# A New Feature Extraction Method for Stroke Classification with Microwave Signals Based on Early Stroke Volume Change

**Meqdad Taheri[1], Mohammad Saeed Majedi[2]**
[1]Communications and Computer Research Center, Ferdowsi University of Mashhad, Mashhad 9177948883, Iran
[2]Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad 9177948883, Iran

Corresponding author: Mohammad Saeed Majedi (majedi@um.ac.ir)

**ABSTRACT** This study proposes a new feature extraction method for microwave-based brain stroke classification using differential scattering parameters. In addition to the differences in dielectric properties, a further clinical distinction between hemorrhagic and ischemic strokes after onset is that hemorrhagic stroke volume tends to evolve during the first few hours, whereas ischemic stroke volume remains relatively stable. This early-stage behavioral disparity leads to measurable variations in scattering parameters across successive measurements. Therefore, according to the proposed method, scattering parameters are acquired from the patient's head at two distinct times, and the subtraction of these measurements is used as a feature for classification. A synthetic database of training and test data is generated using three-dimensional magnetic resonance imaging (MRI)-based numerical head phantoms enclosed by a 21-element modified printed monopole antenna array. The physical dimensions, dielectric properties, stroke sizes, and lesion locations are systematically varied to represent diverse clinical scenarios, and realistic noise is incorporated into the simulated data. A standard support vector machine (SVM) classifier is employed to evaluate the proposed feature extraction method on the generated dataset, yielding average accuracies of 97.11% and 78.55% at noise levels of −115 dB and −95 dB, respectively; these results indicate improvements of 34.15 and 16.68 percentage points compared to the use of raw scattering parameters.

**Keywords** Feature extraction, hemorrhagic stroke, ischemic stroke, microwave system, SVM, brain stroke diagnostics.

## I. INTRODUCTION

Stroke remains a major global health challenge, affecting more than 13 million people each year and resulting in around 5.5 million deaths annually[1]. Strokes are categorized into two primary types: ischemic and hemorrhagic. Over 87% of strokes are ischemic, while about 13% are hemorrhagic [2]. The treatment protocols for these two types differ significantly, as therapeutic strategies effective for one type may be detrimental for the other. Therefore, accurate diagnosis of the stroke type is crucial to ensure that appropriate treatment is promptly administered [2].

Given the importance of rapid diagnosis, current clinical practice relies heavily on advanced imaging modalities such as magnetic resonance imaging (MRI) and computed tomography (CT) scanners. Despite their high accuracy, these equipment are not portable, and their unavailability in remote areas can lead to delays in treatment. This lack of access to diagnostic facilities, especially during the critical first 4.5 hours following an attack, can significantly diminish the chances of patient recovery and lead to permanent damage [3].

Over the past few decades, the application of microwave technology in medical fields has emerged as a significant area of research interest among scientists [4].

About ten years ago, a novel device that combined microwave techniques with machine learning algorithms for stroke type classification was introduced in [5]. The main goal of this system is to enable early and portable stroke classification using machine learning algorithms. To achieve this, suitable training data must be prepared, for which an array of antennas is positioned around the patient's head, where the stroke type has already been identified using conventional methods, to measure the transmission and reflection coefficients, known as scattering parameters. These data are then used to train a supervised machine learning algorithm that can identify the stroke type in new patients. In this technology, classification accuracy depends on the feature extraction method and the classification algorithm. These factors are explored in more detail in the following review of relevant literature.

In [5], the subspace distance classification (SDC) algorithm was employed for stroke classification. The data were obtained from real patients in two clinical cohorts: one with 20 patients and the other with 25. The area under curve (AUC) values for hemorrhagic-versus-ischemic stroke classification for the two cohorts were reported to be 0.88 and 0.85, respectively. In [6], the SDC algorithm was applied with the aim of detecting subdural hematoma using broadband microwave technology. The dataset was obtained

from four human-cranium phantoms with different hematoma volumes (0, 40, 70, and 110 mL) and 1500 simulated observations. The classification accuracies were reported to be 100% for the phantom measurements and between 82% and 96% for the simulated data. A similar technique was also employed in [7], where the goal was to differentiate between healthy brains and those affected by hemorrhagic stroke. The dataset was generated using three-dimensional numerical phantoms derived from a single MRI-based head model. At a noise level of –100 dB, the classification achieved an AUC of 0.93. In [8], a support vector machine (SVM)-based approach was used to detect stroke presence within a simplified three-dimensional physical head phantom. Remarkably, the system achieved 100% detection accuracy, despite being trained on a limited dataset. This work was further extended in [9], where, in addition to stroke type classification, stroke localization was also performed. Both classification and localization achieved a perfect accuracy, reaching 100%. In [10], the SVM classification method, combined with Principal Component Analysis (PCA), was employed to classify data into three categories: ischemic stroke, hemorrhagic stroke, and no stroke, under various scenarios. The dataset was derived from ten two-dimensional numerical phantoms, each containing five distinct tissue types. In one specific scenario, the classification accuracy was reported to be 64.3%. In [11], a similar approach was adopted, utilizing ten three-dimensional numerical phantoms with four distinct tissue types, resulting in an improved classification accuracy of 86%. In [12], scattering parameters were transformed into electric fields, which were then reduced using PCA and subsequently used for stroke classification via Linear Discriminant Analysis (LDA). The study employed a three-dimensional physical phantom and achieved a classification accuracy of 90%. In [13], three algorithms, SVM, Multilayer Perceptron (MLP), and k-Nearest Neighbors (k-NN), were employed to determine the presence or absence of stroke. The dataset was collected from a three-dimensional numerical phantom containing seven different tissue types. Among the algorithms, MLP and SVM demonstrated the most favorable performance, each achieving an accuracy of 100%, whereas k-NN achieved a lower accuracy of 67%. In [14], classification of both ischemic and hemorrhagic strokes was performed using an SVM model in combination with mutual information derived from computed graph degrees. These graph degrees, extracted from signal data, were used as classification features. The dataset was based on ten MRI-derived three dimensional numerical phantoms. The classification achieved an accuracy of 89%. In [15], the Euclidean distance of the reflected signal was used as input to an SVM classifier. The dataset was obtained from ten real rabbit heads, in which the induced strokes varied in both size and position. The classification achieved an accuracy of 96%. In [16], a deep learning technique was employed for hemorrhagic stroke detection. The dataset was obtained using two-dimensional stochastic head models containing five distinct tissue types. This approach achieved a classification accuracy of 94%. In [17], a wavelet convolutional neural network (CNN) was employed for feature extraction and stroke classification. The dataset was generated using 124 MRI-based three-dimensional numerical phantoms. The classification accuracy was 84.38% for noise-free data and 81.65% for noisy data. In [18], the goal was to detect the presence or absence of stroke using a deep learning approach. The dataset was obtained from real human subjects, comprising 235 stroke patients and 196 healthy individuals. The deep learning model takes three types of input: wavelet-transformed microwave signals, clinical information (age), and system configuration parameters (antenna-to-head proximity). The proposed model achieved a classification accuracy of 76.19%. The main characteristics and reported performance of the reviewed microwave-based stroke classification studies are summarized in Table I.

In previous studies, differences in scattering parameters results from brains affected by hemorrhagic and ischemic strokes have been considered as the primary feature for stroke classification. The main drawback in these studies, particularly in those employing MRI-based head phantoms or using human-subject data, is the relatively low classification accuracy. To address this limitation, in this study we propose a new feature extraction method based on early stroke volume change, aiming to enhance the classification performance. Hemorrhagic stroke regions typically exhibit a significant increase in volume during the early hours after onset, whereas the volume of ischemic stroke regions remains relatively constant. This volume change results in corresponding changes in the scattering parameters. Therefore, we measure the scattering data from the patient's head at two different time points and use their difference as the feature for classification. Hence unlike the scattering features used in the reviewed classification methods, the proposed differential feature exploits within-head temporal changes associated with early stroke volume evolution to distinguish between hemorrhagic and ischemic strokes. We will demonstrate that this approach significantly improves classification accuracy.

## II. THE PROPOSED METHOD

### A. THEORY

Aside from the differences in dielectric properties between hemorrhagic and ischemic strokes, another crucial distinction is that the volume of an ischemic stroke remains constant after it occurs, whereas in the case of a hemorrhagic stroke, bleeding can continue during the initial hours following onset.

In the following, we review studies that have investigated this topic in detail. Following an ischemic stroke, the affected brain tissue can be categorized into two distinct regions based on cerebral blood flow thresholds: the ischemic core and the penumbra.

TABLE I. SUMMARY OF EXISTING MICROWAVE-BASED STROKE CLASSIFICATION STUDIES.

| Ref. | Dataset type | Feature | Classifier method | Classification task | Reported performance |
|---|---|---|---|---|---|
| [6] | 2D numerical phantom And Physical phantom | Scattering parameters | SDC | Subdural hematoma detection | 100% accuracy (physical phantom); 82–96% accuracy (numerical phantom) |
| [8] | Physical phantom | Scattering parameters | SVM | Stroke detection | 100% accuracy |
| [9] | Physical phantom | Scattering parameters | SVM | Stroke type and position classification | 100% accuracy for both tasks |
| [12] | Physical phantom | PCA-reduced scattering electric fields | LDA | Stroke type classification | 90% accuracy |
| [10] | 2D numerical phantom | PCA | SVM | Stroke detection and type classification | 64.3% accuracy |
| [16] | 2D numerical phantom | Scattering parameters | Deep learning | Hemorrhagic stroke detection | 94% accuracy |
| [11] | 3D numerical phantom | PCA | SVM | Stroke type classification | 86% accuracy |
| [13] | 3D numerical phantom | Scattering parameters | SVM, MLP, k-NN | Stroke detection | 100% accuracy (SVM and MLP); 67% accuracy (k-NN) |
| [7] | 3D numerical phantom | Scattering parameters | SDC | Hemorrhagic stroke detection | AUC = 0.93 at a noise level of −100 dB |
| [14] | MRI-based 3D numerical phantom | GDMI features | SVM | Stroke type classification | 89% accuracy |
| [17] | MRI-based 3D numerical phantom | Wavelet scalogram features | CNN | Stroke type classification | 84.38% accuracy (without noise); 81.65% accuracy (average SNR = 38.7 dB) |
| [15] | Animal data | Reflection signal Euclidean distance | SVM | Stroke type classification | 96% accuracy |
| [5] | Clinical data | Logarithm of signals | SDC | Stroke type classification | AUC = 0.88 and 0.85 |
| [18] | Clinical data | Wavelet scalograms + age + antenna configuration | Deep learning | Stroke detection | 76.19% accuracy |

Fig. 1 illustrates that the transformation of the penumbra into the ischemic core occurs gradually, with the overall stroke volume remaining relatively stable over time. Positron Emission Tomography (PET) studies have shown that this conversion process takes approximately 18 hours [19]. Several studies have examined changes in hemorrhagic stroke volume within the first hours after onset. Below, we review the literature focusing on those with a sufficiently large patient population. In [20] it was reported that at least 26% of 103 patients experienced more than a 33% increase in hemorrhagic volume between the baseline and the 1-hour CT scans. In [21], 39% of 105 patients showed more than a 33% increase in hemorrhagic volume between the baseline CT (median 74 minutes after stroke onset) and an ultra-early follow-up scan (median 95 minutes later). In [22], out of 419 patients with hemorrhagic stroke, 134 underwent their first CT scan within one hour of symptom onset and had a second CT within 24 hours of admission.

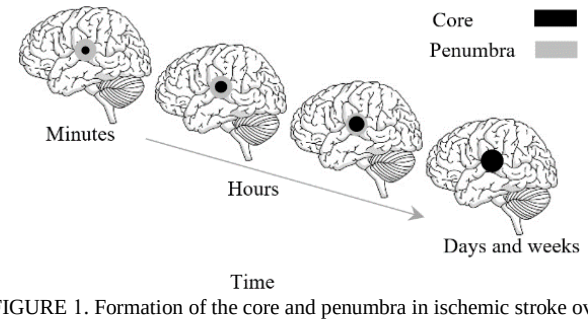



FIGURE 1. Formation of the core and penumbra in ischemic stroke over time[23].

Among them, 26.1% showed more than a 50% increase in hemorrhagic volume. The study demonstrated that the risk of hemorrhagic growth significantly decreases over time, with most of the increase occurring in the early hours, especially within the first six hours. It is worth noting that in [20-22], this hemorrhagic growth occurred even under medical care.

As shown, the volume of ischemic strokes remains almost unchanged during the first few hours after stroke onset, while hemorrhagic strokes tend to increase in volume. Such volumetric changes directly affect the local spatial distribution of dielectric properties within the brain. Accordingly, the difference in the dielectric properties between two time instants $t_0$ and $t_1$ is defined as follows:

$$\Delta\varepsilon(t_0,t_1)=\varepsilon(t_1)-\varepsilon(t_0) \quad (1)$$

Where ε represents the spatial distribution of the complex permittivity at the two time instants $t_0$ and $t_1$. As a result, the corresponding differential permittivity $\Delta\varepsilon(t_0,t_1)$ remains close to zero in ischemic strokes, whereas hemorrhagic strokes exhibit a non-zero $\Delta\varepsilon(t_0,t_1)$ associated with volumetric expansion. As reported in [24], the differential scattering matrix $\Delta S(t_0,t_1)$, obtained by subtracting the scattering parameters measured at two different time instants $t_0$ and $t_1$ is related to the corresponding differential dielectric contrast $\Delta\varepsilon(t_0,t_1)$.

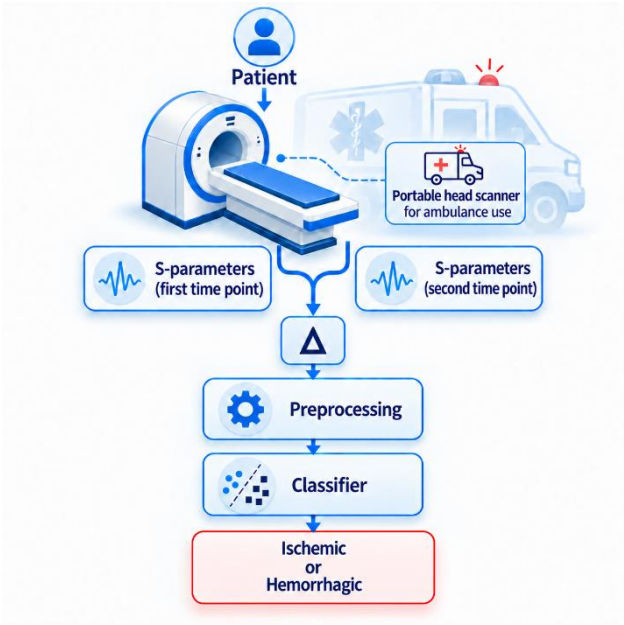



FIGURE 2. Overview of the proposed differential scattering–based framework for stroke type classification.

Therefore, it can be inferred that $\Delta S(t_0, t_1) \approx 0$ for ischemic strokes, in the noiseless case, while $\Delta S(t_0, t_1) \neq 0$ for hemorrhagic strokes.

Accordingly, a stroke type classification framework is proposed, as illustrated in Fig. 2. Within this framework, the head scattering parameters are measured at two successive time points, and their subtraction is used, after preprocessing, as input to a classifier for stroke type discrimination. The specific preprocessing steps and classifier configuration used for evaluation are detailed in Section III-B.

### B. ANTENNA DESIGN

The sensing element employed in the proposed microwave head scanner is a modified printed monopole antenna. The geometry and dimensions of the antenna are presented in Fig. 3[25].

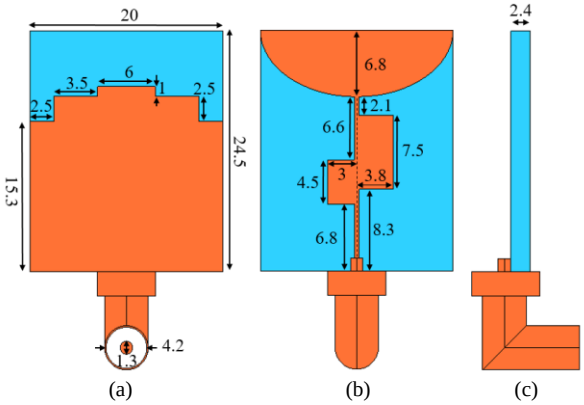



FIGURE 3. Geometry of the antenna. (a) Back view. (b) Front view. (c) Side view (Unit: mm).

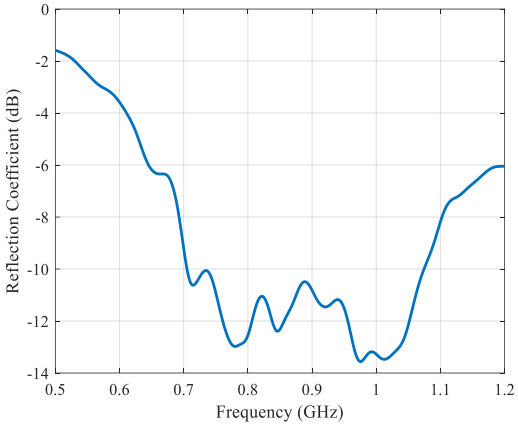



FIGURE 4. Reflection coefficient of antenna in matching medium.

The antenna is printed on an FR-4 substrate with a thickness of 2.4 mm and consists of a semi-elliptical radiating element connected to a microstrip feed line with two matching stubs, with a modified ground plane. The antenna is numerically evaluated using COMSOL Multiphysics [26] in a matching medium with a relative permittivity of 40 and a conductivity of 0.1 S/m. The simulated reflection coefficient of the antenna is evaluated over the frequency range of 0.5–1.2 GHz, as shown in Fig. 4. The antenna exhibits a reflection coefficient below −10 dB over approximately 0.7–1.1 GHz, indicating satisfactory impedance matching over this frequency range.

### C. SETUP CONFIGURATION

The setup configuration utilized to obtain the scattering parameters from the brain is shown in Fig. 5. The operating frequency of this system is 1 GHz, selected to balance the penetration of low-power microwave signals into the head with achieving acceptable accuracy [11] .This configuration consists of 21 identical antenna elements described in the previous subsection and is simulated using COMSOL Multiphysics software. All 21 antennas are included simultaneously in the full-wave model, where each antenna port is excited sequentially using a port sweep and the responses at all 21 ports are recorded, yielding a complete complex 21×21 scattering matrix for each simulation. The antennas are arranged in three concentric rings, such that the two upper rings each consist of eight antennas, while in the first ring three antennas located near the eye region are removed to meet safety considerations. The antennas are placed with an angular spacing of 45° on each ring, and a vertical spacing of 36 mm is maintained between the antenna port centers. The detailed antennas arrangement, including the vertical position, ring radius, and azimuth angle of each antenna port, is provided in Table II. As shown in Fig. 5, to enhance the penetration of electromagnetic waves into the brain, the antennas and the head are placed within a cylindrical matching medium with a relative permittivity of 45 and a conductivity of 0.1 S/m.

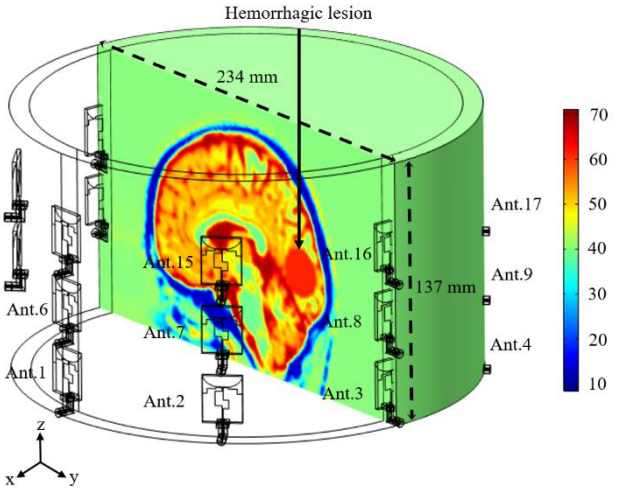



FIGURE 5. A three-dimensional representation of the geometric arrangement of the head scanner.

TABLE II. ANTENNAS ARRANGEMENT.

| Ring | z (mm) | Radius (mm) | Port number and azimuth angle |
|---|---|---|---|
| 1 | 2.25 | 120.25 | 1: 0°, 2: 45°, 3: 90°, 4: 135°, 5: 180° |
| 2 | 38.25 | 120.25 | 6: 0°, 7: 45°, 8: 90°, 9: 135°, 10: 180°, 11: 225°, 12: 270°, 13: 315° |
| 3 | 74.25 | 120.25 | 14: 0°, 15: 45°, 16: 90°, 17: 135°, 18: 180°, 19: 225°, 20: 270°, 21: 315° |

The dimensions of this cylinder are selected to exceed those of the largest head considered, ensuring that the antenna arrangement remains unchanged in each simulation.

### D. DATASET GENERATION AND NOISE MODELING

To create realistic numerical models of the human head, we use data from the BrainWeb database [27]. This database provides 20 phantoms, each represented as a three-dimensional head model in a matrix with dimensions of 362 × 434 × 362. In the base model, each tissue type is assigned a unique numerical identifier, which is then mapped to its dielectric properties using the dataset from [28] at a frequency of 1 GHz. Stroke regions are incorporated as spherical volumes, with their dielectric properties adjusted accordingly: hemorrhagic stroke is modeled using $\varepsilon = 61.1$ and $\sigma = 1.58$ [14], while ischemic stroke is represented by a 10–25% reduction in dielectric parameter values relative to surrounding healthy tissue [29]. Table III summarizes the numerical identifiers and associated dielectric properties for all tissue types included in the model.

Since data-driven methods require a large amount of data and the BrainWeb database contains only 20 different brain models, nine types of random modifications are applied to each head model to enhance diversity. These modifications include scaling the head size in different directions, as detailed in Table IV, as well as random shifts of up to 2 cm in the x, y, and z directions. Additionally, the stroke location and radius are randomly assigned, with the stroke radius ranging between 1 and 2 cm. A random rotation around the z-axis, with angles ranging from -5 to 5 degrees, is also applied. This approach ensures a diverse range of head and stroke configurations. As explained in Section II-A, the volume of hemorrhagic stroke increases in the early hours following its occurrence, so two simulations with different volumes are required to generate the hemorrhagic stroke dataset. In these simulations, the head's position relative to the antenna is assumed to remain constant, with the only difference between the initial and secondary simulations being the stroke volumes. As a result, we generate 2000 hemorrhagic stroke samples: half correspond to the baseline condition and the other half to the increased hemorrhagic volume. The distribution of volume growth used in the simulations is based on the clinical findings in [20] and is illustrated in Fig. 6. For ischemic stroke, where the volume is constant, only one simulation is performed for each case, which results in 1000 data samples.

TABLE III. DIELECTRIC PROPERTIES OF BRAIN REGIONS IN HUMAN HEAD MODELS AT FREQUENCY 1 GHZ [28].

| Tissue Layer | Tissue Label in Brain Web dataset | $\varepsilon_r$ | $\sigma(s/m)$ |
|---|---|---|---|
| Background | 0 | 40 | 0.1 |
| CSF | 16 | 68.4 | 2.46 |
| Gray Matter | 32 | 52.3 | 0.985 |
| White Matter | 48 | 38.6 | 0.622 |
| Fat | 64 | 11.3 | 0.116 |
| Muscle | 80 | 54.8 | 0.978 |
| Skin | 96 | 40.9 | 0.9 |
| Skull | 112 | 12.4 | 0.156 |
| Vessels | 128 | 44.6 | 0.729 |
| Around fat | 145 | 11.3 | 0.116 |
| Dura matter | 161 | 44.2 | 0.993 |
| Bone marrow | 177 | 11.2 | 0.239 |

TABLE IV. HEAD SIZE DATA (UNIT: MM) EMPLOYED IN THIS STUDY [7].

| Direction | Minimum | Maximum |
|---|---|---|
| Ear-Ear | 132 | 169 |
| Forehead-Neck | 162 | 214 |
| Top-Below Eye | 109 | 137 |

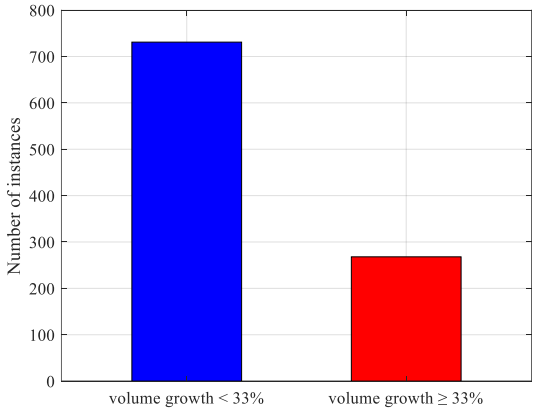



FIGURE 6. Changes in the volumes of hemorrhagic stroke at two different simulation times.

Two distinct time points are then represented by applying independent noise to the simulated data. To reflect more realistic measurement conditions, noise is added to the amplitude and phase of the simulated S-parameters for both hemorrhagic and ischemic datasets**,** based on the model described in [7], where the noisy signal is defined as:

$$s_{noisy}(\omega) = [A(\omega) + A_n(\omega)]\exp[i(\varphi(\omega) + \varphi_n(\omega))] \quad (2)$$

Here, the amplitude noise $A_n(\omega)$ is given by $\sigma N$ and the phase noise $\varphi_n(\omega)$ is given by $\arctan(\frac{\sigma N}{|A(\omega)|})$ where $N$ is a random number drawn from a standard normal distribution with mean 0 and variance 1, and $\sigma$ is the desired standard deviation of the added noise. The standard deviation is chosen such that $\sigma = 10^{(n_L/20)}$, where $n_L$ is the noise level expressed in dB. Two noise levels are selected to reflect the typical performance of a medium-quality vector network analyzer [30]: –115 dB, which corresponds to the analyzer's noise floor, and –95 dB, which is approximately 20 dB higher and represents a more challenging noise condition [13].

## III. EVALUATION AND RESULTS

### A. EARLY EVALUATION

In this section, under a controlled scenario, the effect of stroke volume variation on the differential scattering matrix in hemorrhagic stroke, as well as the effect of noise on the differential matrix in both hemorrhagic and ischemic strokes, is investigated.

To establish a controlled setting, a fixed stroke location is considered in which the initial stroke size is identical for both stroke types at the first time instant. At the second time instant, the radius of the hemorrhagic stroke is increased by 25%, corresponding to approximately a 95% increase in volume, which represents the largest volume growth considered in the present simulation dataset. In contrast, the ischemic stroke size remains unchanged. All other parameters, including head position, dielectric properties, and antenna configuration, are kept constant to ensure that any observed variation in the differential scattering matrix is solely attributed to volume change or noise.

Figs. 7(a) and 7(b) illustrate the simulated magnitudes of the scattering parameters for the hemorrhagic stroke case under noiseless conditions at the first and second time instants, respectively. As observed, despite the significant volume increase, the differences between the two scattering responses are not visually distinguishable. The corresponding differential scattering response is presented in Fig. 7(c). It should be noted that the display scale in this figure is intentionally adjusted; otherwise, these differences would not be observable in a direct comparison. The maximum magnitude of the differential variations is on the order of $16\times10^{-5}$.

Figs. 7(d)–7(f) present the corresponding results under noisy conditions at a noise level of −95 dB. A comparison between Figs. 7(f) and 7(c) indicates that the magnitude of the differential variations remains on the order of $16\times10^{-5}$. However, in the noiseless case, the maximum values are predominantly concentrated around the channels located near the stroke region, whose position is illustrated in Fig. 5. In contrast, under noisy conditions, the distribution of these variations becomes more dispersed and exhibits a less structured pattern.

Figs. 7(g)–7(i) present the noisy ischemic stroke case at −95 dB, where Fig. 7(i) corresponds to the differential scattering matrix. By comparing Figs. 7(i) and 7(f), the distinction between the hemorrhagic and ischemic differential matrices is to some extent visually discernible; specifically, the ischemic differential matrix contains a larger number of near-zero entries, whereas the hemorrhagic differential matrix exhibits more pronounced non-zero variations. However, in scenarios involving smaller volume changes, the visual difference between the differential matrices would be less pronounced, and therefore, the use of data-driven approaches for reliable classification becomes necessary.

### B. NUMERICAL RESULTS

In this study, three classifiers, namely SVM, MLP, and k-NN, are employed to evaluate the performance of the proposed feature. The SVM with a Gaussian kernel is used as the primary classifier due to the limited size of the available dataset, as well as its widespread use in related studies reported in the literature. In addition, MLP and k-NN are included to evaluate whether the effectiveness of the proposed feature generalizes across different classification algorithms. It is emphasized that the focus of this work is not on comparing the classifiers themselves; rather, these classifiers are used as evaluation tools to assess the discriminative capability and generalizability of the proposed feature.

The numerical models, as previously mentioned, include 21 antenna elements in total. Consequently, each simulation generates a complex S-matrix with dimensions of 21 by 21. Due to the Lorentz reciprocity principle, where $s_{i,j} = s_{j,i}$ this S-matrix contains 231 unique and independent elements. These elements are then reshaped into 1×231 vectors for further processing. Subsequently, by stacking 1000 such vectors vertically, the dataset matrices are formed with a size of 1000×231.Therefore, the following matrices are available: two 1000×231 matrices for the hemorrhagic stroke dataset at the first and second time points, as well as a 1000×231 matrix for the ischemic stroke dataset. Note that, by applying independent noise to the ischemic scattering data, two separate matrix sets are generated to represent two distinct time points. Subsequently, the differential matrices are computed by performing row-wise subtraction between the matched pairs at the two time points for both ischemic and hemorrhagic cases.

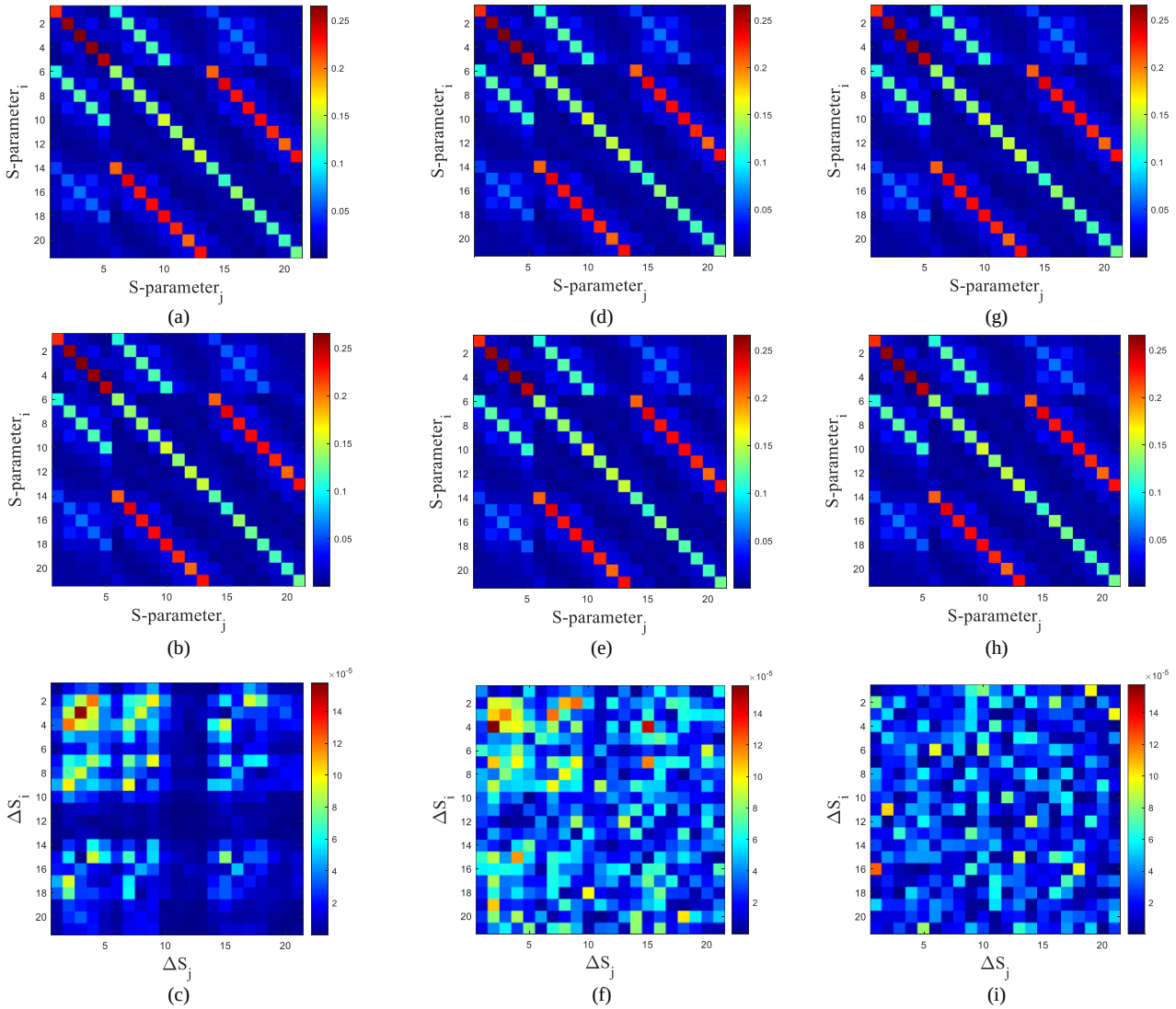



FIGURE 7. Simulated scattering-parameter magnitudes and corresponding differential responses in the controlled scenario. (a)–(c) Hemorrhagic stroke under noiseless conditions: first time instant, second time instant (25% radius increase ≈ 95% volume growth), and the magnitude of the corresponding differential scattering matrix, respectively. (d)– (f) Hemorrhagic stroke at a noise level of −95 dB. (g)– (i) Ischemic stroke at a noise level of −95 dB.

Since the elements of these matrices are complex-valued, they must be transformed into a format suitable for machine-learning-based classification. For this purpose, each complex number is represented by its magnitude and phase. Specifically, the magnitudes of the 231 coefficients in each row form the first 231 columns of the feature matrix, while the corresponding phases constitute the remaining 231 columns. The resulting 1000×462 differential feature matrices for the ischemic and hemorrhagic cases are then used as input to the selected classifiers.

Nested cross-validation is used to separate hyperparameter optimization from performance evaluation. The outer loop estimates classification performance, while the inner loop is used for hyperparameter optimization. In the outer loop, the dataset is divided into five folds, keeping all samples generated from the same BrainWeb head model in a single fold to prevent data leakage. In each iteration, four folds form the outer training set, and the remaining fold is held out as the outer test set. Bayesian hyperparameter optimization is then performed within the outer training set using three-fold inner cross-validation. The inner folds are also grouped by head model, so samples from the same head model cannot appear in both the inner training and validation sets. After hyperparameter optimization, the model is refit on the complete outer training set using the selected hyperparameters. The held-out outer test set is used only for performance evaluation and is not involved in hyperparameter optimization or model selection.

TABLE V. EVALUATION OF THE PROPOSED DIFFERENTIAL FEATURE ACROSS DIFFERENT CLASSIFIERS.

| Classifier | Feature | Noise level (dB) | Accuracy (%) | Specificity (%) | Sensitivity (%) |
|---|---|---|---|---|---|
| SVM | — | −115 | 62.96 | 68.12 | 57.80 |
| | | −95 | 61.87 | 67.72 | 56.02 |
| | Proposed method | **−115** | **97.11** | **99.30** | **94.92** |
| | | **−95** | **78.55** | **85.42** | **71.68** |
| MLP | — | −115 | 63.12 | 66.64 | 59.60 |
| | | −95 | 62.75 | 67.66 | 57.84 |
| | Proposed method | **−115** | **95.94** | **97.96** | **93.92** |
| | | **−95** | **75.42** | **76.42** | **74.42** |
| k-NN | — | −115 | 54.03 | 58.44 | 49.62 |
| | | −95 | 53.77 | 59.10 | 48.44 |
| | Proposed method | **−115** | **90.20** | **98.74** | **81.66** |
| | | **−95** | **74.04** | **83.44** | **64.64** |

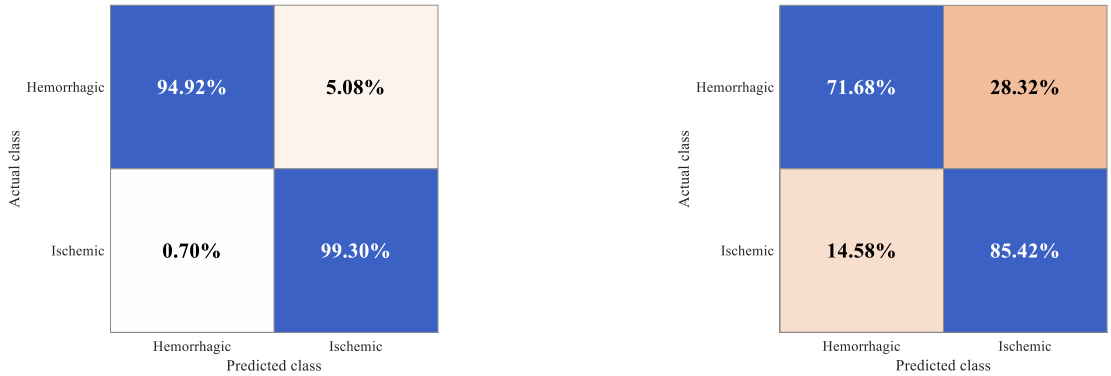



FIGURE 8. Normalized confusion matrices (%) of the proposed method at noise levels of (a) −115 dB and (b) −95 dB, averaged over five nested cross-validation runs with independent noise realizations.

Classification performance is evaluated using accuracy, sensitivity, and specificity, calculated as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (3)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100\% \quad (4)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \times 100\% \quad (5)$$

where TP, TN, FP, and FN denote true positive, true negative, false positive, and false negative predictions, respectively. Hemorrhagic and ischemic strokes are considered the positive and negative classes, respectively. Thus, sensitivity represents the proportion of hemorrhagic cases correctly classified, while specificity represents the proportion of ischemic cases correctly classified.

At each noise level, the nested cross-validation procedure is repeated five times with independently generated noise. In each repetition, performance metrics are calculated using the test predictions for all samples. The final results are reported as the mean of each metric over the five repetitions.

Table V reports the performance obtained using the proposed differential feature and S-parameters from the second time point alone with SVM, MLP, and k-NN classifiers. The differential feature improves classification performance across all three classifiers and both noise levels, supporting its effectiveness beyond the primary SVM classifier. Fig. 8 shows the normalized confusion matrices for the proposed method using SVM at both noise levels.

### C. EFFECT OF HEMATOMA VOLUME GROWTH

The effect of hematoma volume growth on the proposed differential features is evaluated using the outer-test predictions already obtained from the nested cross-validation procedure by grouping hemorrhagic cases into six intervals of relative volume increase (%): (0, 5], (5, 10], (10, 20], (20, 33], (33, 50], and (50, 100]. For each interval, the sensitivity of the primary SVM classifier is calculated from these outer-test predictions at noise levels of −115 and −95 dB. Results are reported as the mean ± standard deviation across the five independent noise realizations.

As shown in Fig. 9, sensitivity increases with hematoma volume growth at both noise levels.

At −115 dB, mean sensitivity rises from 74.05% in the lowest growth interval to approximately 100% for larger expansions. At −95 dB, it increases from 38.92% to 95.47%. These results indicate that larger volume changes produce more distinguishable differential scattering responses, whereas small changes are more difficult to detect, particularly at the higher noise level.

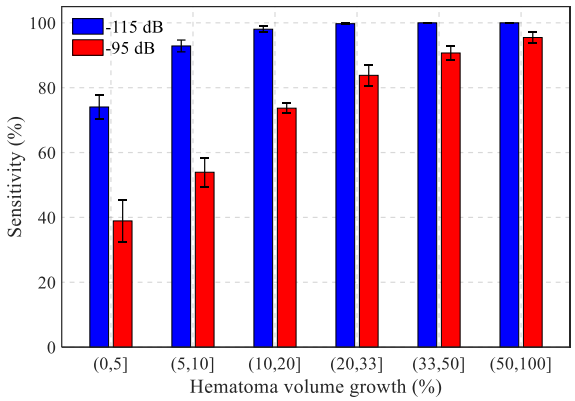


FIGURE 9. Sensitivity across different hematoma volume-growth intervals at −115 and −95 dB.

The no-growth case represents an inherent limitation of the proposed approach. Under the present simulation assumptions, differences between the two measurements in hemorrhagic cases without hematoma expansion arise solely from noise. Consequently, the differential feature provides no class-specific temporal information, and such cases may be misclassified as ischemic.

## IV. DISCUSSION

### A. COMPARISON WITH EXISTING FEATURE EXTRACTION APPROACHES

To quantitatively evaluate the proposed feature extraction method, selected feature extraction approaches from previous microwave-based stroke classification studies are implemented and evaluated on the present dataset. These approaches are selected based on two criteria: they have previously been used for microwave-based stroke classification and are directly applicable to the complex scattering data generated in this study without requiring additional measurements or input modalities. Accordingly, real/imaginary S-parameters, following [9] and [13], and real/imaginary S-parameters combined with PCA, following [10] and [11], are selected for comparison.

To isolate the effect of feature extraction, all approaches are evaluated using the same SVM classifier, dataset, noise conditions, data-splitting strategy, and nested cross-validation procedure described in Section III-B. As shown in Table VI, the real/imaginary S-parameter representation achieves accuracies of 62.51% and 62.43% at noise levels of −115 and −95 dB, respectively, while the PCA-based approach achieves 62.19% and 61.23%. Interestingly, this performance range is close to the 64.63% accuracy reported in the clinical study in [18] when only the microwave signal input was used, whereas the reported accuracy increased to 76.19% when additional clinical and antenna-configuration information were included. Although the simulated and clinical datasets are not directly comparable, this similarity provides additional context for the relatively modest performance of the selected existing approaches.

TABLE VI. COMPARISON OF SELECTED FEATURE EXTRACTION METHODS ON THE PRESENT DATASET.

| Feature extraction approach | Noise (dB) | Accuracy (%) | Specificity (%) | Sensitivity (%) |
|---|---|---|---|---|
| Real/imaginary S-parameters [9], [13] | −115 | 62.51 | 65.98 | 59.04 |
| | −95 | 62.43 | 68.46 | 56.40 |
| Real/imaginary S-parameters + PCA [10], [11] | −115 | 62.19 | 66.60 | 57.78 |
| | −95 | 61.23 | 65.36 | 57.10 |
| Proposed differential magnitude/phase | **−115** | **97.11** | **99.30** | **94.92** |
| | **−95** | **78.55** | **85.42** | **71.68** |

In contrast, the proposed differential magnitude/phase feature achieves accuracies of 97.11% and 78.55% at the same noise levels, corresponding to improvements of 34.60 and 16.12 percentage, respectively, compared with the best-performing selected existing approach.

This reduction in accuracy indicates that the proposed differential feature is sensitive to measurement noise; however, it still maintains higher classification performance than the selected comparison approaches under the more challenging noise condition. To deal with this, signal enhancement techniques can be used to improve the signal-to-noise ratio. Also, using a better-quality vector network analyzer can help reduce the noise, which makes the classification more reliable in real-world situations.

### B. PRACTICAL IMPLEMENTATION CONSIDERATIONS

The proposed feature extraction method is computationally simple, requiring only subtraction of two scattering matrices, selection of the independent S-parameters, and conversion to magnitude- and phase-based features. Since model training is performed offline, the total processing time is on the order of milliseconds, indicating that computational processing is unlikely to limit practical real-time implementation.

Beyond computational processing time, the proposed method also requires an interval between the two scattering measurements to capture hematoma growth. To estimate this interval, the hematoma volume changes in the simulated data are mapped to clinically reported hematoma growth rates. According to [21], the median hematoma growth rate from stroke onset to initial imaging is 10.3 mL/h. Using this rate, the simulated hematoma volume increments in our dataset are converted into estimated inter-measurement intervals. The median estimated interval is approximately 4 min, and about 80% of the simulated hemorrhagic cases correspond to intervals of 10 min or less. These estimations suggest that the required inter-measurement interval may be compatible with prehospital use.

The results also depend on how the volume of hemorrhagic stroke changes over time, as described in [20]. If the actual growth pattern is different, the classification accuracy might be affected. Future work

should consider using a wider range of growth patterns to make the method more general and reliable.

Finally, although this study assumes that the head remains stationary between the two measurements, maintaining this condition in clinical environments can be challenging. To address this, [31] introduces a flexible conformal antenna array that attaches directly to the head, enabling consistent antenna placement even during patient transport.

## V. CONCLUSION

Accurate classification of stroke type is a critical and clinically relevant task. This study presented a new and high-performing feature extraction method based on the change in hemorrhagic stroke volume over time, while the ischemic stroke volume remains nearly constant, which affects the microwave scattering parameters reflected from the head between two measurement time points. Therefore, the head's scattering parameters are measured at two distinct time points, initially at baseline and subsequently after a delay. The subtraction between these measurements is then computed and used as input to a supervised classifier trained to identify the stroke type.

A synthetic dataset was generated based on 20 MRI-based three-dimensional numerical human head models. In total, 3000 samples were produced, comprising 2000 hemorrhagic samples corresponding to 1000 paired cases and 1000 ischemic samples. For ischemic stroke, where the stroke volume was assumed to remain constant over time, a single baseline simulation was performed for each case, and two distinct time points were represented by applying independent noise to the simulated data.

Using SVM as the primary classifier, together with MLP and k-NN to assess the generalizability of the proposed feature across different classifiers, the proposed differential feature consistently outperformed the corresponding S-parameter features from the second time point alone across the evaluated noise levels. Using the primary SVM classifier, average accuracies of 97.11% and 78.55% were achieved at −115 dB and −95 dB, respectively.

These findings confirm that exploiting the difference between successive scattering measurements can significantly enhance stroke type classification performance. The main technical contribution of the proposed approach is the use of temporal differential scattering information to emphasize stroke-related changes while reducing the contribution of static anatomical structures. From a practical perspective, the low computational cost and the relatively short estimated interval between successive measurements support the potential use of the proposed approach in portable prehospital systems for early stroke-type assessment.

To fully evaluate the clinical effectiveness of the proposed approach, validation using measurements from real patients will be necessary.


## Funding

The authors received no financial support for the research, authorship, and/or publication of this article.


## Conflicts of Interest

The authors declare no conflicts of interest.

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## REFERENCES


[1] E. J. Benjamin *et al.*, "Heart disease and stroke statistics—2019 update: a report from the American Heart Association," *Circulation*, vol. 139, no. 10, pp. e56-e528, 2019.

[2] K. Boswell, "Emergency Department Critical Care," in *Emergency Department Critical Care*, J. R. Shiber Ed. Springer Nature Switzerland AG: Springer, 2020.

[3] L. Guo, A. S. Alqadami, and A. Abbosh, "Stroke Diagnosis Using Microwave Techniques: Review of Systems and Algorithms," *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, 2022.

[4] J.-C. Chiao *et al.*, "Applications of Microwaves in Medicine," *IEEE Journal of Microwaves*, vol. 3, no. 1, pp. 134-169, 2022.

[5] M. Persson *et al.*, "Microwave-based stroke diagnosis making global prehospital thrombolytic treatment possible," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 11, pp. 2806-2817, 2014.

[6] S. Candefjord *et al.*, "Microwave technology for detecting traumatic intracranial bleedings: tests on phantom of subdural hematoma and numerical simulations," *Medical & biological engineering & computing*, vol. 55, no. 8, pp. 1177-1188, 2017.

[7] A. Fhager, S. Candefjord, M. Elam, and M. Persson, "3D simulations of intracerebral hemorrhage detection using broadband microwave technology," *Sensors*, vol. 19, no. 16, p. 3482, 2019.

[8] M. Salucci, J. Vrba, I. Merunka, and A. Massa, "Real-time brain stroke detection through a learning-by-examples technique—An experimental assessment," *Microwave and Optical Technology Letters*, vol. 59, no. 11, pp. 2796-2799, 2017.

[9] M. Salucci, A. Gelmini, J. Vrba, I. Merunka, G. Oliveri, and P. Rocca, "Instantaneous brain stroke classification and localization from real scattering data," *Microwave and Optical Technology Letters*, vol. 61, no. 3, pp. 805-808, 2019.

[10] T. Pokorny *et al.*, "On the role of training data for svm-based microwave brain stroke detection and classification," *Sensors*, vol. 23, no. 4, p. 2031, 2023.

[11] T. Pokorny, D. Vrba, O. Fiser, M. Salucci, and J. Vrba, "Systematic Optimization of Training and Setting of SVM-Based Microwave Stroke Classification: Numerical Simulations for 10 Port System," *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, 2024.

[12] Z. Gong, Y. Ding, Y. Chen, and M. J. Cree, "Wearable Microwave Medical Sensing for Stroke Classification and Localization: A Space-Division-Based Decision Tree Learning Method," *IEEE Transactions on Antennas and Propagation*, 2023.

[13] V. Mariano, J. A. Tobon Vasquez, M. R. Casu, and F. Vipiana, "Brain Stroke Classification via Machine Learning Algorithms Trained with a Linearized Scattering Operator," *Diagnostics*, vol. 13, no. 1, p. 23, 2022.

[14] 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 Microwaves in Medicine and Biology*, vol. 5, no. 1, pp. 46-53, 2020.

[15] J. Xu *et al.*, "Noninvasive and portable stroke type discrimination and progress monitoring based on a multichannel microwave transmitting–receiving system," *Scientific Reports*, vol. 10, no. 1, p. 21647, 2020.

[16] L. Alon and S. Dehkharghani, "A stroke detection and discrimination framework using broadband microwave scattering on stochastic models with deep learning," *Scientific Reports*, vol. 11, no. 1, p. 24222, 2021.
[17] S. Hasan, A. Zamani, A. Brankovic, K. Bialkowski, and A. Abbosh, "Stroke Classification with Microwave Signals using Explainable Wavelet Convolutional Neural Network," *IEEE Journal of Biomedical and Health Informatics*, 2023.
[18] W. Yuan, P. Thammasorn, L. Wang, and S. Mo, "A Comprehensive Deep Learning Framework for Microwave Stroke Classification: Combining Signal Analysis, Clinical Variables, and Antenna System Measurements," *IEEE Access*, 2025.
[19] C. M. Ermine, A. Bivard, M. W. Parsons, and J.-C. Baron, "The ischemic penumbra: from concept to reality," *International Journal of Stroke*, vol. 16, no. 5, pp. 497-509, 2021.
[20] T. Brott *et al.*, "Early hemorrhage growth in patients with intracerebral hemorrhage," *Stroke*, vol. 28, no. 1, pp. 1-5, 1997.
[21] C. A. Mutimer *et al.*, "Ultra-early hematoma expansion is associated with ongoing hematoma growth and poor functional outcome," *Stroke*, vol. 56, no. 4, pp. 838-847, 2025.
[22] Y. Fujii, R. Tanaka, S. Takeuchi, T. Koike, T. Minakawa, and O. Sasaki, "Hematoma enlargement in spontaneous intracerebral hemorrhage," *Journal of neurosurgery*, vol. 80, no. 1, pp. 51-57, 1994.
[23] U. Dirnagl, C. Iadecola, and M. A. Moskowitz, "Pathobiology of ischaemic stroke: an integrated view," *Trends in neurosciences*, vol. 22, no. 9, pp. 391-397, 1999.
[24] D. O. Rodriguez-Duarte, C. Origlia, J. A. T. Vasquez, R. Scapaticci, L. Crocco, and F. Vipiana, "Experimental assessment of real-time brain stroke monitoring via a microwave imaging scanner," *IEEE Open Journal of Antennas and Propagation*, vol. 3, pp. 824-835, 2022.
[25] A. Asadi, "Design, Simulation and Fabrication of A Compact Printed Planar Monopole Antenna for Use in Microwave Imaging
System for Stroke Detection " Master of Science, Electrical Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran, 2023.
[26] *COMSOL Multiphysics®*. COMSOL AB, Stockholm, Sweden. Accessed: 9 January 2023
[Online]. Available: http://www.comsol.com
[27] BrainWeb. "Simulated Brain Database." www.bic.mni.mcgill.ca/brainweb (accessed 1 August 2018).
[28] C. Gabriel and S. Gabriel, "Compilation of the dielectric properties of body tissues at RF and microwave frequencies," ed: King's College, London, Department of Physics, 1996.
[29] S. Semenov, T. Huynh, T. Williams, B. Nicholson, and A. Vasilenko, "Dielectric properties of brain tissue at 1 GHz in acute ischemic stroke: Experimental study on swine," *Bioelectromagnetics*, vol. 38, no. 2, pp. 158-163, 2016.
[30] K. Technologies, "Keysight Streamline Series USB Vector Network Analyzer P937XA 2-Port, Up to 26.5 GHz; Data Sheet," Keysight Technologies, Santa Rosa, CA, USA, 2018.
[31] A. S. Alqadami, K. S. Bialkowski, A. T. Mobashsher, and A. M. Abbosh, "Wearable electromagnetic head imaging system using flexible wideband antenna array based on polymer technology for brain stroke diagnosis," *IEEE transactions on biomedical circuits and systems*, vol. 13, no. 1, pp. 124-134, 2018.