This is a repository copy of *Multi-frequency symmetry difference electrical impedance tomography with machine learning for human stroke diagnosis*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/210311/

Version: Accepted Version

Article:

IPEM Institute of Physics and Engineering in Medicine

Multi-frequency symmetry difference electrical impedance tomography with machine learning for human stroke diagnosis

Multi-Frequency Symmetry Difference Electrical Impedance Tomography with Machine Learning for Human Stroke Diagnosis

Barry McDermott*, Adnan Elahi, Adam Santorelli, Martin O'Halloran, James Avery, and Emily Porter

Abstract— Objective: **Multi-Frequency Symmetry Difference Electrical Impedance Tomography (MFSD-EIT) can robustly detect and identify unilateral perturbations in symmetric scenes. Here, an investigation is performed to assess if the algorithm can be successfully applied to identify the aetiology of stroke with the aid of machine learning.** *Methods:* **Anatomically realistic four-layer Finite Element Method models of the head based on stroke patient images are developed and used to generate EIT data over a 5 Hz – 100 Hz frequency range with and without bleed and clot lesions present. Reconstruction generates conductivity maps of each head at each frequency. Application of a quantitative metric assessing changes in symmetry across the sagittal plane of the reconstructed image and over the frequency range allows lesion detection and identification. The algorithm is applied to both simulated and human (n=34 subjects) data. A classification algorithm is applied to the metric value in order to differentiate between normal, haemorrhage and clot values.** *Results:* **An average accuracy of 85% is achieved when MFSD-EIT with Support Vector Machines (SVM) classification is used to identify and differentiate bleed from clot in human data, with 77% accuracy when differentiating normal from stroke in human data.** *Conclusion:* **Applying a classification algorithm to metrics derived from MFSD-EIT images is a novel and promising technique for detection and identification of perturbations in static scenes.** *Significance:* **The MFSD-EIT algorithm used with machine learning gives promising results of lesion detection and identification in challenging conditions like stroke. The results imply feasible translation to human patients.**

Index Terms— **electrical impedance tomography, reconstruction algorithm, stroke imaging**

I. INTRODUCTION

IMPORTANT medical conditions such as stroke feature causative lesions which are essentially static in nature [1]. In stroke, the cause is either a bleed (haemorrhagic stroke) or a clot (ischaemic stroke) [1]. This aetiology must be rapidly and definitively diagnosed before appropriate and highly divergent treatment can be initiated [3]. For example, thrombolytic agents are used to dissolve clots, but have time dependent efficacy, and are contra-indicated in haemorrhagic stroke patients as they are potentially fatal [2]. Stroke diagnosis is achieved by gold standard neuroimaging including computed tomography (CT) and magnetic resonance imaging (MRI) scans. However, challenges with availability of CT and MRI result in delayed imaging and treatment rates as low as 4% [3]. The example of stroke diagnosis highlights the urgent need for a technology that can rapidly differentiate between the two causative lesion types and facilitate early initiation of correct treatment.

Electrical Impedance Tomography (EIT) may represent such a technology. EIT is a low-cost, portable and safe imaging technology [4]–[6]. In EIT, electrodes are placed on the exterior of the body and used to inject innocuous current (of frequencies in the order of Hz to MHz depending on the application), with the resultant voltages measured [4]–[6]. A current injection and voltage measurement pattern is pre-defined as a 'protocol' with the complete set of voltage measurement 'channels' referred to as a 'measurement frame'. Measurement frames are used to reconstruct a map of the interior in terms of the conductivity, σ, of the tissues therein [4]–[6]. EIT when applied to biomedical scenarios has to date enjoyed most success when used in areas featuring a time change [6]. Such applications allow the use of time differencing of the measurement frames and a cancellation of errors to which EIT is highly sensitive [4], [7]–[9].

However, in static scenes (such as in stroke) time differencing is not possible and alternate techniques are required including absolute EIT (aEIT) and frequency difference EIT (fdEIT). aEIT attempts to reconstruct an image from a single measurement frame, and features a high sensitivity to errors [6]. fdEIT differences frames taken at different frequencies where there is a frequency dependent

change in the σ profile of the tissues [6]. The latter has been applied to stroke diagnosis with some success but challenges remain such as sensitivity to errors in electrode position, skull anatomy, and in bleed detection relative to clot detection [9]–[11]. Further challenges exist in applying EIT to stroke, including the attenuation of current due to the highly resistive skull, the shunting effect of the highly conductive cerebrospinal fluid (CSF) layer inside the skull, and the safety limitations on the amplitude of current that can be used (for example a maximum current of 100 µA root mean square for frequencies up to 1 kHz) [5], [12].

As such, the effective application of EIT to static scenes is challenging. However, it is an area where development and innovation is needed. Previously, we demonstrated the application of machine learning (ML) classification algorithms to EIT measurement frames [13], [14]. Significant promise was shown in the use of support vector machine (SVM) classifiers in numerical and experimental studies with regard to the detection and differentiation of normal from static haemorrhagic lesions [13], [14]. A central theme in our previous work has been the move away from images in acceptance of the limits of EIT, where it is unlikely detailed information as to precise location or lesion volume would be possible given the state of the technology at this time [20], [21]. Detection and identification may be possible however, allowing commencement of treatment before gold standard neuroimaging can be employed for more complete patient triage.

Recently we have introduced a novel modality of EIT suitable for static scenes featuring symmetry, which focusses on detection of unilateral perturbations causing a change in this inherent symmetry [15], [16]. This so-called Bi-Frequency Symmetry Difference EIT (BFSD-EIT) was then tested for the application of stroke [17]. The work of [17] helped define the robustness of the technique, maximal allowances with respect to errors, and techniques to improve the efficacy. Specifically, two techniques were identified that could further improve the algorithm efficacy: the use of a quantitative metric called the Global LHS & RHS Mean Intensity (GMI), and a multi-frequency (MF) approach in the band of greatest change in σ of the tissues (≤ 100 Hz [18], [19]).

Therefore, in this work, we advance the utility of BFSD-EIT by implementing these improvements. The number of frequency (f) points considered is increased to achieve a Multi-Frequency Symmetry Difference Electrical Impedance Tomography (MFSD-EIT) algorithm. Then, MFSD-EIT is used to produce GMI metrics at each frequency. This multi-frequency GMI data (MF GMI) is in turn used as the input for a ML classification algorithm. Efficacy of classification is then assessed with respect to identifying and differentiating between normal, haemorrhagic stroke patients, and ischaemic stroke patients.

The human dataset collected by Goren *et al.* in University College London (UCL) is used as the primary data source [3]. Realistic four-layer finite element method (FEM) models constructed from the CT scans of patients are used to generate simulated EIT measurement frames as well as reconstruct and analyse the real measurement frames collected from these patients [3]. The MFSD-EIT algorithm is applied to the simulated and real frames, and the GMI results collected across a band from 5 Hz – 100 Hz. These results are then used as input features for SVM classification to assess if differentiation between stroke types is possible.

The layout of the paper is as follows. In Section II, an overview of the UCL stroke dataset is given followed by a summary of the MFSD-EIT algorithm. Then, the numerical models are described. This section ends with an overview of the generation of MF GMI data from the EIT data. In Section III, the MF GMI data is used as an input to a SVM classification algorithm. Scenarios considered for classification include binary classifications (for example, normal versus lesion, bleed versus clot), as well as multiclass classification. Analysis is performed on both numerical and human data sets with these results then discussed in Section IV. The paper is concluded in Section V.

While the stroke diagnostic problem is the example used in this paper, another motivation for this work is the development of a robust modality of EIT for application in static scenes in general. The extension of EIT into such areas will result in an increased number of biomedical applications where EIT can be effectively employed.

II. METHODOLOGY

In this section, the UCL dataset is presented, as well as a summary of the MFSD-EIT algorithm (including the GMI metric). The rationale for the use of this proposed algorithm in stroke is described, and theoretical ideal results for normal, bleed and ischaemic patients are discussed. Next, a description of the FEM models used in the creation (and reconstruction) of simulated measurement frames and the reconstruction of human frames is presented. The section ends with a description of the MF GMI data obtained from the numerical models and the human frames.

A. UCL MF EIT Stroke Dataset

EIT measurement frames were collected as part of a clinical trial with the Hyper Acute Stroke Unit (HASU) at University College London Hospital (UCLH) [3]. This dataset is the most comprehensive collection of human EIT data related to stroke. EIT measurement frames were collected at 17 frequency points from 5 Hz – 2 kHz, with maximal current adjusted according IEC 60601-1 guidelines [12]. In the published dataset, measurement frames are provided from $N_{Healthy} = 10$ healthy volunteers, and from $N_{Patients} = 18$. Some of the patients had repeat recordings performed. Hence, the final set of recordings comprises $N_{Healthy} = 10$, $N_{Haemorrhagic} = 10$ and $N_{Ischaemic} = 14$ cases. The frames were recorded using the ScouseTom EIT system [22], with 32 electroencephalogram (EEG) electrodes placed on the patient according to EEG 10-20 system with some variants (and identical to the layout described in [16], [17]). In most cases (16 of the 24 patient sets) recording was performed within 48 hours of stroke onset [3]. The injection/ measurement protocol provided 930 separate voltage measurements in a measurement frame at each frequency with the protocol

selected to maximize the magnitude of recorded voltages and the number of independent measurements. The recorded measurement frames had a mean reported signal to noise ratio (SNR) of approximately 45–50 dB which is lower than the recommended minimum 60 dB from the feasibility study of [17]. However, care was taken to maximise quality of recording with removal of voltages excessively contaminated, filtering and other post-processing techniques performed by the UCL investigators [3].

Importantly the electrode layout is symmetric and so the electrodes can be considered as two mirror image orientations, and thus suitable for MFSD-EIT. The protocol used in the UCL dataset was not a MFSD-EIT protocol, but the resultant measurement frames could be post-processed for input in to the MFSD-EIT algorithm.

In addition to the EIT measurement frames, diagnostic imaging with CT and/ or MRI scans of 18 patients is provided in the dataset [3]. These images were collected at a different time point than the EIT recording session and contain the diagnostic report of the radiologist. These neuroimages are used to generate FEM models, described in part E of this section.

B. MFSD-EIT Algorithm and GMI

MFSD-EIT is an extension of BFSD-EIT with the use of multiple (instead of two) f points resulting in a more robust disambiguation of lesion type step. A summary is provided here, with a thorough presentation of the algorithm is given in [16]. Briefly, the algorithm can be considered as consisting of two steps:

(i) Detection in Deviation from Normal Symmetry:

The electrodes are arranged on the body of interest as symmetric pairs with respect to a plane of symmetry. In the case of the head, the sagittal plane divides the head into symmetric left- and right-hand sides (LHS and RHS). Any electrode on the symmetric plane is considered its own pair. A measurement frame is taken from an 'A-orientation' and then from a mirror image 'B-orientation'. A sample channel from both orientations is illustrated in Fig. 1. These frames are differenced and reconstructed using 0th order Tikhonov regularization [23] onto a FEM model corresponding to A-orientation. Each voxel in the reconstruction has a conductivity change assigned. For a given voxel, a positive conductivity change intensity indicates the measurements from B-orientation are more conductive than that from A-orientation at that location. The magnitude of the intensity is proportional to the magnitude of the difference in measurements. If a unilateral perturbation is present, symmetry will be disrupted at the location of the perturbation with an equal but opposite disruption in terms of conductivity change at the symmetric location.

(ii) Disambiguation of Lesion Type

In the case of stroke the two causative lesions of haemorrhage and clot are more and less conductive than the brain respectively. This difference in conductivity results in ambiguity as to which of the disruptions detected from step (i) is the true perturbation, if a perturbation is indeed detected. The
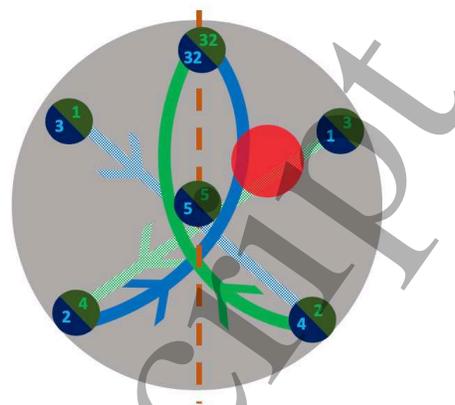
Fig. 1. Illustration of an equivalent channel from A- (green) and B- (blue) orientations. The plane of symmetry is shown as an orange line, with electrodes simultaneously part of A- and B- orientations (for example electrode #1 in A-orientation is also electrode #3 in B-orientation). In the sample channel, current in injected between electrode #2 and #32, with voltage measured between electrode #3 and #4. The 'green' channel represents this channel when A-orientation is used while the 'blue' channel is this channel for B-orientation. The channels are equivalent and theoretically give the same voltage measurement. However, the presence of a perturbation (illustrated as a red circle) upsets the symmetry with this disruption seen as a difference in the voltages recorded between the green and blue channels.

disambiguation step involves repeating step (i) at a different frequency point where there is a known change in the pattern of conductivity of the tissues from the original frequency point. The change in contrast translates as a proportional change in the conductivity difference detected at the candidate perturbation locations. *A priori* knowledge of the change in contrast of the tissues allows identification of the lesion as haemorrhage or clot. The use of multiple frequency points results in a more robust analysis as the pattern of change over a band can be considered as opposed to at only two points [17].

The voxel intensity values in the reconstructions can be analysed and used to generate robust quantitative metrics. The GMI metric has been shown to be particularly robust to errors and can alone be used to identify perturbations [16], [17].

The GMI metric is defined as:
- The average intensity over all the voxels on each side (LHS and RHS) of the sagittal plane. The intensity will have magnitude and negative or positive sign.

At each frequency point the resultant GMI is reported as two separate values, a LHS and a RHS value. If the body under consideration is perfectly symmetrical and no perturbation is present, then the result should be a LHS and RHS GMI both of 0 at any frequency point. The presence of a unilateral perturbation moves the GMI away from 0, with equal but opposite values for LHS and RHS proportional to the contrast between the perturbation and the background brain at that frequency point. The GMI results for all frequency points considered are referred to in this work as MF GMI data sets.

C. Rationale for use in Stroke

In stroke the causative lesion either a bleed or clot in the brain [1]. Bleed is more conductive than brain at a given f point, with clot less conductive in the 5 Hz – 100 Hz range [11], [24]. The

Fig. 2. Conductivity spectra of ischaemia, bleed and brain across the 5 Hz – 100 Hz band (data from [18], [19]). Ischaemia and bleed are respe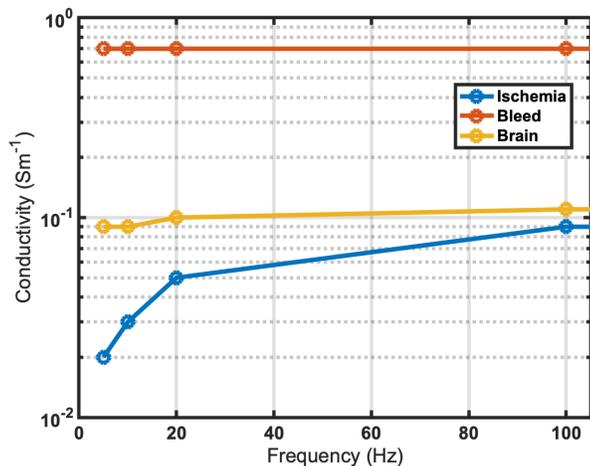ctively less and more conductive than healthy brain. However, across this band the contrast profile in σ between ischaemia and brain is significantly different than that of bleed and brain facilitating disambiguation using MFSD-EIT. The UCL stroke dataset includes measurements taken at 5 Hz, 10 Hz, 20 Hz, and 100 Hz.

Fig. 3. Theoretical MF GMI results for normal (N), bleed (B), and clot (C) cases across the 5 Hz – 100 Hz band at four f points (5 Hz, 10 Hz, 20 Hz, 100 Hz). The GMI for a given case is ideally equally but opposite at each f point and the trend follows the contrast of the tissue with respect to background brain across the band. In the N case the results approach zero (as there is no lesion present). The pattern for B differs to that for C leading to differentiation being possible.

contrast between the tissues in terms of σ shows most change below 100 Hz [11], [24]. These σ spectra from 5 Hz – 100 Hz are shown in Fig. 2. The contrast in σ between ischaemia and brain shows a significant change in across the 5 Hz – 100 Hz band with a high contrast at 5 Hz and low contrast at 100 Hz. Bleed shows an approximate constant contrast with respect to brain across the same band. Hence, theoretically, performing step *(i)* of MFSD-EIT across this band at multiple f points should result in the same GMI result if a bleed is present. If a clot is present the GMI will be maximal at 5 Hz and reduce to the normal (perturbation free) value as 100 Hz is approached and the contrast reduces with respect to the background brain. The UCL stroke dataset includes measurements taken at 5 Hz, 10 Hz, 20 Hz, and 100 Hz with these f points used in this study, and the GMI calculated at each of these four discrete frequencies.

The other tissues of the head (scalp, skull, CSF) have approximately constant σ values in this frequency band [17]–[19].

D. Theoretical Ideal Results

The ideal results for a perfectly symmetrical head in three cases: with no lesion present (normal), a bleed present, and a clot present are shown in Fig. 3. In this figure, the GMI is calculated at four f points (5 Hz, 10 Hz, 20 Hz, and 100 Hz). In all cases, the LHS and RHS components of the GMI at a particular f point are equal but opposite, while across the band the trend in GMI matches the change in contrast between the lesion tissue and background brain. In the normal case, the GMI is theoretically zero at all f points, while for bleed and clot the trends mirror those of the σ spectra seen in Fig. 2. As the patterns for bleed and clot differ, differentiation of these two lesion types should be possible. The magnitude of the GMI for lesions is a function of lesion location and size (with the sign a function of location), but the pattern is unaffected by these properties. In reality these ideal results are confounded by many

factors, which are discussed in detail in [17].

E. FEM Models

The UCL dataset includes neuroimaging studies of the 18 stroke patients. These neuroimages were used to create four-layer computer aided design (CAD) models of the anatomy of the head of each patient above the inion-nasion line [25] using 3D Slicer to segment out the scalp, skull, CSF, and brain layers from CT images before using Autodesk Fusion 360 CAD software to refine and convert the layers to stereolithography (STL) files [26], [27]. Next, each of these four-layer STL models were converted into a fine tetrahedral mesh (~1 million elements each) with refinement around the electrode positions, with the electrodes modelled as 10 mm diameter and 1 kΩ contact impedance [28]. From each of these fine meshes, a coarse mesh (~ 200k elements each) was created with the electrodes in A-orientation for use in reconstruction of both simulated and human measurement frames.

Importantly these patient specific meshes are not exact representations of the true anatomy due to inevitable approximations in the segmentation and meshing process. Further, the electrode positioning would not have been in the theoretical exact location when recording patient voltages. These errors, and others such as assumed electrode contact impedance, are likely unavoidable in practice where patient specific anatomy will not be known at the time of stroke onset (and perhaps not required [29]), electrode placement will need to be performed with a degree of haste, and other compromises a result of the urgency of the condition. It is important that any technology, and thus any algorithm can cope with these types of errors with an adequate degree of tolerance. As discussed thoroughly in [17], MFSD-EIT is robust to a variety of errors within certain limits.

In order to generate the simulated data, numerical models of anatomy are required. In order to increase the size of the data

rode

Fig. 4. Top: All the simulated anatomies are derived from CT neuroimaging of stroke patients. A sample CT slice of the anatomy used to create the fine mesh is shown (from resources supplied in [3]). Bottom: Fine mesh with 50 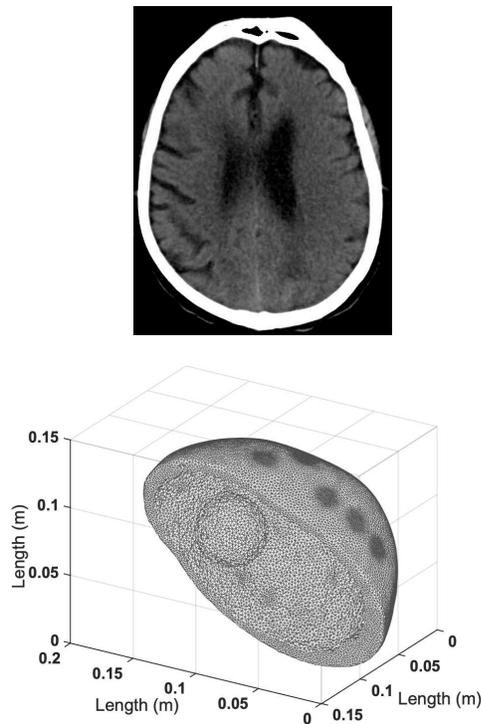ml simulated spherical lesion in the north-east position (where the front of the head represents north). The brain layer has been removed in order to show the lesion. The electrode refinement on the scalp layer is also clearly seen.

set as is recommended for ML [30], each of the 18 fine meshes were distorted to generate new anatomies. Each layer was dilated to 105% and shrunk to 97% of the original volume, which along with the normal geometry generates three possible anatomies for each of the four layers [17]. This range was selected based on the normal variation existing between people of similar height and weight, and the assumption that in practice a "best guess" generic mesh for reconstruction based on the patient's height and weight would be available [17], [31]. Combining every permutation of these layers results in 81 four-layer (i.e. 3^4 possible combinations) anatomies generated from each original patient STL model and hence 1,458 simulated anatomies (18 x 81) overall, each made into a fine mesh.

Into each of these meshes spherical perturbations were placed in one of four different locations (north-east, north-west, south-east, south-west), and as 20 ml or 50 ml volume. A wide spectrum of lesion sizes is possible in stroke. In Fiebach *et al.*, for example, the haemorrhagic volume in acute stoke was reported in as ranging from 1 – 101.5 ml [32], while Payabvash *et al.* reported an average ischaemic lesion as 39.5 ± 84.9 ml [33]. Hence, we selected 20 ml and 50 ml as representative small and large lesions, given the volumes reported in these previous studies.

In total, 1,458 normal simulated anatomies, and 11,664 lesion simulated anatomies were created each as fine FEM models. A representative FEM model is shown in Fig. 4 with a 50 ml lesion in the north-east location along with a sample slice

of the CT image that the model is derived from. In addition, a set of simulated anatomies was developed with lesions in a fifth location lying on the sagittal plane. These anatomies with central lesion were used in one study to analysis the effect of central lesions on classifier performance (Section III, H).

F. MF GMI Data Generation

This section describes the generation of MF GMI data from the simulated anatomies and from the human measurement frames.

Simulated measurement frames were generated from the simulated anatomies. The fine FEM models corresponding to the perturbation-free (i.e. normal) models were assigned σ values to the voxels of each of the four layers: scalp 0.23 Sm^{-1}, skull 0.05 Sm^{-1}, CSF 2 Sm^{-1}, brain 0.1 Sm^{-1}. These σ values are those of these tissues across the 5 Hz – 100 Hz band [17]–[19]. These models were then forward solved using the PEITS solver in order to generate measurement frames in both A- and B-orientations [34]. Noise was added to these frames at a level of 48 dB SNR (matching the average levels in the clinical human dataset [3]). The inverse problem of reconstruction was then performed onto the coarse mesh corresponding to the patient from which the simulated anatomy was derived (i.e. the original undistorted mesh). This emulates a real-world scenario where exact anatomy will be unknown and a best guess mesh may be needed for reconstruction. The GMI data was then computed. This process was repeated three additional times to give MF GMI data at each of the four selected frequency points.

Next, the perturbation models were assigned 0.7 Sm^{-1} to those voxels where the perturbation is located. This value is the σ of bleed across the band (Fig. 2). MF GMI data was generated in the same manner as the normal models to result in simulated data bleed cases. The perturbation models were next modelled as clots with the clot voxels assigned σ values of 0.02, 0.03, 0.05, 0.09 Sm^{-1} in separate forward solves representing the clot at the four frequency points (Fig. 2), generating MF GMI data for each case. Finally, the procedure for the normal models was repeated in order to provide sufficient sets of normal MF GMI data to balance the number of perturbation (bleed and clot) sets. Due to the addition of noise to the measurement frames, each MF GMI data from these normal cases was unique. The simulated data set hence comprised of measurement frames from the 1,458 normal simulated anatomies, and 11,664 lesion anatomies. This large size of the simulated data set with respect to the limited amount of human data set motivated the use of MF GMI data from simulated measurement frames in addition to human measurement frames. Further the use of simulated data allowed control of error sources that may have been present in the human data set such as electrode placement errors. With the exception of added noise, no additional sources of error were modelled in the simulated data set, facilitating an idealised control group that could be compared to the human data set.

An example of the reconstructed images generated at each *f* point and the corresponding MF GMI data is shown in Fig. 5. In this example, the reconstructed images and corresponding MF GMI data for a 50 ml clot in the south-west location is shown. The image is sharpest at 5 Hz, and progressively more
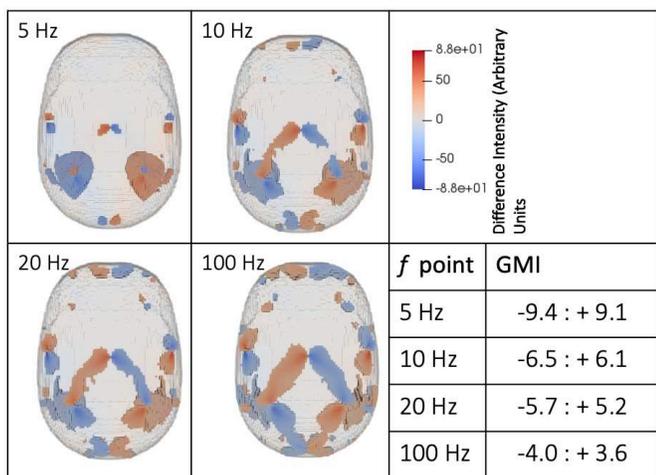
Fig. 5. Reconstructed images, and derived GMI values at the four f points (5 Hz, 10 Hz, 20 Hz, 100 Hz) for the case of a 50 ml clot in the south-west location. The contrast between clot and brain is highest at 5 Hz, reflected in a relatively clean image and corresponding high GMI values. As the contrast reduces across the band, the image becomes noisier, and the GMI values reduce.

noisy at the higher frequencies as the contrast between clot and brain decreases. This change in contrast between clot and brain (which is shown in Fig. 3) is reflected in the MF GMI values.

With regards the human data, the 18 coarse meshes used for reconstruction were derived from neuroimaging studies which were provided for the patients only [3]. Neuroimaging was not provided for the normal healthy volunteers. In order to reconstruct and generate MF GMI data from the healthy volunteer measurement frames, each of these measurement frames were reconstructed using each of the 18 coarse meshes. Hence, 180 normal MF GMI result sets were created. Further, in order to extend the amount of patient data, each of the 18 patient measurement frame sets were individually reconstructed and used to generate MF GMI results from each of the 18 coarse meshes. The result was 180 MF GMI result sets from bleed cases and 252 result sets from clot cases. Hence, in most cases, the mesh used to reconstruct and generate data was not anatomically related to the subject the measurement frame was derived from – the measurement frame from each patient is reconstructed onto each of the 18 possible coarse meshes, with only one of these related to the patient. Even in the case where the frame is related to the patient mesh, the mesh is an approximation as described in Section II-E. This mismatch is a further error source, and a further test as to the robustness of the overall approach proposed in this paper as in clinical scenarios it is unlikely patient neuroimaging or precise anatomy will be available [29].

These MF GMI data sets are then used as features for SVM classifiers, described in the next section.

III. ML CLASSIFICATION APPLIED TO MF GMI DATA

In this section a ML classification algorithm is applied to the MF GMI data derived from simulated and human measurement frames. The section starts with a brief summary of SVM classifiers, the type of classifier used in this study. Next the manner in which classification was carried out is outlined.

Finally, the results of the various data sets in binary classifications, and multiclass classification are reported. In addition, consideration is given to the effect of lesions lying on the sagittal plane, and different levels of simulated noise on classifier performance.

A. SVM Classifiers

SVMs are a group of ML algorithms often used for binary classification but can be adapted for use in multiclass classification [30]. SVM classification has been successfully demonstrated in previous biomedical applications including the use of microwaves for detection of breast cancer [35]–[37], impedance spectroscopy for the detection of prostate cancer [38], and work by our group into EIT measurement frames for the detection of brain haemorrhage [13], [14]. The SVM is trained with features from labelled observations (supervised learning) generating a trained model. This model is used to classify observations previously unseen with unknown labels (test or validation set). The features in this work are the MF GMI data across the four f points (hence the observations have 8 features or dimensions).

The basis of SVM classification is the creation of hyperplane with margins that optimally separates observations from the classes. When training the model, the hyperplane and margins are defined and applied to future observations to classify them as class 0 or class 1 (in binary case). The kernel used by a SVM defines the function used in hyperplane generation. In this study a radial basis function (RBF) is used which offers a flexible and robust hyperplane compared to simpler kernels such as linear kernels [13], [14].

The classification problems considered in this paper are the binary classification of normal versus lesion (bleed or clot), as well as all combinations of normal, bleed, and clot. Further, a multiclass classification of normal, bleed, and clot is performed.

The labels (-1 or +1) correspond to the classes under investigation in a given binary classification task (e.g. 'normal' versus 'lesion'; 'bleed versus 'clot'; 'normal' versus 'bleed'; 'normal versus 'clot'). In the multiclass case, the labels 0, 1, 2 are used for 'normal', 'bleed', and 'clot' respectively.

The performance of the classifier can be reported in terms of the confusion matrix which tabulates the numbers of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) classifications made by the trained model on a test or validation set. These values can be summarised as metrics including sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) of the classifier [30]. These metrics are reported in the results sections of this paper, with the confusion matrices reported for the multiclass classification.

B. Classification Protocol

In this section, the processing of the MF GMI data before use in the classifier is described as well as the procedure for training and testing a SVM classifier. The implementation used MATLAB, and in particular resources provided in the statistics and machine learning toolbox [39]. For each of the scenarios discussed in the following sections, the MF GMI data sets from
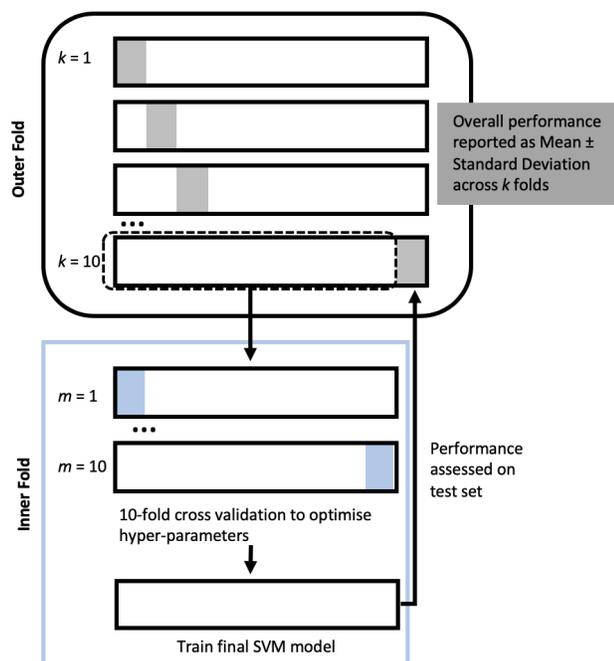
Fig. 6. Nested cross-validation. The complete data set is separated into k =10 folds with 90% as a training set (white) and 10% as a test set (grey). The training set from each of these outer folds is itself divided similarly into m = 10 inner folds divided into training (white) and test (blue) sets, with 10-fold cross validation performed on the inner fold to optimise hyper-parameters. These hyper-parameters are then used to train a final SVM model on the entire training set, with performance assessed on the held out test set. This repeated over the k folds with a final overall performance reported as the mean ± standard deviation.

the simulated data at 48 dB, and the human data (~48 dB) are used in separate classifications with separately trained and tested classifiers, with the results for each set reported. For a given classification, the MF GMI data corresponding to the scenario under study is used as the input features to the SVM classifier. Each observation has 8 features as described in Section III, A, and is labelled by class. In the case of the simulated data sets, the number of cases used are balanced between classes, whereas in the case of the human data set all the data is used due to the relatively small size of the available data.

Nested cross-validation as described in [40] and used in [14], [37], provides a generalised robust indication of classifier performance and helps mitigate against bias. This technique is used in this study to optimise SVM classification performance. The data is separated into k = 10 separate folds. Each of these outer folds is divided into a unique training set and test set made up of 90% and 10% of the total data set, respectively. The ttraining set is used to select optimal SVM classifier hyper-parameters (the box constraint and kernel scaling factor) using a Bayesian optimisation procedure in a m = 10 fold cross-validation process. The performance of the final classifier is assessed using the excluded test set. This process is repeated for each of the k = 10 folds with the final overall performance reported as the mean ± standard deviation across the ten iterations. This approach is summarised in Fig. 6. In each case, z-normalisation of the training set features is performed. The
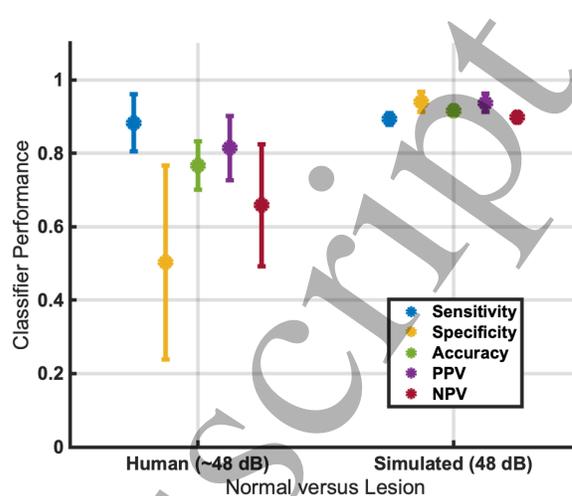

Fig. 7. Classifier performance for normal (negative class) versus lesion (positive class) for human (~48 dB SNR), and simulated (48 dB SNR) data sets. A mean accuracy of 77% is seen for the human data, with a mean sensitivity of 88% and PPV of 81% indicating strong detection of lesion while the mean specificity (51%) and NPV (66%) indicative of a low rate of accurate detection of normal cases. In the case of the simulated results strong performance is indicated by an accuracy of 92% at the 48 dB SNR level.

test set is in terms of z-normalised using the mean and standard deviation from the training set, this process ensures that there is no data contamination between the training and test set from the z-normalisation process [41]–[43].

In the following sections, four cases of binary classification are considered followed by a multiclass classification problem (normal versus bleed versus clot). The multiclass problem uses an error-correcting output codes (ECOC) classifier to effectively deal with the multiclass problem by combining multiple SVM binary classifiers [44]. Aside from using a ECOC approach, the multiclass classification is treated identically to the binary classifications with nested cross-validation used.

C. Normal versus Lesion

This study examined the binary classification of normal (assigned as the negative class) and lesion (assigned as the positive class). Lesion referred to either bleed or clot. The simulation data sets comprised of 10,064 simulated normal cases and 10,064 simulated lesion cases (with an even number of bleed and clot cases). The human data set comprises of 180 normal cases and 432 lesion cases (180 bleed, 252 clot cases). The classifier performance for each data set is presented in Fig. 7. The results of the human data indicate strong lesion detection with a mean sensitivity of 88% and PPV of 81%. Detection of normal cases however proved challenging leading to poor mean specificity of 51% (NPV 66%). The mean accuracy was 77%. In the case of the simulated 48 dB SNR data set all metrics are ≥ 89%, with a mean accuracy of 92%.

D. Bleed versus Clot

In this study, binary classification of bleed (negative class) versus clot (positive class is considered). This scenario may be the most relevant study to the clinical case where the patient is known to have a stroke and so the crucial diagnostic step is the
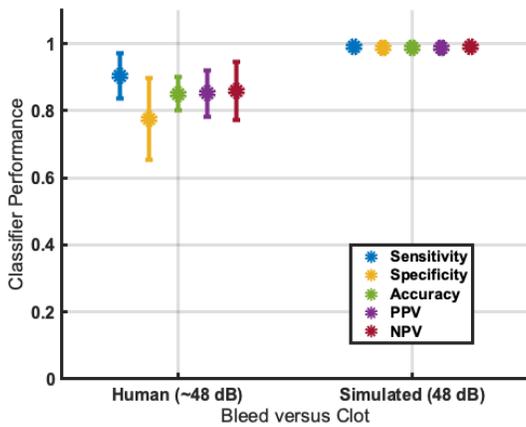
Fig. 8. Classifier performance for bleed (negative class) versus clot (positive class) for human (~48 dB SNR) and simulated (48 dB SNR) data sets. A mean accuracy of 85% is seen for the human data, with clots strongly detected (mean sensitivity 90% and PPV 85%). Bleeds are slightly less well detected with a mean specificity (77%) and NPV (86%). The simulated results are strong at the 48 dB SNR level, with no metric with a mean less than 99%.

identification of the lesion as a bleed or clot. The simulated data sets have an even number of bleed and clot cases with a total of 10,064 cases. The human data set is made up of the 180 bleed and 252 clots. The results are shown in Fig. 8. A mean accuracy of 85% is reported for the human data, with a mean sensitivity of 90% and PPV of 85% indicative of strong correct clot detection. Bleeds are slightly less well detected with a mean specificity of 77% and NPV of 86%. With regards to the 48 dB SNR simulated data set, all metrics have a mean of $\geq 99\%$.

E. Normal versus Bleed

The results for normal (negative class) versus bleed (positive class) are shown in Fig. 9. The simulated data sets have an even number of normal and bleed cases with a total of 10,064 cases. The human data set is made up of the 180 normal and 180 bleed cases. A mean specificity of 90% and NPV of 85%, with a mean sensitivity of 85% and PPV of 90% is reported for the human
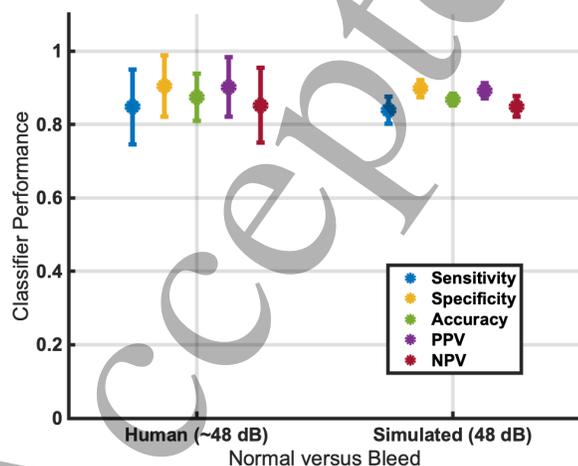

Fig. 9. Classifier performance for normal (negative class) versus bleed (positive class) or human (~48 dB SNR), and simulated (48 dB SNR) data sets. The classifier preforms with a mean accuracy of 87% in both the human and the simulated 48 dB SNR sets.
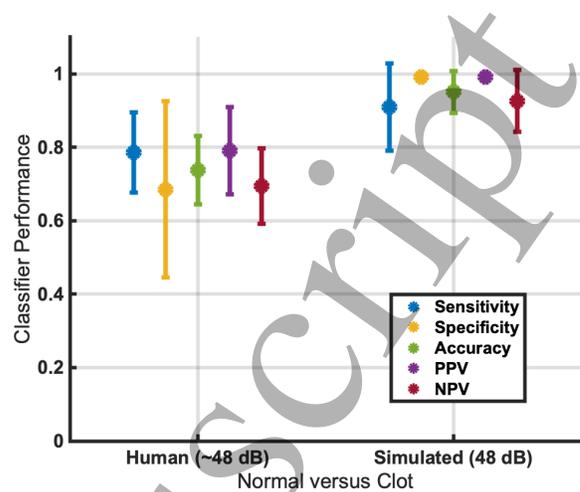
Fig. 10. Classifier performance for normal (negative class) versus clot (positive class) for human (~48 dB SNR), and simulated (48 dB SNR) data sets. The classifier preforms with a mean accuracy of 74% with the human data, with a mean accuracy of 95% in the simulated 48 dB SNR set.

data (overall mean accuracy of 87%). In the simulated data, the mean accuracy of the 48 dB data set is also 87%. As well as being of value in the stroke diagnostic pathway, intracranial bleeds are an important feature in other conditions such as traumatic brain injury (TBI) [13], [14]. Thus, a diagnostic modality that can robustly differentiate normal from bleed would be of significant value when applied to such cases.

F. Normal versus Clot

The results of the final binary classification combination of normal (negative class) versus bleed (positive class) are shown in Fig. 10. The human data set is made up of the 180 normal and 252 bleed cases, with the simulated data sets have 10,064 cases with equal numbers of both case types. In the simulated results, all metrics for the 48 dB SNR set are $\geq 91\%$. The human data set results have a mean accuracy of 74%, with the strongest result being a PPV of 79%.

G. Multiclass Classification

In this study, multiclass classification is performed with 6,290 each of normal, bleed, and clot cases (18,870 cases in total) in the simulated data sets (at 48 dB SNR level). The human data set is made of the 180 normal, 180 bleed, and 252 clot cases. The results are presented as confusion matrices of the average $\pm$ standard deviation as percentages of the classification results of the nested cross-validation in Tables I and II. The mean overall accuracy of the classifier with the human data set is 68.0%, and for the simulated 48 dB set is 90.1%.

TABLES I - II
RESULTS AS CONFUSION MATRICES OF MULTICLASS CLASSIFICATION FOR HUMAN DATA SET, AND SIMULATED DATA SET (48 DB SNR). THE RESULTS ARE THE AVERAGE ± STANDARD DEVIATION OVER 10 ITERATIONS AS PERCENTAGES. THE GREEN CELLS ARE CORRECT CLASSIFICATIONS, WITH RED CELLS INCORRECT. THE LIGHT GREY CELLS ON THE RIGHT ARE THE PPV RESULTS, AND THE LIGHT GREY CELLS ON THE BOTTOM ARE THE SENSITIVITY RESULTS. THE DARKER GREY CELL IN THE BOTTOM RIGHT IS THE OVERALL ACCURACY.

TABLE I
HUMAN DATA SET

Output Class				
Normal	**17.9 ± 7.6**	3.6 ± 2.9	7.9 ± 4.2	61.0 ± 26.1
Bleed	2.2 ± 2.2	**19.8 ± 6.3**	3.0 ± 2.7	79.3 ± 25.3
Clot	9.3 ± 7.4	6.1 ± 5.5	**30.3 ± 5.6**	66.3 ± 12.9
	60.9 ± 26.0	67.2 ± 21.5	68.4 ± 13.3	**68.0 ± 19.9**
	Normal	Bleed	Clot	
		Target Class		

TABLE II
SIMULATED DATA SET (48 DB SNR)

Output Class				
Normal	**29.7 ± 1.6**	4.4 ± 1.0	1.5 ± 1.1	83.4 ± 4.5
Bleed	3.3 ± 0.9	**28.6 ± 1.5**	0.1 ± 0.2	89.3 ± 4.8
Clot	0.2 ± 0.2	0.3 ± 0.4	**31.7 ± 1.6**	98.4 ± 4.9
	89.3 ± 4.9	86.0 ± 4.6	94.7 ± 4.7	**90.1 ± 4.7**
	Normal	Bleed	Clot	
		Target Class		

H. Effect of Centrally Lying Lesions

It is of interest to consider the effect of lesions lying on the sagittal plane on classifier performance. As described in [15], [16], an absolute limitation of the symmetry based algorithm is the inability to detect lesions lying perfectly on the plane of symmetry as there is effectively no difference in the scene presented on either side of the plane. Simulated data sets at 48 dB are used to study the effect of centrally lying lesions using normal versus lesion, and bleed versus clot binary classification problems. In both cases, an even number of each class are used with the number of cases the same as described in Sections III, C and III, D. For the data sets with no centrally lying lesions the four locations and two volume sizes as described in Section II,

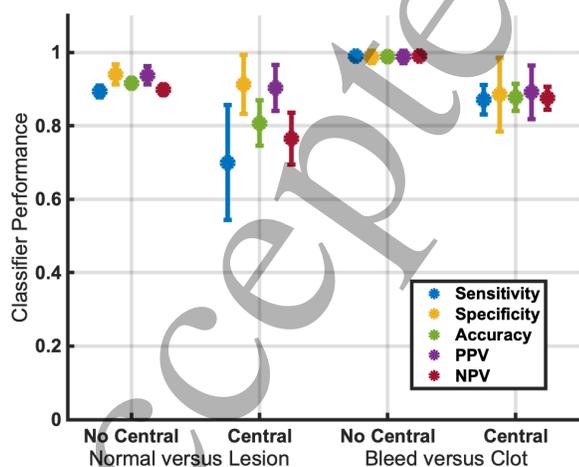Fig. 11. Central lying lesions effect on classifier performance for normal (negative class) versus lesion (positive class), and bleed (negative class) versus clot (positive class) for simulated (48 dB SNR) data sets. The classifier performance is consistently reduced in all metrics when centrally placed lesions are included in the data set compared to when absent.

E are used. For the data sets with centrally lying lesions, 20% of the cases have centrally lying lesions of the two volume sizes. The results are shown in Fig. 11. In both cases considered, the results indicate a drop in classifier performance when central lesions are present in the data set. For example, the mean accuracy drops from 92% to 81% in the normal versus lesion case, with a drop from 99% to 88% in the bleed versus clot case.

I. Simulated Data Sets at Different Noise Levels

In this work, the simulated data had noise added at a level of 48 dB SNR in order to match the SNR of the clinical data [3]. However, the simulations were also performed at noise levels of 80 dB and 20 dB in order to study the effect of noise. The mean accuracy from the simulated sets at each SNR level for each binary classification problem is shown in Table III. Performance drops off at 20 dB compared to at 48 dB as expected, while the 80 dB level out performs the 48 dB. However, these results indicate only minor differences between the 48 dB and 80 dB sets, suggesting that perhaps 48 dB is adequate for accurate results if no other error sources are present.

IV. DISCUSSION

This section presents discussions of the classification results from the various scenarios considered in the previous section.

A. Normal versus Lesion

In the stroke diagnostic pathway, the initial diagnosis of stroke is not usually done using neuroimaging, relying instead on preliminary diagnostics that can be performed by the first responders including physical examination protocols such as Rule Out Stroke in the Emergency Room (ROSIER) [45]. The results of these are used to label a patient as a potential stroke patient with neuroimaging subsequently performed [45]. Although initially thought to have a relatively high sensitivity and specificity (92% and 86% respectively), there is now evidence the specificity of these tests is possibly as low as 15% when performed by first responders [46]. As such there is value in a technology that could be replace such preliminary tests or be used in conjunction. The results seen in Fig. 7 show that with human data a mean accuracy of 77% was achieved for the classification of normal versus lesion (bleed or clot). Although the specificity (mean 51%) and NPV (66%) were lower than the sensitivity (mean 88%) and PPV (mean 81%), these former metrics show a significantly higher performance at identifying normal (negative) patients than the current preliminary diagnostics used. This encouraging result is supported by the results from the simulated data sets, where no metric is < 89%

TABLE III
PERFORMANCE ACCURACY AS AVERAGE ± STANDARD DEVIATION OVER 10 ITERATIONS AS PERCENTAGES FOR EACH BINARY CLASSIFICATION PROBLEM FOR SIMULATED DATA AT 80 DB, 48 DB, AND 20 DB SNR NOISE LEVELS.

SNR Level	Normal vs. Lesion	Bleed vs. Clot	Normal vs. Bleed	Normal vs. Clot
20 dB	67 ± 3	74 ± 5	76 ± 2	60 ± 3
48 dB	92 ± 1	99 ± 1	87 ± 2	95 ± 5
80 dB	93 ± 0	99 ± 4	90 ± 3	99 ± 2
		Binary Classification Problem		

when 48 dB SNR noise is added.

B. Bleed versus Clot

Timely neuroimaging using CT or MRI is needed to differentiate between the lesion types before treatment for stroke can start [1]. It is this differentiation of lesion type that is the primary motivation of this work, with the fundamental idea of MF GMI data being capable of differentiating between bleed and clot as a result of differing σ profiles discussed in Section II, C and shown in Fig. 3. The results as shown in Fig. 8 are encouraging with the classifier performance on the human data set showing a mean accuracy of 85%, with a similar mean NPV and PPV values. While these results demonstrate significant promise on human data, there is evidence that further improvements are possible when the results of the simulated sets showing all metrics with mean values $\geq$ 99% are considered.

C. Normal versus Bleed

A related pathology to stroke is TBI, which may or may not feature intracranial haemorrhage. Initially physical examination protocols are used to triage a patient and decide on case severity with haemorrhage correlated with worse outcomes and a need to neuroimage. The ability to rule in or out haemorrhage early in the triage process would result in better use of CT resources and patient outcomes and is another potential application for the presented algorithm.

Although the σ of bleed does not change appreciably across the EIT band [18], [19], the presence of a bleed results in a significant change in the recorded EIT measurement frame and MF GMI data compared to the normal case due to a disturbance in symmetry [13]–[16]. This change in MF GMI should be distinguishable from the normal, and is in the case where a high level of symmetry exists in the absence of a lesion [17]. However, the presence of asymmetric anatomy in the normal case can confound this detection [17].

The results from the simulated data shown in Fig. 9, when compared to the corresponding results in Fig. 10 (normal versus clot), imply challenge in classifying normal versus bleed compared to normal versus clot. The classifier results in Fig. 9 have a lower mean value compared to those in Fig. 10 for the simulated data. The classifier in the case of normal versus clot has the advantage of the change in σ featured by clot across the band. However, the results are still strong for normal versus bleed with mean accuracy of 87% for both the human and the simulated data sets. These results imply the MFSD-EIT with classification approach may indeed have application in areas such as TBI and build upon the work of [13], [14] where ML was applied to EIT measurement frames in detecting intracranial haemorrhage.

D. Normal versus Clot

Similar to the classification problem of bleed versus clot, normal versus clot is theoretically well suited to the MF GMI approach due to the differing σ profile of clot across the band compared to no change for the normal case (Fig. 3). The results as shown in Fig. 10 confirm this with an accuracy of 74% reported for the human data, and all metrics $\geq$ 91% for the

simulated data at 48 dB. When considering the classifier performance results of the simulated data sets scenarios where there is a divergence in σ profiles across the band between the two classes considered result in better classifier performance. For example, in the simulated data sets the mean accuracy is 95% for normal versus clot and 99% for bleed versus clot, compared to 92% for normal versus lesion and 87% for normal versus bleed.

E. Multiclass Classification

A complete stroke diagnostic workup would comprise of classification as normal, haemorrhagic stroke, or ischaemic stroke. Currently, an initial diagnosis is performed of normal versus lesion (not stroke or probable stroke), before neuroimaging is performed in the patients with a lesion to decide on bleed or clot (haemorrhagic or ischaemic stroke) [45]. The multiclass results in Tables I – II give an indication as to the effectiveness of MFSD-EIT with SVM classification applied to such a one-step diagnosis. An overall accuracy of 68.0% is achieved with the human data, and 90.1% with the simulated data at 48 dB. Analysis of the confusion matrices show that in the simulated data set, the most common misclassification is normal as bleed and vice-versa. This misclassification is a consequence of the lack of contrast change in these tissues across the band. The change in σ of clot results in a consistently high performance (sensitivity > 94%, PPV > 98%).

F. Effect of Centrally Lying Lesions

Centrally lying lesions do not disturb the symmetry of the scene and so inclusion of such lesions in a data set should theoretically show worse performance compared to a set without. This prediction is confirmed by the results shown in Fig. 11, where a consistent drop in performance is seen when the data sets that include the centrally lying lesions are included. While a causative stroke lesion can occur anywhere within the network of blood vessels in the brain, it is the case that vasculature is largely duplicated on either side of the sagittal plane with the major arteries serving the brain (middle, anterior and posterior cerebral arteries) all having left and right hand sides [47]. As such it would be unlikely for a stroke patient to have a sagittally positioned lesion (for example no lesion in the human data set was on the midline [3]) but such lesions would be challenging to detect with this technique.

G. UCL MF EIT Stroke Dataset - Other Remarks

The UCL MF EIT dataset is currently the most comprehensive collection of human stroke data available to researchers [3]. In this paper, simulated data generated from the resources provided in the dataset was used to validate the proposed approach in a controlled environment with limited sources of error. While improved real world results may have been possible with phantom or animal models, the availability of gold standard human EIT data motivated the adoption of the measurement frames provided in [3]. It is acknowledged however that the dataset does however have some limitations.

The size of the dataset is relatively small, especially for use in machine learning applications where sets as large as possible are desirable and often numbering into the thousands [13], [37].

The neuroimaging studies from the 18 patients were used to form 18 unique coarse FEM models for reconstruction with the $N_{Healthy} = 10$, $N_{Haemorrhagic} = 10$ and $N_{Ischaemic} = 14$ human EIT measurement frames reconstructed onto each model to give 180 MF GMI result sets from normal cases, 180 sets from bleed cases, and 252 sets from ischaemic cases,. While this approach increased the data by a factor of 18, it remains a relatively modest number of observations.

Further, the data is biased in terms of classes with more cases of ischaemia than normal or bleed. Such imbalances in data can skew the performance results. In this study, the simulated data sets were kept balanced in terms of numbers of observations from each class but all the human MF GMI data was used due to the limited pool available. Metrics such as accuracy, PPV and NPV are sensitive to imbalanced data, while sensitivity and specificity are not [48], [49]. Further, the dataset provides information on the location and nature of the lesions, but not precise information on size beyond qualitative terms such as 'small', and 'large'. There are however examples of small and large lesions of both types in the set.

The work of [17] presented an analysis of the effect of error sources on the quality of GMI data. These errors included measurement frame noise, electrode contact impedance errors, and electrode positioning errors. While care was taken to maximise the quality of the protocol and data recorded, inevitably these sources of error are present in the human data and have an effect on the collected measurement frames. Of particular interest is electrode positioning errors where differences greater than $\pm$ 5mm in the assumed location of electrodes can severely affect the GMI data [17]. The actual locations of electrodes are not reported in the dataset, and when placing electrodes on patients, placement error in the order of $\pm$ 5 mm is likely.

Other sources of uncertainty also exist in the dataset. For example, the delay from stroke onset to EIT recording ranged from hours to days while neuroimaging was performed at different timepoints which could lead to errors in data interpretation. Further, measurement noise is present in all recordings, with an average SNR of 48 dB across the frequency band despite the use of the ScouseTom which has a theoretical rating of 77.5 dB [3], [22]. The higher SNR was achieved when in use with a resistor phantom, highlighting the challenge in moving from ideal phantom models to human [22].

These limitations must be kept in consideration when considering the results of the study particularly with respect to the human data. Further, these limitations may explain the discrepancy in performance between the results from the human and simulated data sets. In the simulated cases while noise at a level of 48 dB SNR was added to the frames, no other modelling errors were present (for example electrode positioning).

H. Simulated Data Sets at Different Noise Levels

As shown by the results in Table III although the accuracy at 80 dB is better than at 48 dB as expected, the performance is similar. These results suggest that while the higher SNR is desirable, perhaps stroke detection may be possible at lower SNR levels if other errors sources such as electrode placement are controlled.

V. CONCLUSIONS

This work applied the novel MFSD-EIT algorithm to both human and simulated stroke data for the first time. The GMI metric summarises symmetrical differences on either side of the brain, with analysis across multiple frequencies allowing detection and identification of normal, bleed and clot cases. This principle is used for the first time with ML techniques using a robust nested cross-validation approach to generate trained RBF kernel SVM classifiers with the MF GMI data as input features. In binary classification problems, better classification is generally seen where tissues differ in the pattern of σ change across the frequency band. Importantly this divergence in σ pattern exists for bleed versus clot, with differentiation of these lesion types in stroke the main application proposed in this work. Strong results are also reported for more challenging clinically important applications such as normal versus bleed. A mean accuracy of $\geq 85\%$ is reported for the human data in these cases. A complete one-step diagnostic modality for stroke is also considered as multiclass classification of normal versus bleed versus clot, with a mean accuracy of 68% with human data.

Limitations exist with the human dataset, with small set size and possible errors in important parameters such as electrode positioning affecting achievement of accurate MF GMI data. Further, the average data noise had a low SNR of 48 dB. Notably, the results achieved in the human dataset were achieved despite the presence of these limitations and the existence of many error sources. In [17], some methods to mitigate these errors were presented, including with respect to electrode placement. Hence, further work to mitigate these errors should lead to even better algorithm performance with human data. Indeed, there are indications from the simulated data that improvements to reduce these errors in human data would indeed result in better performance. In simulated datasets, mean accuracy is always $\geq 87\%$ for the measurements with 48 dB SNR, and $\geq 90\%$ for measurements with an 80 dB SNR noise level. While the simulated data has limited sources of errors relative to the human data, an important error source as described in [17] is asymmetric normal anatomy, which is present in both datasets.

We believe the results demonstrate that using MFSD-EIT with ML algorithms is a promising approach to diagnosing stroke and related applications.

APPENDIX

The UCL stroke dataset is documented in [3], and available at https://github.com/EIT-team/Stroke_EIT_Dataset (archived at DOI: 10.5281/zenodo.1199523. The meshing software is available at https://github.com/EIT-team/Mesher. The forward solver used, PEITS (Parallel EIT Solver) is described in [34], and available at https://github.com/EIT-team/PEITS (archived at DOI: 10.5281/zenodo.1641128). The reconstruction software is available at https://github.com/EIT-team/Reconstruction (archived at DOI: 10.5281/zenodo.1643416).

REFERENCES

[1] G. A. Donnan, M. Fisher, M. Macleod, and S. M. Davis, 'Stroke', *The Lancet*, vol. 371, no. 9624, pp. 1612–1623, May 2008, doi: 10.1016/S0140-6736(08)60694-7.

[2]   electronic Medicines Compendium, 'SPC Actilyse'. https://www.medicines.org.uk/emc/medicine/308 (accessed Dec. 27, 2018).

[3]   N. Goren *et al.*, 'Multi-frequency electrical impedance tomography and neuroimaging data in stroke patients', *Sci. Data*, vol. 5, p. 180112, Jul. 2018, doi: 10.1038/sdata.2018.112.

[4]   B. Brown, 'Electrical impedance tomography (EIT): a review', *J. Med. Eng. Technol.*, vol. 27, no. 3, pp. 97–108, Jan. 2003, doi: 10.1080/0309190021000059687.

[5]   D. Holder, *Electrical Impedance Tomography: Methods, History, and Applications*. Bristol: Institute of Physics Publishing, 2005.

[6]   A. Adler and A. Boyle, 'Electrical Impedance Tomography: Tissue Properties to Image Measures', *IEEE Trans. Biomed. Eng.*, vol. 64, no. 11, pp. 2494–2504, Nov. 2017, doi: 10.1109/TBME.2017.2728323.

[7]   A. Boyle and A. Adler, 'The impact of electrode area, contact impedance and boundary shape on EIT images', *Physiol. Meas.*, vol. 32, no. 7, pp. 745–754, Jul. 2011, doi: 10.1088/0967-3334/32/7/S02.

[8]   M. Jehl, J. Avery, E. Malone, D. Holder, and T. Betcke, 'Correcting electrode modelling errors in EIT on realistic 3D head models', *Physiol. Meas.*, vol. 36, no. 12, pp. 2423–2442, Dec. 2015, doi: 10.1088/0967-3334/36/12/2423.

[9]   E. Malone, G. S. dos Santos, D. Holder, and S. Arridge, 'A Reconstruction-Classification Method for Multifrequency Electrical Impedance Tomography', *IEEE Trans. Med. Imaging*, vol. 34, no. 7, pp. 1486–1497, Jul. 2015, doi: 10.1109/TMI.2015.2402661.

[10]   E. Malone, G. Sato dos Santos, D. Holder, and S. Arridge, 'Multifrequency Electrical Impedance Tomography Using Spectral Constraints', *IEEE Trans. Med. Imaging*, vol. 33, no. 2, pp. 340–350, Feb. 2014, doi: 10.1109/TMI.2013.2284966.

[11]   E. Malone, M. Jehl, S. Arridge, T. Betcke, and D. Holder, 'Stroke type differentiation using spectrally constrained multifrequency EIT: evaluation of feasibility in a realistic head model', *Physiol. Meas.*, vol. 35, no. 6, pp. 1051–1066, Jun. 2014, doi: 10.1088/0967-3334/35/6/1051.

[12]   International Organization for Standardization, 'International Electrotechnical Commission. Medical Electrical Equipment: Part 1: General Requirements for Basic Safety and Essential Performance.' IEC 60601-1, 2002.

[13]   B. McDermott, M. O'Halloran, E. Porter, and A. Santorelli, 'Brain haemorrhage detection using a SVM classifier with electrical impedance tomography measurement frames', *PLOS ONE*, vol. 13, no. 7, p. e0200469, Jul. 2018, doi: 10.1371/journal.pone.0200469.

[14]   B. McDermott, E. Dunne, M. O'Halloran, E. Porter, and A. Santorelli, 'Brain Haemorrhage Detection Through SVM Classification of Electrical Impedance Tomography Measurements', in *Brain and Human Body Modeling: Computational Human Modeling at EMBC 2018*, S. Makarov, M. Horner, and G. Noetscher, Eds. Cham: Springer International Publishing, 2019, pp. 211–244.

[15]   B. McDermott, E. Porter, M. Jones, B. McGinley, and M. O'Halloran, 'Symmetry difference electrical impedance tomography—a novel modality for anomaly detection', *Physiol. Meas.*, vol. 39, no. 4, p. 044007, Apr. 2018, doi: 10.1088/1361-6579/aab656.

[16]   B. J. McDermott, J. Avery, M. O'Halloran, K. Y. Aristovich, and E. Porter, 'Bi-frequency symmetry difference electrical impedance tomography – a novel technique for perturbation detection in static scenes', *Physiol. Meas.*, Feb. 2019, doi: 10.1088/1361-6579/ab08ba.

[17]   B. J. McDermott, M. Ohalloran, J. Avery, and E. Porter, 'Bi-Frequency Symmetry Difference EIT - Feasibility and Limitations of Application to Stroke Diagnosis', *IEEE J. Biomed. Health Inform.*, pp. 1–14, 2019, doi: 10.1109/JBHI.2019.2960862.

[18]   B. Packham *et al.*, 'Comparison of frequency difference reconstruction algorithms for the detection of acute stroke using EIT in a realistic head-shaped tank', *Physiol. Meas.*, vol. 33, no. 5, pp. 767–786, May 2012, doi: 10.1088/0967-3334/33/5/767.

[19]   L. Horesh, 'Some Novel Approaches in Modelling and Image Reconstruction for Multi-Frequency Electrical Impedance Tomography of the Human Brain', UCL, London, UK, 2006.

[20]   A. Adler, B. Grychtol, and R. Bayford, 'Why is EIT so hard, and what are we doing about it?', *Physiol. Meas.*, vol. 36, no. 6, pp. 1067–1073, Jun. 2015, doi: 10.1088/0967-3334/36/6/1067.

[21]   L. Horesh, O. Gilad, A. Romsauerova, S. Arridge, and D. Holder, 'Stroke type differentiation by multi-frequency electrical impedance tomography - a feasibility study', in *Proc IFMBE*, 2005, vol. 11, pp. 1252–1256.

[22]   J. Avery, T. Dowrick, M. Faulkner, N. Goren, and D. Holder, 'A Versatile and Reproducible Multi-Frequency Electrical Impedance Tomography System', *Sensors*, vol. 17, no. 2, pp. 280–280, Jan. 2017, doi: 10.3390/s17020280.

[23]   J. L. Mueller and S. Siltanen, *Linear and Nonlinear Inverse Problems with Practical Applications*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2012.

[24]   A. Romsauerova, A. McEwan, L. Horesh, R. Yerworth, R. H. Bayford, and D. S. Holder, 'Multi-frequency electrical impedance tomography (EIT) of the adult human head: initial findings in brain tumours, arteriovenous malformations and chronic stroke, development of an analysis method and calibration', *Physiol. Meas.*, vol. 27, no. 5, pp. S147–S161, May 2006, doi: 10.1088/0967-3334/27/5/S13.

[25]   J. Avery, K. Aristovich, B. Low, and D. Holder, 'Reproducible 3D printed head tanks for electrical impedance tomography with realistic shape and conductivity distribution', *Physiol. Meas.*, vol. 38, no. 6, pp. 1116–1131, Jun. 2017, doi: 10.1088/1361-6579/aa6586.

[26]   A. Fedorov *et al.*, '3D Slicer as an image computing platform for the Quantitative Imaging Network', *Magn. Reson. Imaging*, vol. 30, no. 9, pp. 1323–1341, Nov. 2012, doi: 10.1016/j.mri.2012.05.001.

[27]   Autodesk, *Fusion 360*. Mill Valley, CA, USA: Autodesk.

[28]   B. Grychtol and A. Adler, 'FEM electrode refinement for electrical impedance tomography', in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jul. 2013, vol. 2013, pp. 6429–6432, doi: 10.1109/EMBC.2013.6611026.

[29]   M. Jehl, K. Aristovich, M. Faulkner, and D. Holder, 'Are patient specific meshes required for EIT head imaging?', *Physiol. Meas.*, vol. 37, no. 6, pp. 879–892, 06 2016, doi: 10.1088/0967-3334/37/6/879.

[30]   T. M. Mitchell, *Machine Learning*. New York: McGraw-Hill, 1997.

[31]   A. K. D. Nguyen, A. A. Simard-Meilleur, C. Berthiaume, R. Godbout, and L. Mottron, 'Head Circumference in Canadian Male Adults: Development of a Normalized Chart', *Int. J. Morphol.*, vol. 30, no. 4, pp. 1474–1480, Dec. 2012, doi: 10.4067/S0717-95022012000400033.

[32]   J. B. Fiebach *et al.*, 'Stroke Magnetic Resonance Imaging Is Accurate in Hyperacute Intracerebral Hemorrhage: A Multicenter Study on the Validity of Stroke Imaging', *Stroke*, vol. 35, no. 2, pp. 502–506, 2004, doi: 10.1161/01.STR.0000114203.75678.88.

[33]   S. Payabvash, S. Taleb, J. C. Benson, and A. M. McKinney, 'Acute Ischemic Stroke Infarct Topology: Association with Lesion Volume and Severity of Symptoms at Admission and Discharge', *Am. J. Neuroradiol.*, vol. 38, no. 1, pp. 58–63, Jan. 2017, doi: 10.3174/ajnr.A4970.

[34]   M. Jehl, A. Dedner, T. Betcke, K. Aristovich, R. Klofkorn, and D. Holder, 'A Fast Parallel Solver for the Forward Problem in Electrical Impedance Tomography', *IEEE Trans. Biomed. Eng.*, vol. 62, no. 1, pp. 126–137, Jan. 2015, doi: 10.1109/TBME.2014.2342280.

[35]   A. Santorelli, E. Porter, E. Kirshin, Y. J. Liu, and M. Popovic, 'Investigation of Classifiers for Tumour Detection with an Experimental Time-Domain Breast Screening System', *Prog. Electromagn. Res.*, vol. 144, pp. 45–57, 2014, doi: 10.2528/PIER13110709.

[36]   R. C. Conceicao, M. O'Halloran, M. Glavin, and E. Jones, 'Support Vector Machines for the Classification of Early-Stage Breast Cancer Based on Radar Target Signatures', *Prog. Electromagn. Res. B*, vol. 23, pp. 311–327, 2010, doi: 10.2528/PIERB10062407.

[37]   B. Oliveira, D. Godinho, M. O'Halloran, M. Glavin, E. Jones, and R. Conceição, 'Diagnosing Breast Cancer with Microwave Technology: remaining challenges and potential solutions with machine learning', *Diagnostics*, vol. 8, no. 2, p. 36, May 2018, doi: 10.3390/diagnostics8020036.

[38]   M. A. Shini, S. Laufer, and B. Rubinsky, 'SVM for prostate cancer using electrical impedance measurements', *Physiol. Meas.*, vol. 32, no. 9, pp. 1373–1387, Sep. 2011, doi: 10.1088/0967-3334/32/9/002.

[39]   *MATLAB 2019A*. Natick, Massachusetts, USA: The MathWorks Inc.

[40]   G. C. Cawley and N. L. C. Talbot, 'On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation', p. 29.

[41]   Guodong Guo and S. Z. Li, 'Content-based audio classification and retrieval by support vector machines', *IEEE Trans. Neural Netw.*, vol. 14, no. 1, pp. 209–215, Jan. 2003, doi: 10.1109/TNN.2002.806626.

[42]   T. Jayalakshmi and A. Santhakumaran, 'Statistical Normalization and Back Propagationfor Classification', *Int. J. Comput. Theory Eng.*, pp. 89–93, 2011, doi: 10.7763/IJCTE.2011.V3.288.

[43]   B. Kumar Singh, K. Verma, and A. S. Thoke, 'Investigations on Impact of Feature Normalization Techniques on Classifier's Performance in Breast Tumor Classification', *Int. J. Comput. Appl.*, vol. 116, no. 19, pp. 11–15, Apr. 2015, doi: 10.5120/20443-2793.

[44] S. Escalera, O. Pujol, and P. Radeva, 'Separability of ternary codes for sparse designs of error-correcting output codes', *Pattern Recognit. Lett.*, vol. 30, no. 3, pp. 285–297, Feb. 2009, doi: 10.1016/j.patrec.2008.10.002.

[45] NICE (National Institute for Health and Care Excellence), 'Stroke Overview', *nice.org.uk*. https://pathways.nice.org.uk/pathways/stroke (accessed Oct. 19, 2016).

[46] R. T. Fothergill, J. Williams, M. J. Edwards, I. T. Russell, and P. Gompertz, 'Does use of the recognition of stroke in the emergency room stroke assessment tool enhance stroke recognition by ambulance clinicians?', *Stroke*, vol. 44, no. 11, pp. 3007–3012, 2013, doi: 10.1161/STROKEAHA.13.000851.

[47] S. Standring, N. Ananad, and H. Gray, Eds., *Gray's anatomy: the anatomical basis of clinical practice*, 41. ed. Philadelphia, Pa.: Elsevier, 2016.

[48] V. Garcia, R. A. Mollineda, and J. S. Sanchez, 'Theoretical Analysis of a Performance Measure for Imbalanced Data', in *2010 20th International Conference on Pattern Recognition*, Istanbul, Turkey, Aug. 2010, pp. 617–620, doi: 10.1109/ICPR.2010.156.

[49] A. Tharwat, 'Classification assessment methods', *Appl. Comput. Inform.*, p. S2210832718301546, Aug. 2018, doi: 10.1016/j.aci.2018.08.003.