4, pp. 239–259, 2015.

[37] I. Saied and T. Arslan, "Wideband textile antenna for monitoring neurodegenerative diseases," in *2018 IEEE 29th Annual Int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2018, pp. 356–360.

[38] A. Yago, M. Cavagnaro, and L. Crocco, "Deep learning-enhanced qualitative microwave imaging: Rationale and initial assessment," in *2021 15th European Conference on Antennas and Propagation (EuCAP)*. IEEE, 2021, pp. 1–5.

[39] Q. Guo, Y. Qian, H. Wang, W.-M. Yu, F. Xu, T. J. Cui, and Y.-Q. Jin, "Recognition rate vs. substitution rate curve: An objective utility assessment criterion of simulated training data," *IEEE Trans. Geosci. Remote Sens.*, 2022.

[40] B. Lewis, O. DeGuchy, J. Sebastian, and J. Kaminski, "Realistic sar data augmentation using machine learning techniques," in *Algorithms for Synthetic Aperture Radar Imagery XXVI*, vol. 10987. International Society for Optics and Photonics, 2019, p. 109870C.

[41] E. Avuclu, "A new data augmentation method to use in machine learning algorithms using statistical measurements," *Measurement*, vol. 180, p. 109577, 2021.

[42] Z. Zhang, Z. Tian, Y. Zhang, M. Zhou, and B. Wang, "u-deephand: Fmcw radar-based unsupervised hand gesture feature learning using deep convolutional auto-encoder network," *IEEE Sensors J.*, vol. 19, no. 16, pp. 6811–6821, 2019.

[43] Y. Dong, T. Arslan, and Y. Yang, "Real-time nlos/los identification for smartphone-based indoor positioning systems using wifi rtt and rss," *IEEE Sensors J.*, 2021.

[44] M. Aslam, "Introducing grubbs's test for detecting outliers under neutrosophic statistics–an application to medical data," *Journal of King Saud University-Science*, vol. 32, no. 6, pp. 2696–2700, 2020.

[45] "Grubbs' test for outliers," https://www.itl.nist.gov/div898/handbook/eda/section3/eda35h1.htm, (Accessed on 09/11/2022).

[46] "Originlab - origin and originpro - data analysis and graphing software," https://www.originlab.com/, (Accessed on 09/11/2022).

[47] A. Zamani and A. Abbosh, "Hybrid clutter rejection technique for improved microwave head imaging," *IEEE Trans. Antennas Propag.*, vol. 63, no. 11, pp. 4921–4931, 2015.

[48] B. Sohani, A. D. Abdallah, G. Tiberi, N. Ghavami, M. Ghavami, and S. Dudley, "An analytically based approach for evaluating the impact of the noise on the microwave imaging detection," in *2021 Photon. & Electromagnetics Res. Symp. (PIERS)*. IEEE, 2021, pp. 296–304.

[49] A. E. Stancombe, K. S. Bialkowski, and A. M. Abbosh, "Portable microwave head imaging system using software-defined radio and switching network," *IEEE Journal of Electromagnetics, RF and Microw. in Med. and Biol.*, vol. 3, no. 4, pp. 284–291, 2019.

[50] J. M. Elloian, G. M. Noetscher, S. N. Makarov, and A. Pascual-Leone, "Continuous wave simulations on the propagation of electromagnetic fields through the human head," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 6, pp. 1676–1683, 2014.

[51] M. Sarafianou, A. W. Preece, I. J. Craddock, M. Klemm, and J. A. Leendertz, "Evaluation of two approaches for breast surface measurement applied to a radar-based imaging system," *IEEE Trans. Antennas Propag.*, vol. 64, no. 2, pp. 609–617, 2015.

[52] D. Kurrant, J. Bourqui, and E. Fear, "Surface estimation for microwave imaging," *Sensors-basel.*, vol. 17, no. 7, p. 1658, 2017.

Rahmat Ullah did his Master in Software Engineering from COMSATS University, Islamabad, Pakistan. He worked as a lecturer at Iqra National University, Peshawar, Pakistan. He is currently pursuing his Ph.D. Degree in the School of Engineering, University of Edinburgh, United Kingdom. He is investigating the use of microwave imaging for detecting and monitoring neurodegenerative diseases such as Alzheimer's disease. His research interest includes Medical Image Processing, Machine Learning, Cloud & Cluster computing and Big-health data.

Yinhuan Dong received his B.Eng.(hons) in Electronic Engineering (Telecommunications) at De Montfort University, UK, in 2018, and his M.Sc. degree in Electronics from the University of Edinburgh, Scotland, UK, in 2019. He is currently working toward the Ph.D. degree with the Embedded Mobile and Wireless Sensor Systems (Ewireless) Group at the School of Engineering, University of Edinburgh. His research focuses on geospatial data science, sensor fusion, and ML for automatic crowdsourced indoor positioning systems.

Prof. Tughrul Arslan holds the Chair of Integrated Electronic Systems in the School of Engineering, University of Edinburgh, Edinburgh, UK. He is a member of the Integrated Micro and Nano Systems (IMNS) Institute and leads the Embedded Mobile and Wireless Sensor Systems (Ewireless) Group in the University (ewireless.eng.ed.ac.uk). His current research focuses on developing low power radio frequency sensors for wearable and portable biomedical applications. He is the author of over 500 refereed papers and inventor of over 20 patents. Prof. Arslan is currently an Associate Editor of IEEE Transactions on VLSI Systems and was previously an Associate Editor for the IEEE Transactions on Circuits and Systems I (2005-2006), IEEE Transactions on Circuits and Systems II (2008-2009). He is also a member of the IEEE CAS executive committee on VLSI Systems and Applications (1999 to date), and is a member of the steering and technical committees of a number of international conferences. He is a co-founder of the NASA/ESA conference on Adaptive Hardware and Systems (AHS) and currently serves as a member of its steering committee.

Prof. Siddharthan Chandran holds the MacDonald Chair of Neurology with the Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, U.K. He works in the emerging discipline of Regenerative Neurology. His research combines laboratory activity that includes human stem cells with specialist clinics to both study several neurodegenerative diseases as well as undertake early phase clinical trials. He is also the Director of the Anne Rowling Regenerative Neurology Clinic and Programme Lead of the UK Dementia Research Institute with the University of Edinburgh, where he leads and develops research activities and clinical experimental infrastructure for longitudinal studies.