Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 3, pp. 1681–1701, DOI 10.12694/scpe.v25i3.2713

LOCALIZATION OF DIELECTRIC ANOMALIES WITH MULTI-LEVEL OUTLIER DETECTION THROUGH MEMBERSHIP FUNCTION AND ENSEMBLE CLASSIFICATION FRAMEWORK

MD. NAJUMUNNISA *, ASCS SASTRY †, AND B T P MADHAV ‡

Abstract. This research presents an innovative method for real-time detection of dielectric anomalies, with a primary focus on evaluating apple quality and ripeness using dielectric tomography. The study involves the development of an advanced tomography system within an anechoic chamber, harnessing electromagnetic wave technology and sophisticated antenna systems for data acquisition. The proposed framework encompasses critical stages, including data collection, range bounds computation, threshold determination, class membership assignment, and ensemble classification. By seamlessly integrating statistical methods, density-based clustering, and ensemble learning, this approach significantly enhances precision and reliability in anomaly detection. The integration of available statistical methods, density-based clustering, and ensemble learning may demand substantial computational resources, limiting the scalability and real-time applicability of the proposed framework. Empirical results demonstrate the superior performance of the method, with an accuracy rate of 98.9 %, precision of 0.989, F-measure of 0.989, dielectric anomaly recall rate of 0.99, and a low error rate of 0.18. Overall, this research introduces an advanced approach with the potential to revolutionize apple quality assessment and industrial processes across various sectors.

Key words: Dielectric anomaly; machine learning; statistical analysis; Fast Fourier transform; near-field imaging; spatial smoothing; VSWR, ensemble learning model.

1. Introduction. Microwave imaging shows promise in detecting breast cancer early due to the differences in dielectric properties of anomaly and non-anomaly exposed to microwave frequencies. This technology offers advantages over X-ray mammography and MRI, including non-ionizing radiation, non-compressive imaging, and cost-effectiveness. Despite these advantages, accurately identifying tumors within the background medium from scattered field data remains a challenge. Further research is needed to develop innovative solutions for precise and efficient breast cancer detection using microwave imaging. The use of microwave imaging has been shown to improve clinical outcomes for breast cancer patients by detecting the disease effectively. However, the nonlinearity and ill-posedness of inverse scattering problems pose significant challenges in this field. Nonetheless, recent advancements have revealed that a small number of observations can detect small scatterers and pinpoint their location with accuracy. But this method requires specific assumptions about the target geometry and data accuracy. Theoretical insights from literature sources [1].can be used to ensure the clear identification of multiple scatterers that remain unidentified. The MUSIC algorithm, a well-established linear technique, has been proven effective in detecting breast cancer using microwave imaging [2]. This method is the top choice for tomography of small inclusions in both full and limited-view inverse scattering problems[3]. It provides quasi-real-time qualitative microwave imaging, which is valuable for those in the field [4]. Additional research and development of these techniques show potential for advancing understanding of inverse scattering problems and their applications [5]. The multiple signal classification (MUSIC) algorithm is a popular technique that employs a well-established method. A precise determination of the nuareer of targets for spectrum computation is necessary to prevent the formation of artificial peaks and to maintain high performance in target localization using pseudo spectrum [6]. While the MUSIC algorithm has demonstrated success in detecting targets of varying sizes, a comprehensive structural analysis is essential to accommodate unforeseen circumstances, such as the emergence of anomalous signals or artifacts that cannot be explained through conventional means [7]. Thus, a detailed structural analysis is strongly advised to gain a more profound comprehension of the system

*Department of ECE, Koneru Lakshmaiah Education Foundation, AP, India. (2002040003@kluniversity.in).,
†Department of ECE, Koneru Lakshmaiah Education Foundation, AP, India. (ascssastry@kluniversity.in).
‡Department of ECE, Koneru Lakshmaiah Education Foundation, AP, India. (madhu.newlook@gmail.com).

and to improve the accuracy of target localization with the application of pseudo spectrum [8]. Gaining a comprehensive understanding of these complex phenomena is crucial for enhancing the overall effectiveness of the methodology [9]. Deriving analytical expressions for the scattered field caused by a slender inclusion is a challenging task that necessitates a deep comprehension of the relationship between the Bessel function of the first kind and the Hankel function of the second kind [10]. Rigorous numerical methods were employed to calculate the eigenvectors and eigenvalues of the MSR matrix to overcome this challenge [11].

The utilization of computers in tomography is referred to as Computerised Tomography, Computer Tomography, or Computer Assisted Tomography (CAT). Tomography has widespread use in medical diagnostics, including MRI, CAT, PET, and ultrasound [12]. Proton Computer Tomography was created specifically for imaging objects the size of a head. Unlike traditional X-rays, which only provide a basic view of an object's outline, tomography offers a comprehensive view of the object's internal composition [13]. The process of reconstruction is utilized to identify the distribution of attenuation in all directions encompassing an object, and CT technology presents the object's cross-sectional data [14]. Industrial tomography has its own set of system requirements for sensor configurations and imaging modalities that differ from those of medical tomography. The X-ray beam intensity is detected in a single direction, which is called a projection [15]. The method of creating an image from projections is used to produce cross-sectional data about an object [16]. Tomography is a diagnostic method that avoids overlapping shadows of structures in front of and behind the region under examination. A tomogram in diagnostic medicine provides a visual representation of the examined area. The exploration of utilizing microwave radiation to capture images of living organisms (at frequencies ranging from 1 to 10 GHz) is a recently developed field. Some people have mistakenly believed that wavelengths in the decimetric range are too lengthy to produce precise spatial resolution or that shorter wavelengths would cause excessive attenuation in the organisms of interest, resulting in limited advancement in this area. Nevertheless, these calculations failed to acknowledge that the wavelength can be reduced when passing through materials with high relative permittivity [17].

Furthermore, enhancing the resolution can be achieved by reducing the aperture size. By using suitable mathematical models, the dispersion exhibited in these measurements can be corrected [18]. The dielectric conductivity of the biosystem is a significant factor in the interaction between microwave energy and biosystems, which is dependent on the complicated permittivity and frequency of the interrogating radiation. Filling the antenna with a higher permittivity material can improve spatial resolution while maintaining its radiating characteristics, by reducing the aperture size of the antenna [19]. The near field of view of an antenna can be utilized to increase the resolution of an imaging system. The utilization of the complete time delay and attenuation measurement through an object can be applied for this purpose at frequencies used in microwaves. X-ray tomography methods rely on the assumption of a straight path approximation, where it is assumed that the ray travels directly to the receiver [20]. The only straight line connecting the transmitter and the observation point is where the ray arriving at each point on the observation plane can be linked to certain properties of the medium. Image reconstruction has utilized various mathematical techniques that have been crucial in medical imaging. There are two main types of reconstruction: analytical reconstruction and iterative reconstruction. Analytical reconstruction involves using precise mathematical solutions to image equations. A common method for this is filtered back projection, which combines all ray sums passing through a point to estimate its density. This technique is widely used in X-ray scanners, but its ability to produce high-resolution images is limited by band-limiting. This means that the image cannot contain spatial frequencies greater than a certain maximum frequency. Iterative techniques can also be used to solve image equations [21]. The Simultaneous Iterative Reconstruction Technique (SIRT), which is a modified version of ART, divides the object into pixels and calculates their values. In situations where there are not enough projections, the samples are noisy, or the projections are taken at limited angles, the iterative method is more efficient than the analytical method [22]. It is becoming less common to conduct mass screening for female breast cancer using X-ray CT due to increased awareness of the risks associated with ionizing radiation. Modern scanners typically acquire data at regular intervals and use analytical reconstruction algorithms that are faster and work well with sufficient sampling. Two types of tomographic methods, diffraction tomography and nonlinear deterministic techniques, are based on the inverse scattering problem. Computed tomography using X-rays has transformed biomedical imaging. The use of waves or low-level microwaves can result in lower quality reconstructions. As a

result, alternative energy sources like ultrasound and low-level microwaves have become popular [23]. However, tomographic reconstructions made using electromagnetic or sound waves are of worse quality than those made using X-rays. This is because X-rays move in straight lines and do not diffract, which enables the transmission data to measure the line integral of a particular object parameter along straight line. The Fourier Slice Theorem can be applied in this case. Nonetheless, tomographic imaging using sound waves or low-level microwaves may not result in energy flowing in a straight line, leading to lower quality reconstructions. Imaging objects with large inhomogeneities requires accounting for refraction and various pathways of energy transmission.

An algebraic reconstruction approach with digital ray tracing and ray linking algorithms can address this issue. However, when the object's inhomogeneities match the wavelength, understanding energy transmission requires considering the wavefronts and fields scattered by the inhomogeneities. Therefore, reconstructing an object's constitutive parameter distributions using electromagnetic waves is an inverse scattering problem [24].

2. Related Work. Traditional CT algorithms like ART are inadequate for imaging objects with large inhomogeneities as they do not consider these effects. The analysis of wavefields can be facilitated by using the iterative method. To explain the propagation of waves in a uniform material, a wave equation, which is a second-order linear differential equation, is used. Directly solving the problem of wave propagation in a non-uniform medium is currently not feasible. Nevertheless, approximate methods based on the theory of wave propagation in uniform media can be applied to obtain solutions in the presence of weak inhomogeneities [25]. The iterative method appears to be a viable option when there are inadequate projections, noisy samples, and restricted projection angles for wavefield analysis. Low-dose radiation techniques are utilized for mass screening of female breasts, resulting in a significant reduction of risks associated with traditional X-ray mammography [26]. The analytical technique is widely preferred among other methods due to its speed and effectiveness with adequate sampling. Two types of tomographic methods, diffraction tomography and nonlinear deterministic technique, are based on the inverse scattering problem. X-ray computed tomography has transformed biomedical imaging. Using X-ray CT for imaging is not suitable anymore. Consequently, alternative energy sources like ultrasound and low-level microwaves have become more popular. However, tomographic reconstructions created with electromagnetic or sound waves are not as good as those made with X-rays [27]. This is because X-rays move in straight lines and do not diffract, allowing the transmission data to measure a specific object parameter along these lines. The Fourier Slice Theorem can be used now [28]. Nevertheless, during tomographic imaging, energy does not always flow in a straight line. For this task, sound or microwaves can be utilized. Energy propagation is characterized by refraction and various pathways, particularly when the objects in homogeneities are greater than the wavelength.

The algebraic reconstruction approach, in combination with digital ray tracing and ray linking algorithms, can partially address the bending caused by refraction. However, when the object's inhomogeneities match the wavelength, the energy transmission must be explained in terms of the wave fronts and fields that the inhomogeneities scatter. Consequently, the reconstruction of an object's constitutive parameter distributions using electromagnetic waves is more appropriately considered as an inverse scattering problem [29]. Traditional CT algorithms do not consider these effects. Certain wave fields have characteristics that make the ART method unsuitable. In order to describe wave propagation in a uniform object, a wave equation, which is a second-order linear differential equation, is utilized. Direct methods are currently unable to resolve the issue of wave propagation in a non-uniform medium. However, approximate formalisms can be used to generate solutions in the presence of weak inhomogeneities by utilizing the theory of homogeneous medium wave propagation [29].

The reconstructed images feature dual mesh technology, which is well-suited for systems that use finite element methods. On the other hand, systems that utilize finite difference methods require a thick mesh. Exposure to high-frequency electromagnetic sources is crucial, and displaying the electric fields over a body accurately is necessary due to their frequent and rapid changes. However, the complex wave numbers tend to remain constant in several sub regions, which may result in incomplete sampling of this characteristic in certain body parts. In areas where twin mesh technology was initially implemented, electric fields are calculated using a uniformly dense mesh type. The second mesh, which is less uniform and less dense overall, reveals the complex wavenumber k distribution within the intended region. The Maxwell equation is utilized in these computations. The log-magnitude and phase of the electric signal observed through field measurements were incorporated [30]. This change allows for the detection of phase variation and its spread across multiple

Rieman sheets in the complex plane. The simulation study and microwave imaging tests conducted showed notable enhancements in image quality for significant, high-contrast objects. It was suggested to apply simple visualization and unwinding methods to determine phase values based on the transmitter and receiver positions. A new approach was introduced to restore images of high quality. The objective is to capture images of objects without any prior assumptions about their size and contrast. Several optimization algorithms have been recorded recently to accomplish this task. Nevertheless, identifying targets remains a significant challenge in general. Medical imaging strives to determine an object's size, location, and constitutive parameters, including its refractive index, specifically for intricate objects. For example, when measuring an object from the dispersed field, electromagnetic waves illuminate it successively. In recent times, effective reconstruction methods have been formulated to address this non-linear and ambiguously defined issue. In earlier investigations concerning microwave imaging, the primary focus was on obtaining images of objects. The use of temporal domain techniques helped to simplify the assumptions of wave propagation. These techniques relied heavily on integral equations in the field, which acted as both interior and exterior integral representations of the scattering object [30]. that many of these methods are iterative, with each iteration requiring the solution of the forward problem. The network and location receive dispersed electric field data from an item, resulting in the production of dielectric permittivity output and the cylinder's radius. The evaluation of results considers input and output parameters, as well as various test data sets.

The algorithm's effectiveness suggests that it can be beneficial for real-time remote sensing applications by solving the inverse scattering problem. Additionally, body reconstruction is also a part of this research. Sharp contrasts can be attained by utilizing a genetic algorithm that concentrates on images. This necessitates converting the Inverse Problem into a global nonlinear optimization problem. In [30], a hybrid genetic algorithm was employed to optimize the configurations of dielectrics for tomographic imaging. Another technique was also proposed, which included representing concealed inhomogeneities using multilayer infinite dielectric cylinders with elliptic cross sections. A cost function was formulated, and field terms were generated using Mathieu's series solution. The functionality was minimally reduced through an innovative optimization method. This method was the first to incorporate back projections with linear filters. They utilized wavelet data that was back projected to decompose the projection into one dimension and discovered an alternative basis for the "Natural pixel" formulation of image reconstruction. The wavelet transform uses modulus sum in the phase-indicated direction. This algorithm utilizes wavelets' localization capability to create a local image reconstruction of the radon transform, which significantly reduces exposure and calculations for X-ray imaging by using nearly local data that corresponds to a region of the body's cross-section. The initial step of the process involves obtaining the quincunx approximation and specific coefficients of a function based on projections. The outcomes of the simulations have indicated an improvement in comparison to the reconstruction that employs discrete wavelets that can be separated. The reconstruction of an image from its sampled projections in the form of Radon Transform Values was accomplished by using a raised-cosine wavelet. However, the projected data slices were found to be noisy when reconstructed with these data due to computed tomography measurements.investigated a new family of regularization techniques for reconstruction that utilizes wavelet and wavelet packet decomposition thresholding. This method is based on the concept that the decomposition almost diagonalizes the inverse and includes prior knowledge of the Radon transformation. In relation to medical images, it was found that these methods had better performance compared to filtered back. Iterative techniques such as OSEM and projection were utilized. Reconstruction of a low dose computed tomographic image leads to instability in its inverse, which then causes problems with noise. This reconstruction technique combines an algorithm with the regularized theory of filtered back projection.

The proposed method, known as the Dual Tree-Complex Wavelet Transform (DT-CWT), can effectively decrease jitter and noise without relying on any specific assumptions about the noise model. Moreover, a relatively simple thresholding technique for recovering functions from noisy data was suggested. In the context of coordinates, the non-linear soft thresholding was subjected to the empirical wavelet coefficient using a method referred to as "Visu shrink." This approach provides an optimal estimation of the mean square error for unknown smoothness functions at a particular location, along with a good visual quality estimator. The original data can be restored by utilizing pyramid filtering after flipping it. A system has been established for this purpose. The wavelet-based technique "SURE shrink" provides minimax rates of confluence across all spaces. The

shrink thresholds are adaptively based, and Stein's Unbiased Risk Assessor (SURE) was employed to determine them. Additionally, a practical spatially adaptive method has been developed. The "Risk shrink" technique is utilized by an adaptive system to decrease the size of empirical wavelet coefficients. This technique involves using different mapping techniques and a discriminative strategy to transform coefficients into functions that produce an open source, noise-free image. used a bilateral denoising filter for images that lacked edge rounding. This technique involves using a spatial non-linear filter to average the data. The approximate (low frequency) sub bands of a signal are filtered bilaterally after being decomposed. The effectiveness of removing noise in real, noisy photographs can be achieved using wavelet filter bank and bilateral multiresolution filter.

A new denoising technique has been developed by measuring the importance of noise-free wavelet coefficients using a local window to define significance, which is known as the "signal of interest." Efficient denoising techniques have also been developed by utilizing the rarity and decorrelation features of DWT, with the method relying on an empirical Bayes estimate based on Jeffrey's uninformative prior. Objective wavelet-based Bayesian denoising can be achieved through a simple fixed non-linear shrinkage rule that outperforms alternative methods that require extensive calculations. proposed the use of Bayesian analysis to construct bivariate non-Gaussian distributions and associated non-linear models, which can be used as new shrinkage functions. Assuming that the coefficients of shrinkage functions have no relation to wavelets is not recommended. repaired damaged images of Gaussian noise by combining a bilateral filter with a trivariate shrinkage filter based on wavelets, and they took into account the wavelet coefficients of the trivariate Gaussian distribution in the wavelet domain. The Maximum-a-Posteriori (MAP) estimator was then used to generate a result, considering the statistical correlations between intrascale wavelet coefficients. introduced the trivariate shrinkage function. A new method for denoising medical images in the wavelet domain was proposed, which uses a technique called guided complex shrinking. This method preserves edges and corners while considering features such as orientation. Additionally, an efficient statistical approach for analyzing medical image wavelet coefficients was presented using a combined denoising method. They developed a novel bivariate Laplacian probability density function model with heavy tails to mimic the statistical data of the wavelet coefficients. A straightforward nonlinear shrinkage function was derived to produce noise-free images. A powerful method has been developed to address both noise and blur in reconstructed images. The algorithm includes two steps. The first step involves using a modified version of the Dual Tree-Complex Wavelet Complex Fourier Wavelet Regularized Transform (DT-CWT) called ComForWaRD for pure denoising, along with a generalized Wiener filter and global blur correction. In the second step, a new technique called BiComForWaRD, which is a variation of ComForWaRD, is used specifically for medical imaging. This two-step procedure includes a denoising algorithm corrected using a generalized Wiener filter and a local adaptive bivariate shrinkage function. Overall, this novel algorithm is a powerful tool that can improve the quality of reconstructed images by addressing both noise and blur. The denoising algorithm called dependencies makes use of wavelet coefficients and statistical dependence statistics. The Laplacian probability density function-based pyramid with local dependencies is a denoising technique that can be utilized. Wavelet soft thresholding was suggested by Grace Chang and S. Al as a flexible data-driven approach for picture denoising. Within the Bayesian framework, a prior is frequently utilized to determine the threshold applied to wavelet coefficients. The Generalized Gaussian Distribution (GGD) was proposed as a framework for image processing applications [44]. They introduced a novel statistical signal processing approach in the statistical domain that is wavelet-based, utilizing Hidden Markov Models.

1. Scanning: Scanning is the initial step in the image reconstruction process. It involves acquiring raw data from the object or scene of interest using a suitable imaging device, such as a CT (Computed Tomography) scanner, MRI (Magnetic Resonance Imaging) machine, or a similar medical or industrial imaging system. During scanning, data is collected in the form of signals or measurements that capture the properties of the object being imaged.

2. Discretization: Once the raw data is collected, it often needs to be discretized. Discretization involves converting the continuous data obtained from the scanning process into a discrete format that can be processed and reconstructed digitally. This step typically involves techniques such as sampling and quantization, where the continuous measurements are transformed into discrete data points or pixels, allowing for digital processing.

3. Back-Projection: Back-projection is a crucial step in image reconstruction, especially in the context of tomographic imaging techniques like CT scans. It involves mathematically reconstructing a two-dimensional

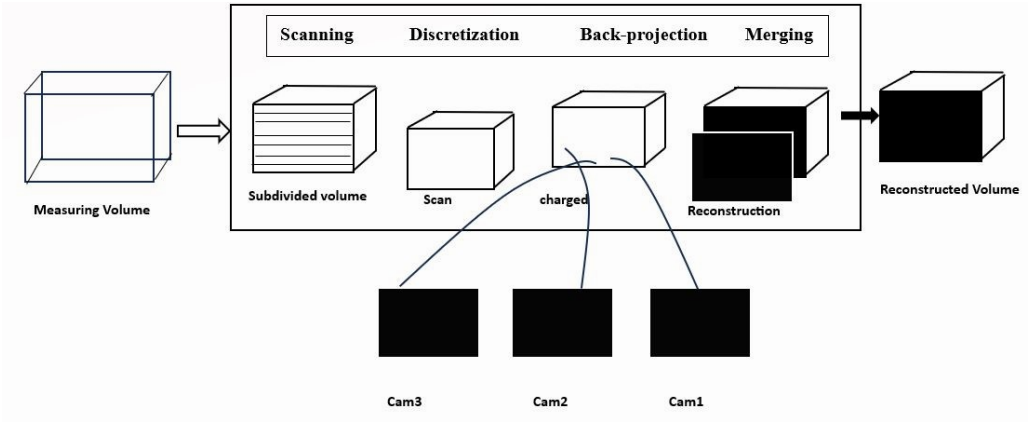

Fig. 3.1: Reconstruction process

or three-dimensional image from the collected data. Back-projection algorithms distribute the discretized data back into a grid or image space, attempting to recreate the distribution of properties within the imaged object. This process is iterative and involves mathematical computations to refine the image reconstruction.

4. Merging: In some cases, multiple sets of data from different scans or imaging angles are merged to create a more comprehensive and accurate image. This merging process combines the reconstructed images from various perspectives to create a single, fused image with improved spatial resolution and reduced artifacts. This is particularly important in techniques like CT, where multiple X-ray projections are acquired from different angles and merged to form a 3D image.

Each of these steps plays a crucial role in the image reconstruction process, allowing us to transform raw data into meaningful and interpretable images. The choice of algorithms, the quality of data acquisition, and the precision of the reconstruction techniques significantly impact the final image's accuracy and clarity, making this process essential in fields such as medical imaging, industrial inspection, and scientific research.

Main contribution of the research:

- Presents a comprehensive approach for dielectric anomaly detection. The study begins by applying the FFT data transformation approach to the raw input data.
- Followed by proposing a gamma quartile distribution-based method for anomaly detection on the transformed data.
- Additionally, a hybrid ensemble learning classifier is introduced to identify dielectric anomalies in real-time test data. As the conventional dielectric anomaly detection models primarily rely on static raw signal data with signal values
- Developed method enhances the overall anomaly detection process of dielectric raw datasets by incorporating statistical range bound analysis and an ensemble classification approach.

The structure of the paper is as follows: Section 3 follows with the proposed framework describing with its system, antenna design, tomographic image reconstruction methodology with their algorithms. Section 4 presents the proposed ensemble learning model. Section 5 illustrates the experimental results and section 6 concludes the research work.

3. Proposed Framework. In this work, primarily an antenna is designed which operates at 0.915GHz and 2.45GHz frequencies. Here apple is taken as the object under test (OUT) and the anomaly on the defected apple has to be reconstructed. The values of dielectric properties like dielectric constant and dielectric loss factor of the object at 0.915 GHz frequency has been considered.

Figure 3.1 illustrates the reconstruction process of an image, which typically involves several key steps, including scanning, discretization, back-projection, and merging. The dielectric constant and dielectric loss factor values of apple at 0.915GHz are 57 and 8. The object is placed in between the proposed antenna and the transmitting antenna, and the signals are collected from the VNA (Vector Network Analyzer).

(a) Healthy Object (b)Defected Object

Fig. 3.2: Experimental setup for tomography reconstruction

3.1. System design. In this section, we provide details about the developed tomography system, which was carried out within an anechoic chamber and, in this context represents a healthy apple and a defective apple. The object under examination is positioned between two antennas: the proposed antenna (acting as the receiving antenna) and the transmitting antenna, which are positioned opposite each other, as illustrated in this section. Specifically, the transmitting antenna is a horn antenna, responsible for emitting signals, while the proposed antenna functions as the receiving antenna.To initiate the experiment, a source generator is connected to the transmitting antenna to provide the necessary excitation, and the receiving antenna is linked to an Agilent Vector Network Analyzer (VNA) to capture the scattered field data. During data collection, the position of the receiving antenna is incrementally shifted in 5-degree phases, and the signals are then routed from the VNA for analysis. It's important to note that this experimental setup takes place within an anechoic chamber, which is equipped with microwave-absorbing materials. These materials serve the crucial role of blocking external radio frequency interference, thus reducing the influence of environmental electromagnetic effects. The use of these absorbers is essential to optimize the performance of the prototype system and achieve accurate and reliable results.

In summary, this section outlines the configuration of the tomography system, its placement within an anechoic chamber, and the equipment and procedures involved in experiments, all aimed at minimizing external electromagnetic interference and ensuring the effectiveness of the developed system which was shown in Figure 3.2.

3.2. Antenna design. In microwave tomography systems, the use of highly directional antennas is common to ensure effective illumination of the object under test (OUT). Figure 3.2 illustrates the design of a monopole antenna created on FR4 material with specific properties, including a dielectric constant of 4.3 and a thickness of 1.6 mm.The antenna's main radiating element, known as the patch, is fabricated from copper material with a thickness of 0.035 mm and incorporates a partial ground plane on the opposite side of the antenna. Additionally, a microstrip feedline is integrated into the design to facilitate the efficient transmission of electromagnetic signals to the patch.To ensure proper impedance matching between the patch and the 50-ohm probe line, careful consideration was given to impedance matching techniques. Furthermore, metamaterials were strategically placed on both sides of the feed line to enhance the antenna's performance.The Computer Simulation Technology (CST) software was employed to fine-tune and optimize the antenna design. The simulated and measured reflection coefficients, often represented as S11 in antenna engineering, were analyzed and compared. The computational model presents the results of this comparison, showcasing how closely the simulated data aligns with the measured data, which is critical for validating the antenna's efficiency. The performance measures, on the other hand, display the antenna patterns. These patterns illustrate how the antenna radiates electromagnetic energy in different directions, highlighting its directional characteristics. Accurate antenna patterns are essential for directing and focusing the emitted signals toward the object under

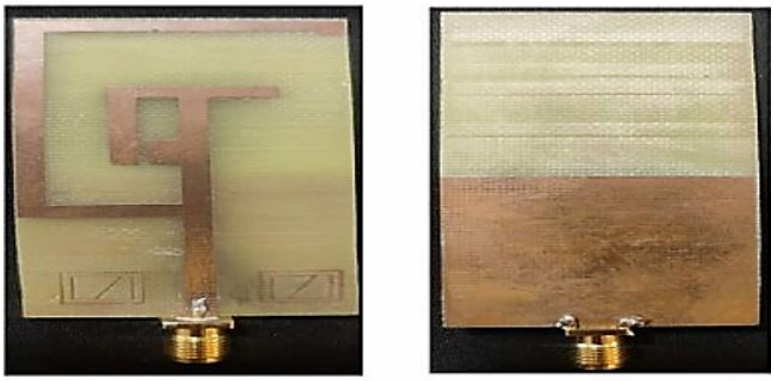

Fig. 3.3: Fabricated antenna designed on FR4. (a) front view, (b) back view

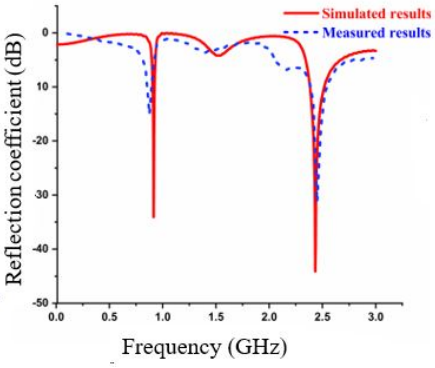

Fig. 3.4: Simulated Measure S11 parameter

test during the tomography process, ensuring optimal data acquisition and image reconstruction.In summary, the described monopole antenna design, constructed with specific material properties, impedance matching techniques, and metamaterial enhancements, underwent rigorous simulation and measurement procedures to validate its performance in terms of reflection coefficients and radiation patterns, all of which are vital for its successful application in microwave tomography systems.

In Figure 3.3 (a), we examine the results of simulated and measured reflection coefficients and VSWR (Voltage Standing Wave Ratio) for the antenna operating at two distinct frequencies: 0.915 GHz and 2.45 GHz. The bandwidth considered for both frequencies is 20 MHz. These parameters are crucial in assessing the antenna's performance and impedance-matching characteristics. Figure 3.3 (b), on the other hand, provides insights into the radiation patterns exhibited by the proposed antenna at the aforementioned operating frequencies, observed along both the E plane (Electric Plane) and H plane (Magnetic Plane). At the lower frequency of 0.915 GHz, the radiation pattern in the E plane indicates a bidirectional nature, meaning that the antenna emits signals in two primary directions.

Simultaneously, along the H plane, the radiation pattern is omnidirectional, implying that the antenna radiates signals in all directions evenly. As the frequency increases to 2.45 GHz, the radiation pattern changes. In the E plane, it exhibits dipole characteristics, suggesting a preference for signal emission in two opposing directions. Along the H plane, it remains nearly omnidirectional, signifying uniform radiation in all directions. These findings highlight the frequency-dependent nature of the antenna's radiation patterns, with its behavior

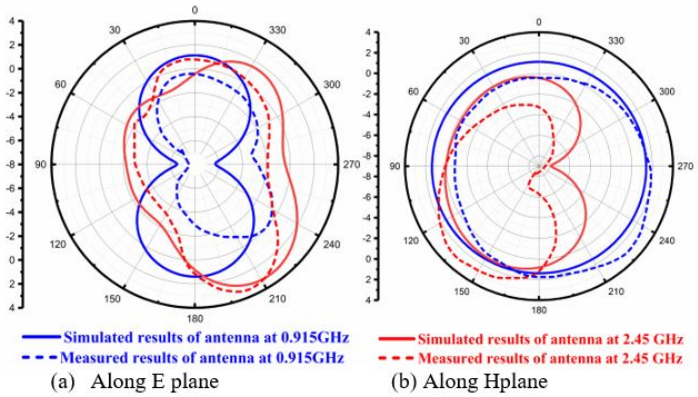

Fig. 3.5: Simulated and Measured radiation pattern results of proposed antenna

transitioning from bidirectional and omnidirectional at lower frequencies to dipole and omnidirectional at higher frequencies which was explained in Figure 3.4.

These insights are essential for optimizing the antenna's performance and its suitability for specific applications in different frequency bands shown in Figure 3.5

3.3. Tomographic image reconstruction methodology. The provided framework outlines a novel approach for dielectric anomaly detection using real-time sensor data. Let's break down the technical details step by step:

1. Data Collection: Initially, dielectric raw data is collected from real-time sensors. These sensors are responsible for capturing the electrical properties of materials under test, which are crucial for anomaly prediction.

2. Range Bounds Calculation: In this step, a gamma distribution-based method is employed to calculate the range bounds. These bounds are used to establish upper and lower thresholds for anomaly detection. Gamma distribution is a probability distribution often used to model data with positive skewness, which makes it suitable for modeling certain types of dielectric data.

3. Thresholding: For each signal in the raw data, both lower and upper bounds are computed based on the gamma distribution parameters. These bounds serve as reference points to filter out the extreme values in the data, specifically the upper outliers and lower outliers. This filtering process helps in identifying data points that deviate significantly from the expected range.

4. Class Membership Function: The filtered outlier data is then subjected to a class membership function. This function assigns a class label to each data point, categorizing them as either normal or anomalous based on their relationship to the computed bounds. This classification step is crucial for distinguishing between expected and unexpected behavior in the dielectric data.

5. Ensemble Classification: Finally, an ensemble classification framework is proposed for predicting dielectric anomalies in real-time test data. Ensemble methods typically combine the predictions of multiple classifiers to improve overall accuracy and robustness. This ensemble approach leverages the outputs of the class membership function to make predictions about dielectric anomalies, enhancing the reliability of the detection process.

visually represents this comprehensive dielectric anomaly detection framework. It showcases the entire workflow, from data collection and threshold calculation to the final ensemble classification step. This framework is designed to effectively identify anomalies in real-time dielectric data, which is valuable for various applications such as quality control, fault detection, and industrial processes shown in Figure 3.6.

Fundamental steps for the tomographic image reconstruction process. The proposed system involves a series of fundamental steps as follows:

1. Data Acquisition: The initial step revolves around gathering data from an array of sensors, such as microphone arrays or antenna arrays. This data typically comprises a collection of time-domain samples.

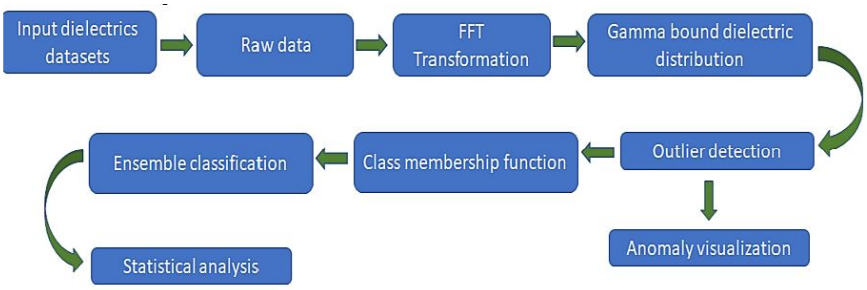

Fig. 3.6: Proposed distributed student dropout prediction model

2. FFT Transformation: Subsequently, the Fast Fourier Transform (FFT) is applied to the acquired data to yield its frequency-domain representation, enabling analysis in the spectral domain.

3. Covariance Matrix Estimation: Estimation of the covariance matrix is carried out using the received signal. This matrix reflects the correlation among various sensor outputs and is instrumental in deriving eigenvalues and eigenvectors.

4. Eigenvalues and Eigenvectors Computation: Eigenvalues and eigenvectors are computed based on the estimated covariance matrix. Eigenvectors signify the directions from which signals arrive, while eigenvalues provide insights into the signal strength in each direction.

5. Spatial Spectrum Construction: Leveraging the obtained eigenvectors and the FFT of the received signal, a spatial spectrum is constructed. This spectrum portrays the signal power across different directions, allowing for directional analysis.

6. Application of Dielectric Anomaly Detection Approach: The system then applies a dielectric anomaly detection approach to assign class labels based on the spatial spectrum and other pertinent information.

7. Validation of Class Labels: The assigned class labels undergo validation through inter and intra similarity distance assessments, ensuring the accuracy and reliability of the labeling process.

8. Data Storage as Training Data: The validated data is stored for subsequent use as training data, facilitating the system's learning and improvement over time.

9. Ensemble Classification Framework Application: An ensemble classification framework is employed to predict anomalies in dielectric test data. This framework incorporates multiple classification techniques for robust prediction.

10. Statistical Metric Analysis: The system performs an in-depth analysis of statistical metrics to evaluate the performance of the ensemble classification framework, providing insights into its accuracy and effectiveness. In essence, these technical steps collectively form a comprehensive system for processing and analyzing data from sensor arrays, with a focus on dielectric anomaly detection, classification, and performance evaluation.

3.4. Algorithm: Dielectric anomaly detection approach. Compute the following hybrid IQR method as

$$DF[x] = features(); \tag{3.1}$$

$$R1 = (V(|DF|/4)); \tag{3.2}$$

$$R2 = (V(|DF|/2) + V(|DF|/2 + 1))/2; \tag{3.3}$$

$$U\_E[x] = R3 + \eta.log(\Gamma(R3 - R1)) \tag{3.4}$$

$$L\_E[x] = R1 - \eta.log(\Gamma(R3 - R1)) \tag{3.5}$$

$$U\_Outlier = R3 + \eta.min(\chi(R3 + R1, 9)), exp(\Gamma(R3 - R1))) \tag{3.6}$$

$$L\_Outlier = R3 - \eta.min(\chi(R3 + R1, 9)), exp(\Gamma(R3 - R1))) \tag{3.7}$$

Algorithm: Interquartile Range outlines (IqR)

- **Initialization:** Begin with setting the loop index variable i to 1.
- **Loop Start:** Start a loop that iterates from $i = 1$ to the size of the set $|FS|$.
- **Access Element:** Retrieve the i-th element ($F(i)$) from the set F.
- **Check Range:** Examine whether the value of $F(i)$ falls above or below the predefined upper and lower range.
- **Anomaly Check:** If $F(i)$ is found to be above or below the specified range, mark it as an anomaly.
- **Non-Anomaly Check:** If $F(i)$ is within the acceptable range, label it as a non-anomaly.
- **Perform Actions:** Execute any necessary actions or logging related to the anomaly or non-anomaly status of the current element.
- **Increment Index:** Increment the loop index variable i to move on to the next element in the set.
- **Loop Continuation:** Repeat the loop until all elements in the set F have been processed.
- **Process Completion:** Finish the process, having labeled each element in the set F as either an anomaly or a non-anomaly based on its position relative to the upper and lower ranges.

The above algorithm outlines a hybrid method for outlier detection using the interquartile range (IQR) for dielectric data. Here's a step-by-step explanation of the algorithm:

Initialize the array DF[x] as the input features. (Assuming DF[x] contains the dielectric data.) Compute the first quartile range R1, which is the value at the 25th percentile of the absolute values of DF[x]. Compute the second quartile range R2, which is the average of the values at the 25th and 75th percentiles of the absolute values of DF[x]. Compute the third quartile range R3, which is the average of the values at the 75th percentile and 25th percentile from the end of the absolute values of DF[x].Compute the upper outlier threshold U_E[x], defined as R3 plus a parameter eta multiplied by the logarithm of the gamma function evaluated at R3 minus R1.Compute the lower outlier threshold L_E[x], defined as R1 minus a parameter eta multiplied by the logarithm of the gamma function evaluated at R3 minus R1. Compute the upper outlier limit U_Outlier, defined as R3 plus eta times the minimum of the chi function evaluated at R3 plus R1 with 9 degrees of freedom, and the exponential of the gamma function evaluated at R3 minus R1. Compute the lower outlier limit L_Outlier, defined as R3 minus eta times the minimum of the chi function evaluated at R3 plus R1 with 9 degrees of freedom, and the exponential of the gamma function evaluated at R3 minus R1. Perform outlier detection: Iterate through each feature F(i) in the dielectric data FS[x]. If F(i) lies above the upper outlier threshold U_E[x] or below the lower outlier threshold L_E[x], label it as an outlier. Otherwise, do further processing or analysis with the non-outlier data. The algorithm combines the IQR method with additional statistical measures (gamma function, chi function) and thresholds (U_E[x], L_E[x], U_Outlier, L_Outlier) to identify outliers in the dielectric data.

3.5. K-class membership validation approach.

Step 1: To each dielectric labeled partition data PD.

Step 2: To each data point p[i] in PD.

Step 3: Perform weight age density to each data object using the Gaussian transformation function as

$$wd_i = \sum_{j\epsilon, i!=j} exp(-(\frac{d_{ij}}{dc})^2) \tag{3.8}$$

$$d_{i,j} = \sqrt{\sum_{t=1}^{m}(x_i^t - x_j^t)^2} \tag{3.9}$$

Where d_c: Threshold

Step 4: To each object, find the highest density using the following measures

$$hd_j = max\{wd_i\} \tag{3.10}$$

Step 5: Computing the k nearest neighbors by using the k-randomized centres as k initial clusters as. The set of k nearest neighbors of center CPi is defined

$$NP_i^k = \{P_j/min(d_{ij}), i! = j\} \tag{3.11}$$

Step 6: Compute the inter cluster similarity and intra cluster similarity to each k-neighbor initial clusters using the following formula.

$$\lambda 1 = IntraClu(p_c, p_i) = \frac{1}{n_i - 1} hd_c . \sum_{m=1} d(p_c, p_m) \tag{3.12}$$

$$\lambda 2 = IntraClu(p_c, p_i) = min_{1<=m<=k}(\frac{1}{n_m} hd_c . \sum_{r=1} d(p_c, p_r)) \tag{3.13}$$

alpha =Q1, beta =Q2, chi =Q3

$$UppOutlier = \chi + \Gamma max\left\{\lambda 1, \lambda 2\right\}, (\chi - \alpha) \tag{3.14}$$

$$LowerOutlier = \chi - \Gamma max\left\{\lambda 1, \lambda 2\right\}, (\chi - \alpha) \tag{3.15}$$

Step 7: Iterate until all points are assigned to k=clusters or no more changes in clusters.

For each data point in a partition, a weightage density is computed using a Gaussian transformation function. The highest density among neighboring data points is then determined for each object. Initial clusters are formed by randomly selecting k centers and finding their k nearest neighbors. Inter-cluster and intra-cluster similarity measures are calculated for each initial cluster. The algorithm iteratively updates the cluster assignments until convergence or a stopping condition is reached. While specific formulas and details are missing, this algorithm aims to cluster data based on density and proximity, refining the assignments through iterative steps.

4. Proposed Ensemble learning model. Ensemble learning is a robust and effective technique that harnesses the collective intelligence of multiple base models to enhance overall performance and generalization in solving complex problems. When employed in the context of localizing dielectric anomalies, which entails the detection of regions within a material exhibiting atypical electrical properties, an ensemble approach can significantly improve the accuracy and reliability of the results.

1. Base Model Diversity: Ensemble learning starts by creating a set of diverse base models, each trained to recognize dielectric anomalies from different perspectives or with distinct algorithms. These models could include various machine learning techniques like decision trees, support vector machines, or neural networks, each with its own strengths and weaknesses.

2. Aggregated Predictions: Once the base models are trained, their individual predictions are aggregated to make a collective decision. This aggregation can be performed using methods like majority voting, weighted averaging, or stacking, depending on the specific ensemble strategy chosen.

3. Anomaly Identification: The aggregated predictions from the ensemble are then used to identify and localize dielectric anomalies within the material. This involves examining the consensus among the base models. Regions where a significant portion of the models agree on the presence of anomalies are more likely to be actual anomalies.

4. Thresholding: To improve the precision of anomaly localization, thresholding techniques can be applied. This means considering only those regions where the level of agreement among the base models surpasses a predefined threshold. This helps filter out false positives and enhances the robustness of anomaly detection.

5. Visualization: To make the results more interpretable, the localized dielectric anomalies can be visualized. This often involves overlaying the anomaly predictions onto the original dielectric property maps or other relevant visual representations. This allows for a clear visualization of where the anomalies are located within the material.

6. Performance Evaluation: The performance of the ensemble is rigorously evaluated using appropriate metrics like precision, recall, F1-score, and accuracy. This step ensures that the ensemble effectively identifies and localizes dielectric anomalies while minimizing false alarms.

7. Model Refinement: The ensemble's performance can be further improved by refining the base models or adjusting the ensemble strategy based on the evaluation results. This iterative process fine-tunes the ensemble for optimal performance.

Ensemble learning leverages the strengths of multiple models to provide accurate and robust dielectric anomaly localization. It combines their predictions, applies thresholding for precision, and visualizes the results, ultimately enhancing the understanding of regions with unusual electrical properties within the material. This approach is especially valuable in applications such as quality control, defect detection, and materials science.

Initialization of ensemble learning process "Set of proposed classifier and base classifiers are represented as ensemble classifiers as"

$$EC = \{KNN, HDT, Proposedmodel\} \tag{4.1}$$

$$MC = \{\}\,; //Modelclassifier \tag{4.2}$$

$$CO = \{\}\,; //classifierOutput \tag{4.3}$$

Procedure The procedure outlines a sequence of technical steps involved in constructing a Bayesian network using input data and estimating Bayesian network node variables based on two different cases, one for discrete attributes and the other for continuous attributes. Additionally, it introduces a novel feature ranking measure aimed at optimizing the decision tree construction process.

1.Data and Variables: The input data is represented as D. PAi represents the set of input random variables. PIj represents an instance of attribute.

2.Bayesian Network Construction: For each attribute PAi in data PD. Construct the Bayesian network graph using the input data and naïve Bayesian estimations.

3.Estimation of Discrete Attributes: Estimate Bayesian network node variables using the two cases. If the attribute type is discrete, then the Bayesian discrete parameters are estimated using the following measure as

$$P(A_i = I_{k/C_j}) = \frac{N_{ijk}}{N_i} \tag{4.4}$$

where N_ijk is the number of instances of class cj having the value Ik in attribute Ai.

4.Estimation of Continuous Attributes: If the attribute type is continuous type, then the Bayesian continuous parameters are estimated using the following measure as

$$P(A_i = I_k/C_j) = G(I_k, \mu_{ij}, \sigma_{ij}) \tag{4.5}$$

$$G(I_k, \mu_{ij}, \sigma_{ij}) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} e^{\frac{(I_{k-\mu_{ij}})^2}{2\sigma_{ij}^2}} \tag{4.6}$$

5. Feature Ranking for Decision Tree Optimization: In order to optimize the decision tree construction process, a novel feature ranking measure is proposed to optimize the problem of pruning.

To each feature in DF[]

do

Table 5.1: (Sample data) Dielectric FFT transformed data

S No	Frequency Sample	s1, 1/Rate	s2, 1/Rate	s3, 1/Rate	s4, 1/Rate
1	0.0	0.9943	0.1188	0.2415	0.2723
2	0.001	0.9945	0.1268	0.2412	0.2743
3	0.002	0.9950	0.1519	0.2397	0.2723
4	0.003	0.9957	0.1877	0.2380	0.2711
5	0.004	0.9965	0.2239	0.2365	0.2700

Finding rank of the feature as

$$S = D.log(D); \tag{4.7}$$

$$p1 = -s/((\sqrt{\sum D[i]})^3 * \sqrt{\chi(D) * \sum D[i]}) \tag{4.8}$$

$$p2 = -(s * CE(D))/(\chi(D) * \sum D[i])^3 \tag{4.9}$$

$$Rank(A[i]) = Max\{p1, p2\} \tag{4.10}$$

6.Optimizing Pruning: This feature ranking measure is utilized to optimize the decision tree construction process, specifically in the context of pruning. In essence, this technical process involves constructing a Bayesian network, estimating parameters for both discrete and continuous attributes, and introducing a novel feature ranking measure to enhance the decision tree construction process, ultimately optimizing the problem of pruning. These steps are fundamental in various data-driven applications, including machine learning and data analysis.

5. Experimental Results. Dielectric tomography is a non-invasive imaging method that relies on the interaction between electromagnetic waves and matter to create detailed images of an object's internal structure. In the context of apple dielectric tomography, this technique is employed to ascertain the distribution of dielectric properties within an apple's interior. This information is valuable for assessing the apple's quality and degree of ripeness.

The data provided in the question seems to be a table of dielectric tomography measurements conducted on multiple apples. Each row within the table corresponds to a single measurement, while each column represents a different sensor or a group of sensors. The "Frequency Sample" column likely denotes the frequency of the electromagnetic wave employed during the measurement. Meanwhile, columns labeled "S1" through "S8" are associated with various sensors or sensor clusters utilized in the measurements. The "cluster" column appears to categorize each measurement into specific clusters, possibly based on certain criteria or sensor configurations. Lastly, the "Outlier" column serves to indicate whether a particular measurement is considered an outlier or not.

Table 5.1 encapsulates a portion of the collected data, which undergoes processing involving FFT (Fast Fourier Transform) transformation. This transformation is applied to convert the data from the frequency domain into a time-domain series. It's important to note that this data processing step is essential for further analysis and interpretation of the dielectric properties within the apples.In essence, this dataset and the described processing steps are instrumental in leveraging dielectric tomography to gain insights into the internal properties of apples, aiding in quality assessment and ripeness evaluation shown in table 5.1.

Table 5.2 presents the dielectric anomaly prediction values of the ranked attributes computed from the sample data of the defected apple which is placed in between the transmitter and receiver antennas. This table contains values related to the prediction of dielectric anomalies specifically computed for attributes that have been ranked. These attributes likely pertain to the sample data of a defected apple positioned between transmitter and receiver antennas. The prediction values could represent the likelihood or probability of a dielectric anomaly being present in relation to these attributes. These values are crucial for assessing the presence and severity of anomalies in the apple's dielectric properties.

Table 5.2: Ranked features for the dielectric anomaly prediction

Ranked Attributes	**Ranked Attributes**
0.9477	38.S5,5/rate
0.9363	20 S3,3 /rate
0.7971	11 S2,2 /rate
0.4374	65 S8,8 /rate
0.4157	29 S4,4 /rate
0.4071	47 S6,6 /rate
0.4004	56 S7,7 /rate
0.3339	59 S2,8 /rate
0.3339	17 S8,2 /rate
0.2736	45 S4,6 /rate
0.2736	31 $6,4 /rate
0.2665	62 S5,8 /rate
0.2665	41 S8,5 /rate
0.2657	53 S4,7 /rate
0.2622	32 S7,4 /rate
0.2471	35 S2,5 /rate
0.2471	14 $5,2 /rate
0.2408	2 S1,1 /rate
0.233	22 S5,3 /rate
0.233	36 S3,5 /rate
0.2235	51 $2,7 /rate
0.2232	16 S7,2 /rate

Table 5.3: Statistical analysis of proposed model for dielectric anomaly detection on Realtime training data

Correctly classified instances	99.6004%
Incorrectly classified instances	0.3996%
Kappa statistics	0.9713
Mean absolute error	0.006
Root mean square error	0.0629
Relative absolute error	4.2547%
Root relatively squared error	23.8332%

Table 5.3 provides a detailed breakdown of the statistical analysis conducted on our proposed model for detecting dielectric anomalies in real-time training data. Table 3 offers a detailed statistical analysis of a proposed model that has been developed for the purpose of detecting dielectric anomalies. This model has presumably been trained on real-time data.The statistical analysis may encompass various metrics and evaluations that assess the model's performance. It aims to provide a comprehensive understanding of how well the model performs in detecting dielectric anomalies.

Figure 5.1 explicates the plot of feature-based correlation using dielectric anomaly detection process. In this the first rows gives the features of outlier; second row gives the feature of cluster and from third row it gives the features of the sampled data. The red colour depicts the anomaly object features correlation with respect to rows parameters and columns parameters.

In this section explains that the model demonstration at a high accuracy of 99.6 %, indicating that the majority of instances are classified correctly. he Kappa statistic measures the agreement between the model's predictions and the actual classifications. A value of 0.9713 suggests a very high level of agreement. These metrics quantify the average magnitude of errors in the model's predictions. Low values indicate that the model's predictions are close to the actual values on average shown in Table 5.3.

Table 5.4: Detailed accuracy by class

	TP Rate	FP Rate	Precision	Recall	F-call	MCC	ROC	PRC	Class
	0.998	0.026	0.998	0.998	0.998	0.971	0.981	0.997	no
	0.974	0.002	0.974	0.974	0.974	0.971	0.981	0.930	yes
Weighted average	0.996	0.025	0.996	0.996	0.996	0.971	0.981	0.992	

Table 5.5: Confusion matrix

a	b	Classified as
1847	4	a=no
4	147	b=yes

In Table 5.4, explains various key parameters, including precision, recall, frequency measure, accuracy, error rate, and the weighted average of the sampled data. In Table 4, a detailed examination of various critical parameters related to the model's performance is presented.

Precision: Indicates the accuracy of positive predictions.

Recall: Measures the ability to correctly identify positive instances.

Frequency Measure: Represents a harmonic mean of precision and recall.

Accuracy: Reflects the overall correctness of the model's predictions.

Error Rate: Indicates the proportion of incorrect predictions.

Weighted Average: Averages the sampled data with weighted consideration.

Table 5.5 illustrates the confusion matrix generated as a result of the anomaly detection process. Table 5 provides a visualization in the form of a confusion matrix, which is a fundamental tool in evaluating the performance of a classification model. The confusion matrix likely breaks down the model's predictions into true positives, true negatives, false positives, and false negatives. It aids in understanding the model's ability to correctly classify anomalies and non-anomalies.

Figure 5.1 illustrates feature-based correlation within the context of the dielectric anomaly detection process. Let's break down the details of this plot:

1. Feature-Based Correlation: Feature-based correlation refers to the analysis of how different features or attributes within a dataset are related to each other. It's a fundamental step in understanding the relationships and dependencies among various aspects of the data.

2. Row Structure: The figure is organized into rows, with specific information presented in each row. The first row appears to represent the features of outliers. These are data points or instances that deviate significantly from the expected or normal behavior and are often considered anomalies.The second row is dedicated to the features of clusters. Clusters typically group data points with similar characteristics or properties together.

3. Columns Parameters: The columns likely represent various parameters or attributes within the dataset. These parameters could be related to dielectric properties, sensor readings, or other relevant factors.

4. Color Coding: The use of color, specifically the color red, is employed to visually depict the correlation between different features or attributes. The red color often signifies a strong positive correlation or association between variables.

5. Visualization: The figure visually represents how each feature, whether it belongs to outliers, clusters, or sampled data, correlates with the parameters or attributes represented by the columns.

By examining the color coding, one can discern the strength and direction of the correlation. For example, if a feature in the first row (outliers) is strongly red-colored about a specific column parameter, it indicates a notable correlation between outliers and that parameter. Overall, Figure 5.1 serves as a visual aid to help understand the interrelationships between different features and parameters in the context of dielectric anomaly detection. It provides insights into which attributes are more closely related to anomalies, clusters, or the sampled data, which is valuable for further analysis and decision-making in anomaly detection processes.

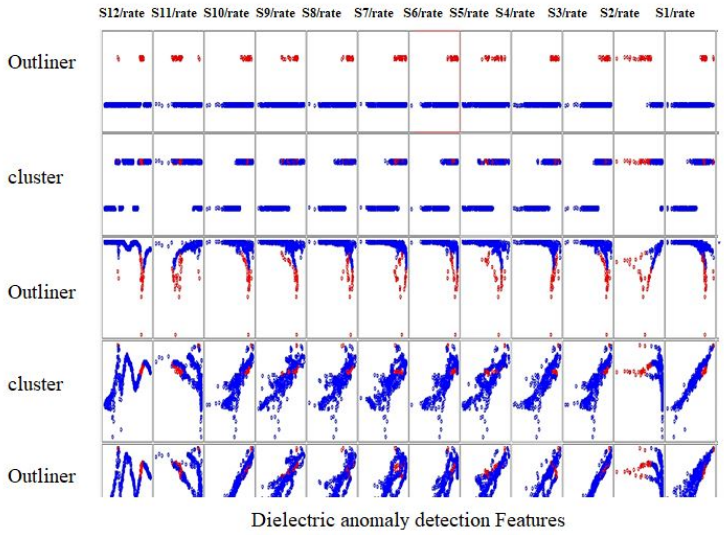

Fig. 5.1: Features-based correlation plot for the dielectric anomaly detection

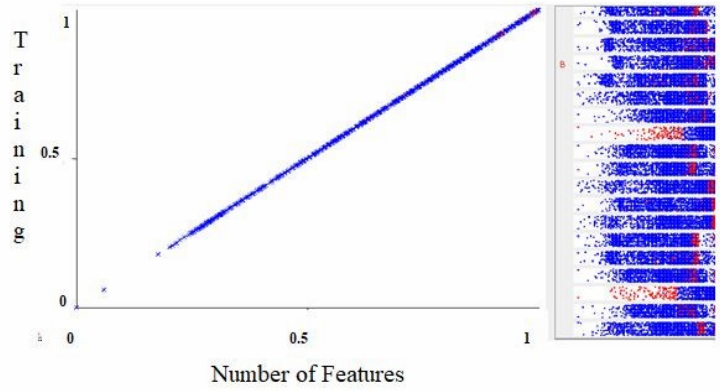

Fig. 5.2: Single feature-based correlation plot for the dielectric anomaly detection.

In Figure 5.2, presents a graphical representation of a single feature and its correlations within the context of our anomaly detection process. The red marks visible in Figure 5.2 represent potential anomaly objects, which have been identified as data points exhibiting unusual behavior or characteristics. These anomalies could signify deviations from expected patterns in the feature space. Conversely, the blue points in figure 5.2 correspond to non-anomaly objects, which are data points displaying normal behavior within the feature space. This visual depiction allows us to observe the distinct separation between potential anomalies (red) and typical data points (blue), demonstrating the effectiveness of our anomaly detection method in identifying aberrations within the single feature and its associated correlations.

In Figure 5.3, we present a comparative analysis of our proposed ensemble dielectric anomaly detection model alongside established models such as KNN, Naive Bayes, logistic classifier, and Random Forest classifier. This analysis encompasses various characteristic metrics including accuracy, precision, Frequency measure, recall, and error rate, all evaluated on the dielectric dataset. The proposed ensemble dielectric anomaly detection approach achieved an impressive accuracy of 98.9%, signifying its ability to correctly classify anomalies and

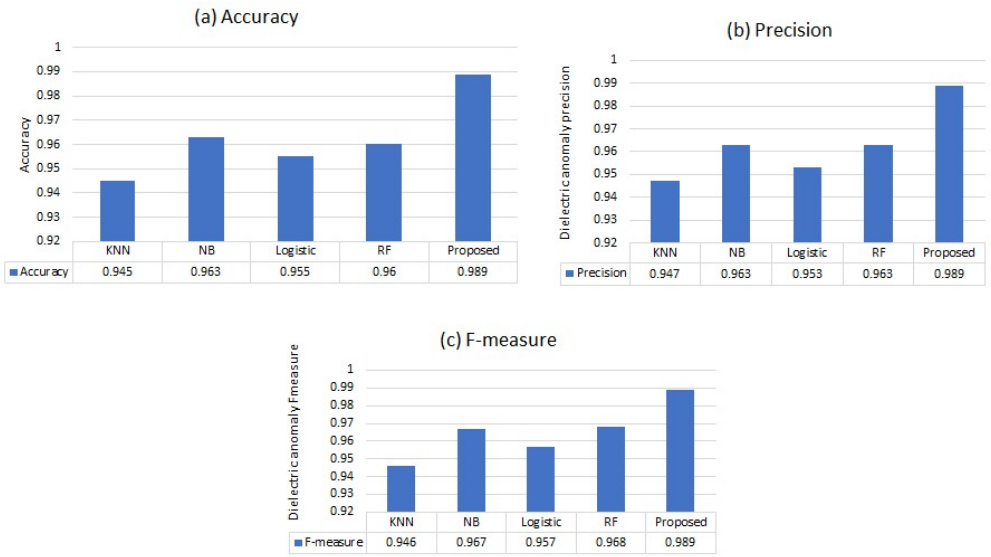

Fig. 5.3: Comparative analysis of proposed ensemble dielectric anomaly detection characteristics to the conventional anomaly detection classifier characteristics on the dielectric training dataset.

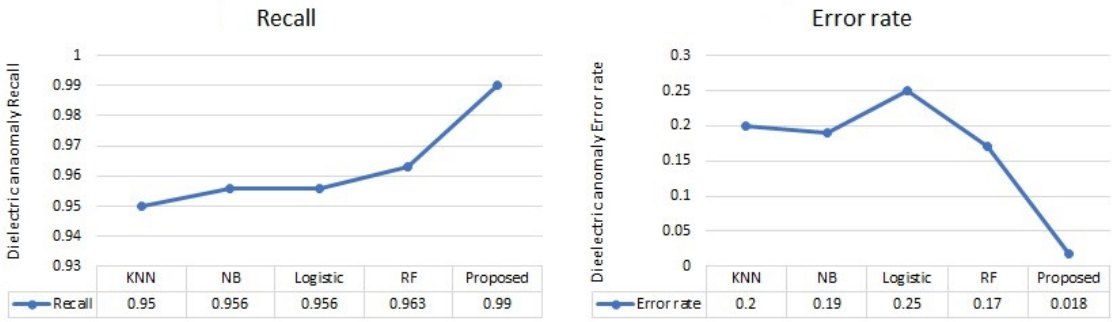

Fig. 5.4: Recall and Error-rate

normal instances within the dataset.

The precision value of 0.989 indicates a high proportion of correctly identified anomalies among all the anomalies detected, minimizing false positives which is shown in Figure 5.3.

The F-measure value of 0.989 demonstrates a balanced harmonic mean of precision and recall, emphasizing the model's effectiveness in handling both precision and recall simultaneously. Dielectric anomaly recall stands at 0.99, indicating a strong ability to capture a significant portion of actual anomalies, which is crucial for comprehensive anomaly detection. Lastly, the error rate, at 0.18, showcases the model's capability to minimize misclassifications and overall classification errors, further underlining its robustness in the context of dielectric anomaly detection. In summary, the proposed ensemble model outperforms traditional classifiers in terms of accuracy, precision, recall, and overall error rate, making it a promising choice for dielectric anomaly detection tasks shown in Figure 5.4.

In Figure 5.5(a, b), we present images of both a selected healthy apple and a defective apple, offering visual representations of these two distinct conditions. Figure 5.5(a) depicts a photograph of a selected healthy apple.

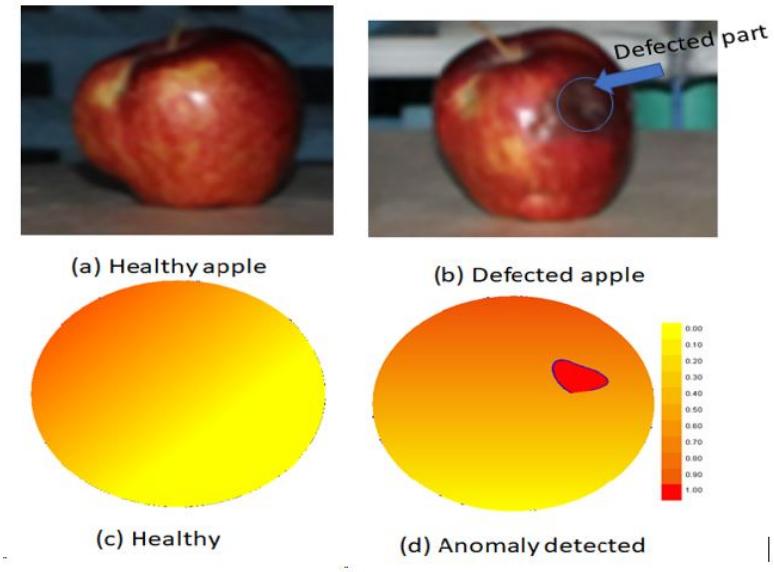

Fig. 5.5: Tomographic image reconstruction, (a) Healthy apple, (b) defected apple, (c, d) image reconstructions of healthy and defective apple.

This image provides a clear and detailed view of a well-formed, undamaged apple, showcasing its external appearance without any apparent defects or blemishes. In contrast, Figure 5.5(b) displays a photograph of a defective apple. In this image, we can observe visible imperfections or irregularities on the surface of the apple, indicating the presence of defects such as bruises, spots, or other forms of damage. Moving on to Figure 5.5 (c, d), we delve into the internal characteristics of these fruits through image reconstruction. Figure 5.5 (c) presents the image reconstruction of the healthy part of the fruit. This reconstruction allows us to visualize the internal structure of the healthy apple, revealing details of its interior composition without defects.In Figure 5.5(d), we focus on the image reconstruction of the defective part of the fruit. This isolated reconstruction provides a closer look at the specific region within the apple that exhibits abnormalities or defects. It enables a detailed examination of the internal damage or irregularities present in the defected portion. These images and reconstructions are instrumental in quality control and assessment processes, aiding in the identification and differentiation of healthy and defective apples, which is vital for ensuring product quality in the agricultural and food industries. This research utilizes sophisticated methodologies, incorporating a multi-level outlier detection strategy and a membership function, for the precise identification of dielectric anomalies. The application of an ensemble classification framework further improves the accuracy and robustness of the process for localizing anomalies. In summary, this study significantly contributes to the advancement of efficient techniques in detecting and pinpointing dielectric anomalies within intricate systems.

6. Conclusion. This study introduces a robust method for dielectric anomaly detection, particularly for assessing apple quality and ripeness using dielectric tomography. The developed tomography system, featuring advanced antennas within an anechoic chamber, enables precise data collection. The framework incorporates statistical analysis, density-based clustering, and ensemble learning to enhance anomaly detection accuracy.The research underscores the importance of accounting for the frequency-dependent characteristics of dielectric properties, which impact antenna radiation patterns and subsequently affect data collection and anomaly detection processes. This insight is critical for optimizing system performance across various frequency bands.Empirical results validate the effectiveness of our approach in accurately identifying dielectric anomalies in real-time sensor data. The ensemble classification framework enhances anomaly detection reliability, making it valuable for quality control, fault detection, and industrial applications.In summary, this research contributes significantly

to dielectric tomography and anomaly detection, providing a validated framework. This innovation has the potential to revolutionize non-invasive quality assessment across industries, particularly for perishable goods like apples. In conclusion, our approach can transform quality control, reduce waste, and ensure product consistency, benefiting both consumers and industries.

REFERENCES

[1] G. Kaur,, *"A comparison of two hybrid ensemble techniques for network anomaly detection in spark distributed environment"*,Journal of Information Security and Applications, vol. 55, p. 102601, Dec. 2020, doi: 10.1016/j.jisa.2020.102601.

[2] H. Bhatt and M. Shah , *"A Convolutional Neural Network ensemble model for Pneumonia Detection using chest X-ray images"*, Healthcare Analytics, vol. 3, p. 100176, Nov. 2023, doi: 10.1016/j.health.2023.100176.

[3] J. Jiang et al.,, *"A dynamic ensemble algorithm for anomaly detection in IoT imbalanced data streams,"* Computer Communications, vol. 194, pp. 250–257, Oct. 2022, doi: 10.1016/j.comcom.2022.07.034.

[4] J. Fuentes-Velazquez, E. Beltran, E. Barocio, and C. Angeles-Camacho. , *"A fast automatic detection and classification of voltage magnitude anomalies in distribution network systems using PMU data,"* Measurement, vol. 192, p. 110816, Mar. 2022, doi: 10.1016/j.measurement.2022.110816.

[5] A. K. Dey, G. P. Gupta, and S. P. Sahu, , *"A metaheuristic-based ensemble feature selection framework for cyber threat detection in IoT-enabled networks,"* .Decision Analytics Journal, vol. 7, p. 100206, Jun. 2023, doi: 10.1016/j.dajour.2023.100206.

[6] I. Piekarz et al., , *"A microwave matrix sensor for multipoint label-free Escherichia coli detection,"* Biosensors and Bioelectronics, vol. 147, p. 111784, Jan. 2020, doi: 10.1016/j.bios.2019.111784.

[7] M. Vishwakarma and N. Kesswani,, *"A new two-phase intrusion detection system with Naïve Bayes machine learning for data classification and elliptic envelop method for anomaly detection,"* Decision Analytics Journal, vol. 7, p. 100233, Jun. 2023, doi: 10.1016/j.dajour.2023.100233.

[8] Y. Kayode Saheed, O. Harazeem Abdulganiyu, and T. Ait Tchakoucht, *"A novel hybrid ensemble learning for anomaly detection in industrial sensor networks and SCADA systems for smart city infrastructures,"* Journal of King Saud University - Computer and Information Sciences, vol. 35, no. 5, p. 101532, May 2023, doi: 10.1016/j.jksuci.2023.03.010.

[9] K. B. Sahay, B. Balachander, B. Jagadeesh, G. Anand Kumar, R. Kumar, and L. Rama Parvathy, , *"A real time crime scene intelligent video surveillance systems in violence detection framework using deep learning techniques,"* Computers and Electrical Engineering, vol. 103, p. 108319, Oct. 2022, doi: 10.1016/j.compeleceng.2022.108319.

[10] A. Muhammad and F. Külahci, , *"A semi-supervised total electron content anomaly detection method using LSTM-auto-encoder,"* Journal of Atmospheric and Solar-Terrestrial Physics, vol. 241, p. 105979, Dec. 2022, doi: 10.1016/j.jastp.2022.105979.

[11] E. Mushtaq, A. Zameer, and A. Khan, , *"A two-stage stacked ensemble intrusion detection system using five base classifiers and MLP with optimal feature selection,"* Microprocessors and Microsystems, vol. 94, p. 104660, Oct. 2022, doi: 10.1016/j.micpro.2022.104660.

[12]] J. Jiang et al., , *"AERF: Adaptive ensemble random fuzzy algorithm for anomaly detection in cloud computing,"* Computer Communications, vol. 200, pp. 86–94, Feb. 2023, doi: 10.1016/j.comcom.2023.01.004.

[13] L. A. Souto Arias, C. W. Oosterlee, and P. Cirillo, , *"AIDA: Analytic isolation and distance-based anomaly detection algorithm,"* Pattern Recognition, vol. 141, p. 109607, Sep. 2023, doi: 10.1016/j.patcog.2023.109607.

[14] O. Abu Alghanam, W. Almobaideen, M. Saadeh, and O. Adwan , *"An improved PIO feature selection algorithm for IoT network intrusion detection system based on ensemble learning,"* Expert Systems with Applications, vol. 213, p. 118745, Mar. 2023, doi: 10.1016/j.eswa.2022.118745.

[15] A. Copiaco et al., , *"An innovative deep anomaly detection of building energy consumption using energy time-series images,"* Engineering Applications of Artificial Intelligence, vol. 119, p. 105775, Mar. 2023, doi: 10.1016/j.engappai.2022.105775

[16] W. Khan and M. Haroon,, *"An unsupervised deep learning ensemble model for anomaly detection in static attributed social networks,"*International Journal of Cognitive Computing in Engineering, vol. 3, pp. 153–160, Jun. 2022, doi: 10.1016/j.ijcce.2022.08.002.

[17] Saikumar, K., Arulanantham, D., Rajalakshmi, R. et al, *"Design and Development of Surface Plasmon Polariton Resonance Four-Element Triple-Band Multi-Input Multioutput Systems for LTE/5G Applications"* Plasmonics, vol. 18, p. 1949–1958, 2023, doi: 10.1007/s11468-023-01922-w.

[18] Vasimalla, Y., Pradhan, H.S., Pandya, R.J. et al, *"Titanium Dioxide-2D Nanomaterial Based on the Surface Plasmon Resonance (SPR) Biosensor Performance Signature for Infected Red Cells Detection,"* Plasmonics , vol. 18, p. 1725–1734,2023, doi: 10.1007/s11468-023-01885-y.

[19] P. J. E. Peebles, *"Anomalies in physical cosmology,"* Annals of Physics, vol. 447, p. 169159, Dec. 2022, doi: 10.1016/j.aop.2022.169159.

[20] Revathi, R., Vatambeti, R., Saikumar, K., Alkhafaji, M. A., Khairy, U. R., & Noori, S., *" An advanced online mobile charge calculation using artificial intelligence,"* AIP Conference Proceedings, vol. 2845, p. 050026, Dec. 2023, doi: 10.1063/5.0170422.

[21] Vasimalla, Y. , Pradhan, H.S. , Pandya, R.J. , ... Rashed, A.N.Z. , Hossain, M.A., *"Anomalies in physical cosmology,"* Annals of Physics, vol. 447, p. 169159, Dec. 2022, doi: 10.1016/j.aop.2022.169159.

[22] S.-H. Son, K.-J. Lee, and W.-K. Park, , *"Application and analysis of direct sampling method in real-world microwave imaging,"* Applied Mathematics Letters, vol. 96, pp. 47–53, Oct. 2019, doi: 10.1016/j.aml.2019.04.016.

[23] W.-K. Park, , *"Application of MUSIC algorithm in real-world microwave imaging of unknown anomalies from scattering matrix,"* Mechanical Systems and Signal Processing, vol. 153, p. 107501, May 2021, doi: 10.1016/j.ymssp.2020.107501.

[24] G. Ludeno, I. Catapano, A. Renga, A. R. Vetrella, G. Fasano, and F. Soldovieri, , *"Assessment of a micro-UAV system for microwave tomography radar imaging,"* Remote Sensing of Environment, vol. 212, pp. 90–102, Jun. 2018, doi: 10.1016/j.rse.2018.04.040.

[25] M. S. Benmoussat, M. Guillaume, Y. Caulier, and K. Spinnler, , *"Automatic metal parts inspection: Use of thermographic images and anomaly detection algorithms,"* Infrared Physics & Technology, vol. 61, pp. 68–80, Nov. 2013, doi: 10.1016/j.infrared.2013.07.007.

[26] S. Liu et al., , *"CEUS Versus MRI in Evaluation of the Effect of Microwave Ablation of Breast Cancer,"* Ultrasound in Medicine & Biology, vol. 48, no. 4, pp. 617–625, Apr. 2022, doi: 10.1016/j.ultrasmedbio.2021.11.012.

[27] A. Chohra, P. Shirani, E. B. Karbab, and M. Debbabi, , *"Chameleon: Optimized feature selection using particle swarm optimization and ensemble methods for network anomaly detection,"* Computers & Security, vol. 117, p. 102684, Jun. 2022, doi: 10.1016/j.cose.2022.102684.

[28] J. Yu and J. Kang, *"Clustering ensemble-based novelty score for outlier detection,"* Engineering Applications of Artificial Intelligence, vol. 121, p. 106164, May 2023, doi: 10.1016/j.engappai.2023.106164.

[29] Maji, A.,Choubey, G, *"Improvement of heat transfer through fins: A brief review of recent developments,"* Heat Transfer, vol. 49, p. 1658-1685, Feb 2020, doi: 10.1002/htj.21684.

[30] Baskar, M., Ramkumar, J., Karthikeyan, C., Anbarasu, V., Balaji, A.,Arulananth, T. S, *"Low rate DDoS mitigation using real-time multi threshold traffic monitoring system,"* Journal of Ambient Intelligence and Humanized Computing, p. 1-9, Jan 2021, doi: 10.1007/s12652-020-02744-y.

Edited by: Polinpapilinho Katina
Special issue on: Scalable Dew Computing for Future Generation IoT Systems
Received: Oct 19, 2023
Accepted: Dec 20, 2024