

# Chapter 12

## Brain Haemorrhage Detection Through SVM Classification of Electrical Impedance Tomography Measurements



Barry McDermott, Eoghan Dunne, Martin O'Halloran, Emily Porter, and Adam Santorelli

### 12.1 Introduction

An important medical problem is the accurate and timely detection and diagnosis of the presence of a brain haemorrhage in a patient. Brain haemorrhages can be present in pathologies such as stroke and traumatic brain injury. Stroke (also known as a cerebral vascular accident (CVA)) features a disruption in the flow of blood to an area of the brain and a subsequent sudden loss of neurological function [1]. Stroke is the main cause of adult disability in the United States, the fourth largest killer, and costs the country in the region of \$70 billion annually in direct and indirect costs [2]. The aetiology of an incidence of CVA will either be related to a blockage of a blood vessel (ischaemic stroke) or the rupture of a blood vessel and subsequent bleed (haemorrhagic stroke). Crucially, as the treatment is radically different depending on the stroke type, it is vital to differentiate the cause as ischaemic or haemorrhagic [3]. For example, the use of the drug tissue plasminogen activator (tPA) is indicated for ischaemic patients but may be lethal to haemorrhagic patients [3]. Further, the patient outcomes following a CVA are directly linked to the length of interval between stroke onset and the start of treatment, with a worse prognosis associated with a delay. This underlines the need for both accurate and rapid detection of the presence, and equally the absence, of brain haemorrhage in stroke patients. Currently, definitive diagnosis is dependent on imaging modalities such as computed tomography (CT) and magnetic resonance imaging (MRI), which often suffer from accessibility issues for patients [4]. A device based on electrical impedance tomography (EIT), and augmented with machine learning (ML), may result in expedited initial diagnosis for CVA patients.

Traumatic brain injury (TBI) is any of a range of injuries that results from an external force impacting the head with a consequent disruption in brain function.

---

B. McDermott (✉) · E. Dunne · M. O'Halloran · E. Porter · A. Santorelli  
Translational Medical Device Lab, National University of Ireland Galway, Galway, Ireland  
e-mail: [b.mcdermott3@nuigalway.ie](mailto:b.mcdermott3@nuigalway.ie)

TBI results in an annual cost of \$61 billion in the United States [5]. Initial triage of TBI usually involves subjective assessment of severity with use of metrics such as the Glasgow Coma Scale [6]. Imaging (usually CT) is indicated for more severe TBI cases including incidents featuring haemorrhage [6, 7]. Better initial triage, including improved early detection of brain haemorrhage, potentially with the use of a modality like EIT coupled with ML, would improve the efficiency of the patient pathway through more objective selection of patients for gold standard imaging like CT. This need is illustrated by the estimation that a 10% reduction in the use of CT for minor TBI patients could save \$10 million annually in the United States [8].

It is emphasised that in both of these motivating clinical examples, the imaging of a bleed is unnecessary; it is the definitive ruling in or out of the presence of the bleed that is essential to the progress of the patient in the work-up.

The use of machine learning applied to medical diagnostics and other medical areas has been the scene of significant and important growth recently [9]. The fact that computers can process large amounts of data at high speed, combined with the rapidly increasing ability of machines to learn and improve performance over time, makes the technique amply suited to analysis and interpretation of biological data. A popular biomedical application for ML and the closely related and complimentary area of data mining (DM) has been interpretation of diagnostic imaging which includes data from such modalities as CT, MRI, and ultrasound [10–12]. However, ML and DM are now being used in a range of other areas such as genetic analysis, monitoring of physiology, and the evaluation of disability [13]. In this work, we examine the potential for EIT to be used to assess anatomy and physiology of the body, coupled with the ML technique of support vector machine (SVM) classification to be used in medical diagnostics, denoted as EIT-SVM.

Fundamental to this research is the use of computational (numerical) models. Computational models allow controlled development of a technology or algorithm with the ability to experiment and test parameters resulting in progression and a better final product before translation to patients.

In the next section, the basis behind EIT, including the nature of EIT measurement frames, which are the input to the classifiers, is described. Description of the SVM classifier and the computational modelling techniques used are also presented in Sect. 12.2. Section 12.3 then summarises the application of a linear SVM classifier to raw and minimally pre-processed EIT measurement frames, investigating the performance of the classifier in detecting bleeds in different scenarios, including variations in simulated noise, bleed size, bleed location, electrode positioning, and anatomy of the model. Section 12.4 presents methods to improve the performance and efficiency of the classifier, including changing the kernel function, selective pre-processing of the frames (including the use of sub-frames), dimensionality reduction and selection of specific features (using Laplacian scores and principal component analysis (PCA)), and finally using an ensemble classifier. Section 12.5 ends the chapter with a discussion and conclusion.

The content of this chapter builds on the research presented in [14], which was expanded upon in [15]. This previously published material from [14, 15] forms the core of Sect. 12.3, before new content in Sect. 12.4 is presented, which aims to improve the classifier performance.

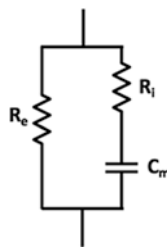
## 12.2 Technologies

This section introduces the core technologies used in this study; EIT and SVM classifiers. In the final part of the section, the computational modelling techniques and tools centred on a two-layer computational model of the head with variants, designed to emulate various test scenarios, is described.

### 12.2.1 Electrical Impedance Tomography

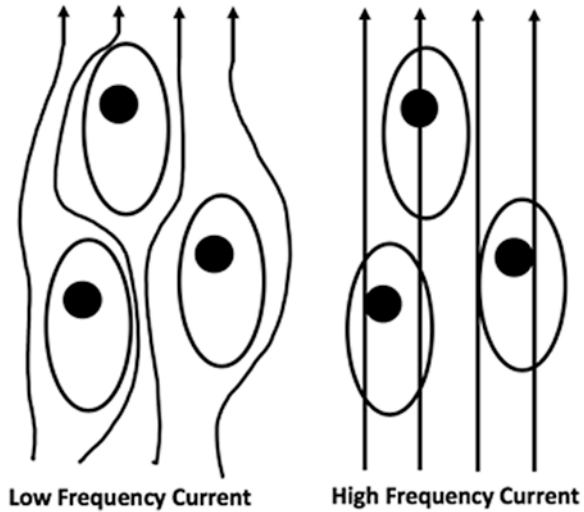
Electrical impedance tomography is an imaging modality and the basis of an ever-increasing and vibrant area of active research with a number of applications in the biomedical sphere [16]. EIT is based on the feature of biological tissue of electrical conductivity, as a result of the ion containing extracellular fluid (ECF) and intracellular fluid (ICF). The ECF bathes and surrounds cells, while the ICF refers to the fluid within cells. The cell membrane that surrounds the individual cells represents the border between the two compartments [17]. The conductivity is characteristic to each particular tissue. For example, blood is a good conductor, owing to the high ion content of the tissue, whereas bone is a poor conductor [16]. The conductivity is quantified in  $\text{Sm}^{-1}$  and is the inverse of the resistivity. Closely related is the concept of electrical impedance, which is the extension of the idea of resistance to alternating current (AC) circuits with a real (resistance) and complex (reactance) part. A biological tissue can be modelled as a three-part electrical circuit as shown in Fig. 12.1, where  $R_e$  is the resistance of the ECF,  $R_i$  the resistance of the ICF, and the cell membrane is modelled as a capacitor with capacitance  $C_m$  [18]. At low AC frequencies, the capacitive reactance of the cell membrane is high with the result of the overall impedance of the system being effectively  $R_e$ . At higher AC frequencies, current can pass through the cell as the capacitive reactance drops and, consequently, the overall impedance of the system drops. This concept is illustrated in Fig. 12.2 [18]. As conductivity is inversely related to impedance, it follows that the electrical conductivity of a tissue will increase with increase in AC frequency. The exact nature of the conductivity profile is a characteristic of the tissue in question.

EIT makes use of the difference in conductivity profiles of tissues. This difference in conductivity profiles is often used to generate an image of the region of



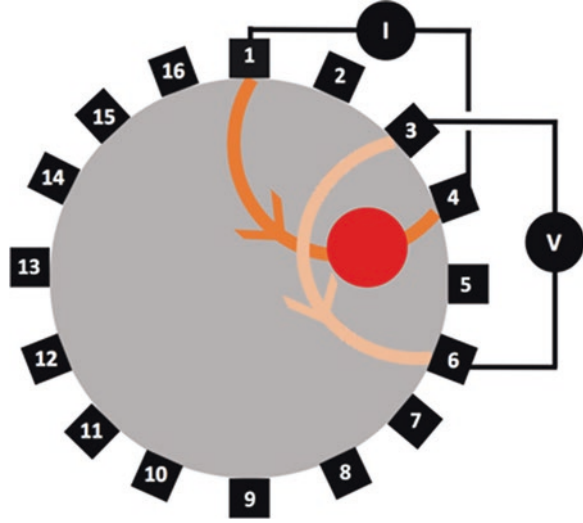
**Fig. 12.1** An electrical model of biological tissue with current having two paths of flow. One path is the ECF with resistance  $R_e$ , while the other is through the cell which has a capacitor like membrane with capacitance  $C_m$  and the ICF with resistance  $R_i$ . (Adapted from [18])

**Fig. 12.2** Current movement through tissues at low and high frequencies. At low frequencies, the capacitor effect of the cell membrane impedes current flow through the interior of the cell. At high frequencies, the capacitor effect becomes negligible and current can flow through the ICF. (Adapted from [18])



interest (ROI). EIT characteristically involves an array of electrodes positioned on the boundary of the ROI. A popular electrode configuration is that of a ring of electrodes, typically with 8–64 electrodes surrounding the region [18]. Electrical current is then injected through a pair of electrodes (“stimulation”) and the resultant voltages measured at all other electrode pairs. The injection pair is then changed, and voltage measurements are taken between the new measuring pairs. The overall pattern of stimulation and measurement constitutes an EIT “protocol”, with each individual measurement referred to as a “channel”. The complete set of channels comprises the EIT measurement “frame”. EIT systems typically operate in the 1 kHz–2 MHz frequency range, with injected currents of the order of  $\mu\text{A}$  to low mA [16]. Importantly, international safety standards limit the current to 100  $\mu\text{A}$  rms for currents up to 1 kHz with the limit rising to an absolute limit of 10 mA when operating above 100 kHz [16, 19]. The electrode configuration and number, protocol, current amplitude, and frequency are application dependent. In Fig. 12.3, a sample EIT measurement channel, with a “skip 2” protocol and the electrodes arranged in a 16-electrode ring surrounding a circular body, is illustrated. In this protocol, each electrode is paired to the electrode three positions away from it (i.e., with 2 in-between electrodes skipped over). The ROI illustrated in Fig. 12.3 is of homogenous tissue with one region of differing conductivity present (illustrated as a red circle). The presence of this tissue affects the voltage at the different measurement electrodes. For example, at 50 kHz, a bleed is more electrically conductive than the surrounding brain parenchyma [15]. Hence, for a given channel with a constant injection current, the measured voltage will be smaller in magnitude if a bleed is present than if there is only healthy brain tissue present. This trend follows from Ohm’s law, described in Eq. (12.1) where  $V$  is the voltage,  $I$  is the current; and  $\sigma$  is the electrical conductivity,

**Fig. 12.3** An EIT measurement channel from a “skip 2” protocol involving a 16-electrode ring around a circular ROI. Current is injected between electrodes #1 and #4 (orange arc) with voltage measured between electrodes #3 and #6 (beige arc). A tissue with different conductivity to the bulk tissue is illustrated by the red circle



$$V\sigma = I \tag{12.1}$$

If a bleed is larger, the measured voltages will be smaller. Further, channels nearer to the bleed are affected more by the presence of the bleed than those further away, as EIT is more sensitive to changes where current density is higher [19]. Hence, information regarding the presence, nature, and location of the various tissues in the ROI are theoretically encoded in the final measurement frame.

For a 16-electrode ring, a given injecting pair results in 16 measurement pairs. However, it is common practice not to take measurements from either of the injecting electrodes hence 13 measurements are taken [19]. Over the course of a complete protocol, there will be 16 injecting pairs and so a complete frame will be made up of a total of 208 channels. The number of channels in the frame is summarised in Eq. (12.2):

$$N_M = N_E(N_E - 3) \tag{12.2}$$

where  $N_M$  is the number of measurements when using  $N_E$  electrodes.

The relationship between the conductivity profile of the ROI and the values in EIT measurement frames is given by the EIT forward and inverse problems. The EIT “forward problem” refers to the prediction of the measured values given the complete conductivity profile of the body [19]. In the computational model used in this study (described in Sect. 12.2.3), the finite element method (FEM) was used to solve the forward problem for the geometry of interest, which is that of the human head. An important calculated parameter is the sensitivity matrix (the Jacobian,  $J$ ). The Jacobian gives the sensitivity of each measurement to a conductivity change within the ROI [19]. The “inverse problem” of EIT involves calculating the conductivity profile of the interior of the body of interest given a set of measurements. This

is an ill-posed inverse problem (the number of “voxels” to be assigned conductivity values is typically larger than the number of measurements) with the need for regularisation techniques in order to obtain the most reasonable solution [19]. The result is a conductivity map of the interior of the ROI.

EIT is a non-invasive modality with a high temporal resolution [19]. However, it has drawbacks, including poor spatial resolution, low sensitivity to conductivity changes at a depth from the boundary, and high sensitivity to electrode modelling errors [19, 20]. Attempts to overcome these challenges and reconstruct useful images have seen different EIT modalities established, many of which rely on difference imaging in order to minimise errors. The most successful EIT modality to date is that of time difference EIT (tdEIT), which reconstructs an image based on differencing frames of a “before” and “after” measurement. This modality has been applied to the monitoring of physiological functions in regions such as the thorax where there is a large contrast between inspiration (air in the lungs) and expiration (air emptied from the lungs) [19]. Static scenes are more challenging, without a satisfactory modality for imaging established to date. In a complex region such as the head, where the high impedance of the skull severely dampens the stimulating current, the imaging of static pathologies such as an established bleed has been proven to be difficult [18, 21].

In this work we examine the viability of using EIT measurement frames in a more direct manner, without the mathematically difficult and challenging image-reconstruction step. In scenarios that do not require an immediate image, such as stroke classification or TBI triage, it may be sufficient to definitively rule in or out a bleed. The information relating to the presence or absence of such a perturbation in the body of interest is encoded in the EIT measurement frame. The basis for this is the a priori knowledge that there is a notable difference in conductivity between blood and normal brain parenchyma [22]. ML offers techniques that can potentially learn from raw or processed EIT frames and classify the frame as positive or negative for a bleed. In the next section we examine such a ML technique: SVM classifiers.

### ***12.2.2 Support Vector Machine (SVM) Classifiers***

A definition of ML proposed by Mitchell is that of a “computer program that improves its performance at some task through experience” [23]. Different types of “tasks” exist when referring to ML. One of the major task types is classification. In a classification task, each observation is assigned to one of a number of designated classes or labels. Each observation consists of several features (traits) that define it. These features are used as the inputs to the ML algorithm. The algorithm will then use this information to create a trained model that can be used to predict the class that future observations belong to. In the context of the work presented here, the input features are the EIT measurement frames (processed or un-processed) obtained from numerical simulations of the head in which a bleed is or is not present. The two

classes defined in this scenario are “bleed” or “normal”, denoted as +1 and -1, respectively. The task of the classifier is to use the measurement frames, with the channel measurement values (or equivalent if processed) as features, to correctly predict whether future observations belong to the “bleed” or “normal” class.

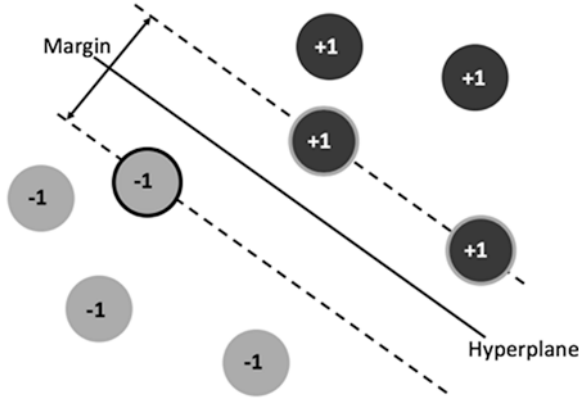
SVMs are a group of popular ML algorithms commonly employed for binary classification. They have been used in previous biomedical applications, including the use of microwave signals to classify whether a breast scan is considered healthy or tumourous [24–26], and electrical impedance spectroscopy signals for classification of breast [27–29] and prostate [30] as diseased or normal. The use of EIT measurements in ML algorithms is a relatively new area of research. Some work has been done in the area of bladder volume estimation [31, 32] and the focus of this chapter, brain haemorrhage detection, has been explored by our group [14, 15, 33].

As is typical in the use of SVMs and related classifiers, the basis of the algorithm is the creation of a model using a training set. This training set consists of observations with the true class known (supervised learning) or unknown (unsupervised learning). The performance of the trained classifier can be assessed by analysis of the results of classifying a test set of previously unseen observations. The trained and tested classifier can then be used to classify new observations; assuming the training and testing process was properly implemented, the classifier will perform in-line with expectations even on new observations.

The core of the SVM model is the creation of a hyperplane that best separates observations from the two classes. A representation of a two-dimensional (2-D) hyperplane (a line) separating the observations classified as +1 or -1 is shown in Fig. 12.4. In the training phase, a mathematical model of the hyperplane and margin is developed with the training observations having  $n$ -dimensions ( $n$  number of features). The hyperplane is used to decide whether future observations belong to either the +1 or -1 class. When the data is not perfectly separable (there exists no margin that guarantees no observations between it and the hyperplane), “soft” margins can be used to ignore those outliers [34]. An important parameter when using SVM classifiers is the kernel, which defines the function used to generate the hyperplane. A linear kernel is the simplest type of kernel, which offers potential advantages including speed, low computational overhead, and an ease of implementation [35]. Other kernel functions, including the non-linear Gaussian Radial Basis Function (RBF), can be used to define the hyperplane [24, 27]. Additional information about the mathematical formulations governing the various SVM algorithms can be found in [34, 36].

The performance of a classifier can be reported by a number of different metrics. A key result is the confusion matrix, which compares the expected and predicted classes. An example of a confusion matrix, for a binary classifier, is shown in Fig. 12.5. As shown, a true positive (TP) refers to observations where the expected and predicted classes are +1, and a true negative (TN) where the expected and predicted classes are -1. A false positive (FP) is where the expected class is -1 but is predicted as +1, with a false negative (FN) the opposite.

Two key metrics of performance derived from the confusion matrix are the sensitivity and specificity. Sensitivity (TP Rate) is the proportion of observations clas-



**Fig. 12.4** Visualisation of a SVM classifier. The trained SVM classifier model calculates the optimal hyperplane that separates the two classes (shown here as black circles for the +1 class and grey circles for the -1). The margin of the hyperplane is as wide as possible (for a “hard” margin), with the borders of the margins defined by the cases called “support vectors” represented here as circles with a visible outer shell. The hyperplane in this case is 2D (a line)

**Fig. 12.5** The confusion matrix for a binary classifier with classes ±1. The expected (true) class and predicted class assigned to cases are compared

		Expected	
		+1	-1
Predicted	+1	TP	FP
	-1	FN	TN

sified as +1 out of the total that are truly +1. Specificity (TN Rate) is the proportion of observations classified as -1 out of the total that are truly -1. Accuracy is the proportion of correctly classified cases out of the total number of cases. These metrics are defined in Eqs. (12.3)–(12.5),

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{12.3}$$

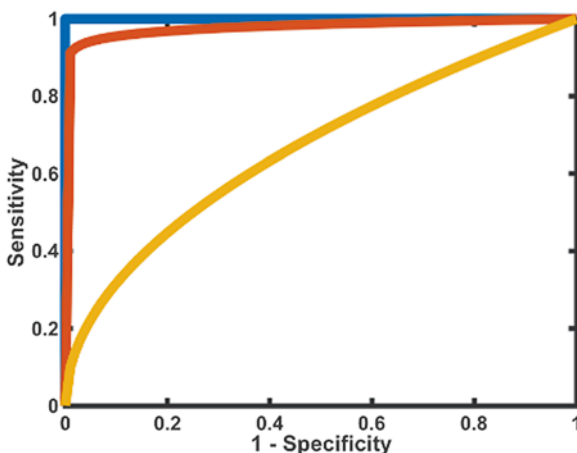
$$\text{Specificity} = \frac{TN}{TN + FP} \tag{12.4}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{12.5}$$



The above Eqs. (12.3)–(12.5) imply the values of sensitivity, specificity, and accuracy range between 0 and 1, with 1 indicating perfect performance for that metric (this range is equivalent to 0–100%).

The receiver operating characteristic (ROC) curve is a plot of sensitivity versus  $(1 - \text{specificity})$  [35]. It is a useful tool to illustrate the trade-off between sensitivity and specificity. If a classifier is 100% sensitive and 100% specific, as is ideal, then the ROC curve is said to have an Area Under the Curve (AUC) of 1. In the proposed application of brain haemorrhage detection applied to stroke and TBI, this is the ideal performance of a trained classifier. However, in cases where the performance is imperfect, this is reflected in a ROC curve where the AUC is  $<1$ . In such cases, it is possible to adjust the operating point of the classifier with a trade-off between sensitivity and specificity. For brain haemorrhage detection, it could be proposed that sensitivity is more important than specificity. A reduced specificity indicates an increased level of FPs which is not ideal but the alternative of reduced sensitivity with a consequent increased level of FNs would result in patients with bleeds being classified as normal and potentially receiving a dose of lethal tPA in the case of stroke or not receiving timely CT scan in the case of TBI. Hence, for brain haemorrhage detection, the optimal point of operation of the classifier is the point on the ROC curve where sensitivity is 1 while minimising  $(1 - \text{Specificity})$ . An example of three ROC curves is shown in Fig. 12.6.



**Fig. 12.6** Receiver operating characteristic (ROC) curves are a plot of sensitivity versus  $(1 - \text{specificity})$  and show the trade-off possible between sensitivity and specificity. An ideal ROC curve has an area under the curve (AUC) of 1 with an example of this shown as the blue trace. Here, an operating point where both sensitivity and specificity are both 100% is at  $(0,1)$ . The red and yellow traces show imperfect ROC curves where  $\text{AUC} < 1$ . In this case it is possible to maximise sensitivity by moving to the operating point shown with the penalty of reduced specificity. At any given point, the red curve gives a better sensitivity/specificity trade-off compared to the yellow curve. The yellow curve only offers a sensitivity of 1 where specificity is 0, which would result all observations being classified as +1

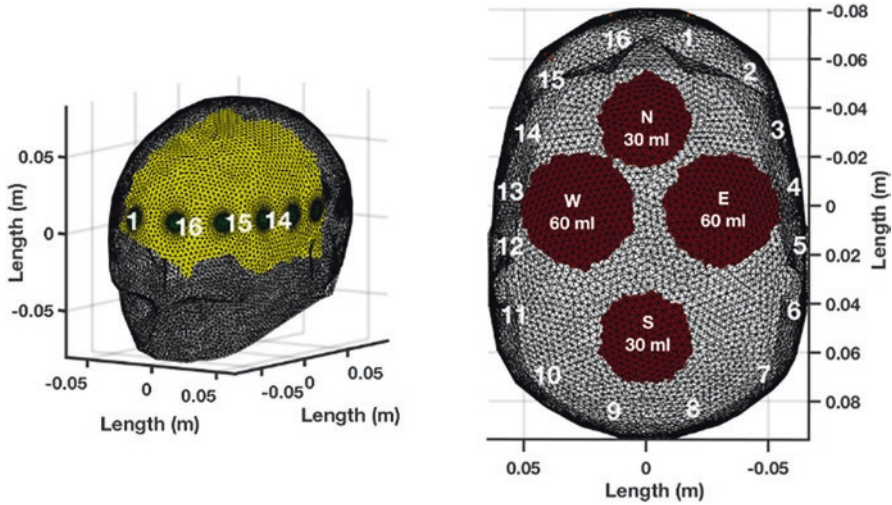
### 12.2.3 *Computational Modelling Techniques*

The core computational model used in this work was a FEM model of the human head and brain. The head is an anatomically complex and intricate structure [37], but for the purposes of EIT, simplifications can be made by focusing on those tissues that have a significant effect on the conduction of electrical current. Typically, EIT simulations use a four-layer model, which includes the brain as the innermost layer, the electrically conductive cerebrospinal fluid (CSF) layer immediately external to it, the highly resistive skull, and the moderately resistive scalp [18]. Naturally, more complex models exist and may be relevant depending on the research question. For example, physical phantom models which model the differing resistivity across the skull are reported in the literature [38, 39].

In this work the head was designed as a two-layer structure. The layers were anatomically accurate representations of the brain and an aggregate outer layer comprised of the tissues external to the brain (the scalp, skull, and CSF layers), derived from anatomically realistic stereolithography (STL) files of the head [40] and brain [41]. As described in [15], this simplified model facilitated the development of an equivalent physical phantom, allowing comparison between the computational results and the phantom results. Further, it was computationally “light” and allowed rapid development of variant test models.

The STL files were meshed into a FEM model using the software packages EIDORS [42], which itself uses Netgen [43] and Gmsh [44] for meshing. EIDORS is an open source set of tools designed to aid the development of EIT (and the related area of diffuse optical tomography), and is written for use with MATLAB [45] and Octave [46]. Using EIDORS, a 16-electrode ring was placed on the exterior surface of the FEM model at the approximate level of the inion-nasion line symmetrically across the sagittal plane. The electrode ring defined a transverse plane, and a refinement of the mesh at the contact points [47] was carried out. This constituted the “base numerical model”. Modifications were made to expand this model to create a total of 243 models of the “normal” (bleed free) head. These 243 models were created by varying the head and brain anatomy ( $\pm 5\%$  in size in each Cartesian axis), and modifying the electrode position ( $\pm 2$  mm in the positioning of the ring in terms of height). More complete details on these 243 models can be found in [33].

Bleeds were modelled as spheres within the brain layer using the computer-aided design package Autodesk Fusion 360 [48]. The two primary bleed sizes used were 30 ml and 60 ml, with some experiments using bleeds of smaller volume (down to 5 ml). In stroke patients, a 30 ml bleed is a threshold size associated with worse outcomes, with 60 ml a threshold for significant mortality [49, 50]. These bleeds were placed in each of the 243 normal models at each of the 4 cardinal points of north (‘N’, front), south (‘S’, back), east (‘E’, right), and west (‘W’, left) in the plane of the ring, at the exterior of the brain. This resulted in 1944 “bleed” head models, each model with one bleed of a given size and location. The electrical conductivity, fundamental to EIT, can be assigned to each FEM model element depending on which tissue is being modelled. The realistic conductivity values of  $0.1 \text{ Sm}^{-1}$ ,



**Fig. 12.7** Computational (numerical) model of the head. Left: The base numerical model is an anatomically accurate two-layer model of the brain and aggregated tissues external to the brain. The 16-electrode ring is shown with electrode contact areas in green and white numbering of some electrodes for orientation. Right: Removal of the brain layer to illustrate the size and positioning of the bleeds. The positioning of the electrodes #1–16 are shown as a ring of white numbers. Bleeds of volume 30 ml and 60 ml are positioned in the north, south, east, and west locations as shown. A given model will contain either no bleed or only one bleed. The bleeds are positioned immediately at the exterior of the brain layer in the plane of the ring. The different colouring of the layers represents the different electrical conductivities;  $0.1 \text{ Sm}^{-1}$  for the aggregate outer layer (white),  $0.3 \text{ Sm}^{-1}$  for the brain (yellow), and  $0.7 \text{ Sm}^{-1}$  for bleed (burgundy). (Adapted from [15])

$0.3 \text{ Sm}^{-1}$ , and  $0.7 \text{ Sm}^{-1}$  were used for the aggregate outer layer, the brain layer, and the bleeds, respectively [15]. EIDORS allows defining of the EIT protocol (“skip 2” for this work) and the subsequent generation of measurement frames from a FEM model. This suite of 243 normal and 1944 bleed heads allowed the emulation of a wide variety of test situations, with these experiments and results described in later sections. In Fig. 12.7 the base numerical model is shown along with the positioning of the 30 ml and 60 ml bleeds within the model.

### 12.3 SVM Applied to Raw EIT Measurement Frames with Analysis of the Effect of Individual Variables on SVM Performance

Initial experiments focussed on the effect of individual variables such as measurement noise, bleed size and location, electrode position, and anatomy. These variables constitute important parameters. Understanding the effect they have on EIT measurement frames, and consequent performance of the SVM classifier, can help

inform future research experiment decisions. The results and conclusions from these experiments are briefly summarised herein; for more detail, refer to [15].

In each experiment, measurement frames generated from a subset of FEM models were used to train and test a linear SVM with no (raw) or minimal processing of the frames performed prior to use of the classifier. Minimal processing constituted sorting the values in the measurement frames in order of numerical value. This simple pre-processing step was found to aid performance in certain scenarios (see Sect. 12.3.2). In all cases, the training and test sets comprise an equal number of measurement frames from normal models and models with bleeds present. In this section, a linear SVM classifier was implemented for all experiments. The classifier was trained with 80% of the data set and then tested with the remaining, unseen, 20%. The classifier is optimised by generating a ROC curve in training. The generalised accuracy in training is used to choose a point on the ROC curve that maximises sensitivity. The final classifier is re-trained at this operating point and the performance of the trained classifier on the test set data is used to obtain the performance metrics presented in this section.

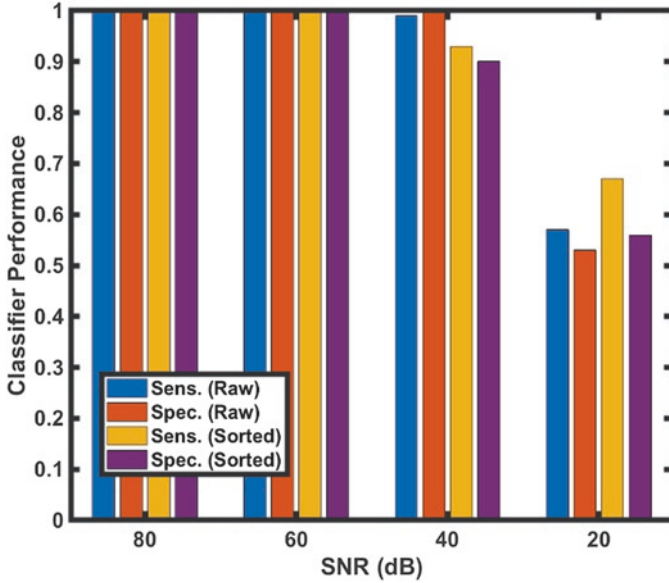
### 12.3.1 *The Effect of Noise*

The amount of noise in a measurement frame can be controlled by adjusting the signal-to-noise ratio (SNR) using tools supplied by EIDORS. The SNR is defined in Eq. (12.6), where the noise is a numerical value in dB,

$$\text{SNR} = 20\text{Log}_{10}(\text{Signal/Noise}) \quad (12.6)$$

In order to add noise to a measurement frame, EIDORS generates a vector (of same size as the measurement frame) of normally distributed random numbers with the values in this vector then scaled by multiplication of the ratio of the Euclidean norms of the measurement frame and noise vector, before further scaling by division by the desired SNR value. This final scaled vector of noise values is added to the measurement frame, resulting in a “noisy” frame.

EIT applications such as thoracic imaging may be successful with a system offering a SNR of 30–40 dB, whereas more demanding neural applications, that may involve smaller changes and issues such as the skull dampening, may require systems capable of 80 dB and higher [51]. In order to study the effect of noise on performance, the base numerical model was used to generate normal frames, with the 30 ml and 60 ml bleeds placed in the north location to generate bleed frames. Noise was added to the measurement frames so that a SNR of 80 dB, 60 dB, 40 dB, and 20 dB was obtained. These measurement frames at the four SNR levels were used as the input features for a linear SVM classifier. Separate experiments were performed with the raw and sorted frames. The results for the sensitivity and specificity are shown in Fig. 12.8. The results show that the classifier performs well at a

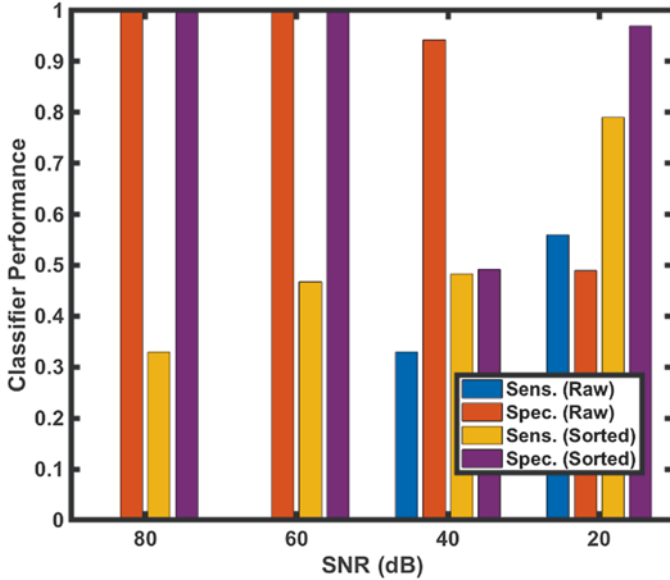


**Fig. 12.8** Effect of noise on classifier performance. Measurement frames from normal and bleed cases have SNR levels of 80 dB, 60 dB, 40 dB, and 20 dB. The frames are unaltered (raw) or sorted by numerical values. The results of the classifier in terms of sensitivity (Sens.) and specificity (Spec.) for each scenario are reported above

SNR of 80 dB and 60 dB (sensitivity and specificity at or near 1), with a falloff in performance at 40 dB and poor performance at 20 dB.

### 12.3.2 Effect of Bleed Location

The base numerical model was used to generate normal frames, with 30 ml and 60 ml bleeds placed at the north location in the training set. The test set was created from frames generated by placing 30 ml and 60 ml bleeds at the three other cardinal points. Hence, the test set had novel bleed locations in comparison to the training set. The results for sensitivity and specificity for the raw and sorted frames are reported in Fig. 12.9, with the experiment performed at SNR levels of 80 dB, 60 dB, 40 dB, and 20 dB. The classifier is seen to fail at bleed detection (sensitivity of 0) at 80 dB and 60 dB when using raw measurement frames. This indicates an inability to cope with bleeds in locations different to that of the training set. The specificity is near 1 at 80 dB and 60 dB as expected as it is a measure of the ability to detect normal cases, which are the same in the training and test sets. The sensitivity then paradoxically increases at lower SNR levels, but an explanation may be the introduction of general inability to differentiate normal from bleed at lower SNR levels as evidenced by the drop in specificity. The simple pre-processing step of sorting the frames by channel value helps increase the sensitivity from 0 to 0.33 and 0 to 0.47



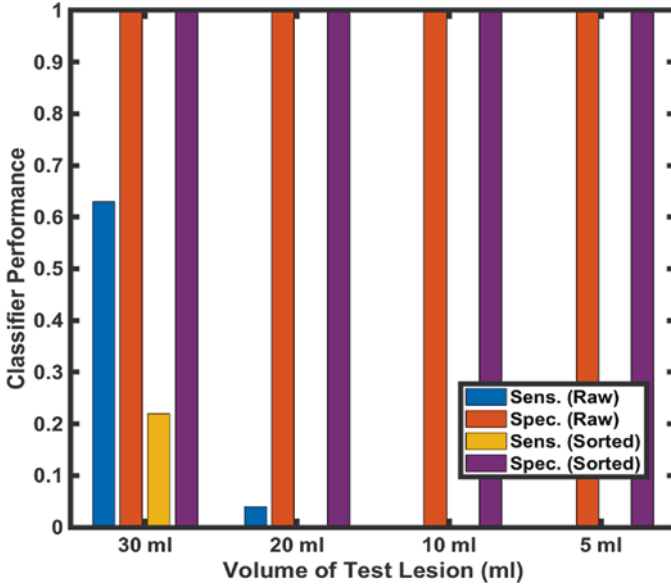
**Fig. 12.9** Effect of bleed location on classifier performance. The classifier performs poorly at detecting bleeds, as judged by the sensitivity, in novel locations to that used in the training set at 80 dB and 60 dB with a paradoxical improvement seen at 20 dB SNR. The implementation of a simple pre-processing step, sorting the frames, improves sensitivity from 0 at 80 dB and 60 dB to values of 0.33 and 0.47 respectively. Specificity is not affected as severely, but this is expected as the normal cases are the same in both the training and test sets

at 80 dB and 60 dB, respectively. The sorting results in channels located near the bleed location (with smaller measured voltages as explained in Sect. 12.2.1) to cluster in the same area of the frame regardless of location. In the absence of the bleed, this area of “clustered” channels will have higher values characteristic of the no bleed case.

However, effectively these results suggest that accurate detection of bleeds in unseen locations is challenging. As described in [15], it is possible to improve performance by working at an adjusted point on the ROC curve which improves sensitivity at cost to specificity.

### 12.3.3 Effect of Bleed Size

As described in Sect. 12.2.1, the larger the size of the bleed, the greater the voltage measurements will deviate from normal values. To investigate this effect, measurements from the base numerical model without a bleed and then with bleeds of 60 ml, 30 ml, 20 ml, 10 ml, and 5 ml at each of the four locations were generated at 60 dB SNR. The 60 ml bleed subset was used to train the classifier, which was then tested with each of the smaller volumes in turn at 60 dB SNR. These results are shown in



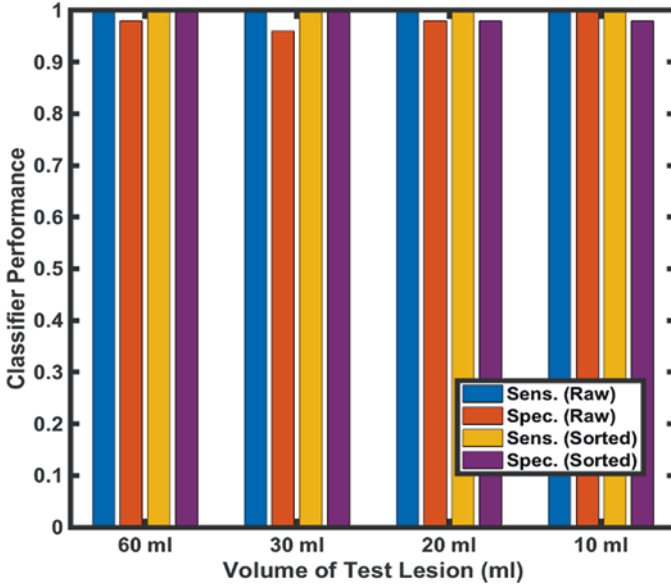
**Fig. 12.10** Effect of bleed size on classifier performance. A 60 ml bleed size is used in the training set with the test set comprised of bleeds of smaller volume. All experiments are performed at 60 dB SNR. The SVM classifier is unable to detect smaller bleeds than those trained with; however, the 30 ml bleed is detected with a sensitivity of 0.63 when using the raw measurement frames

Fig. 12.10, which indicates a general inability to detect bleeds smaller than those trained with. The best value for sensitivity observed was 0.63 when using raw frames to detect the 30 ml bleed. Again, the TN rate (specificity) is not affected as the normal cases are the same in both the training and test sets.

Repeating the experiment using the 5 ml bleed in the training set and testing with each of the larger bleeds gives the results shown in Fig. 12.11, which shows generally good performance (sensitivity and specificity near 1) for detection of each of the larger bleed sizes. As discussed in Sect. 12.2.1, the size of voltage measurements is related to bleed size, with larger bleeds affecting measurements more than smaller ones. Hence, training with a small bleed “sensitises” the classifier to the bleed type, with larger bleeds resulting in even more pronounced changes in voltages and hence easier classification as bleeds.

### 12.3.4 Effect of Electrode Positioning

Recent literature suggests that EIT is sensitive to errors in electrode positioning [52]. In this experiment, the base numerical model is used to generate measurement frames with and without all permutations of the 30 ml and 60 ml bleed at all four positions. The test set then comprises of measurement frames from equivalent



**Fig. 12.11** Effect of Bleed Size on Classifier Performance. A 5 ml bleed size is used in the training set with the test set comprised of bleeds of larger volume. All measurement frames have a SNR of 60 dB. The SVM classifier performs well (Sensitivity and Specificity near 1) for all test volumes

models that differ only in the position of the electrode ring, with the ring displaced  $\pm 2$  mm with respect to the original, parallel to the plane of the original. This was to replicate operator error in placing a ring on a patient's head. This analysis was performed at a SNR of 60 dB. This small error in electrode positioning causes a decrease in the sensitivity by 0.05 and 0.03, for the raw and sorted measurement frames, respectively. There is no impact on the specificity from this small electrode displacement.

### 12.3.5 Effect of Normal Variation in Between-Patient Anatomy

The ability of the classifier to classify normal from bleed in unseen anatomies is assessed in this experiment. The training set is made up of measurement frames calculated from the base numerical model with and without the 30 ml and 60 ml bleed at all four locations. The test set is comprised of measurement frames from 80 other anatomies that differ in the size of both the aggregate outer layer and brain layer by  $\pm 5\%$  in the three Cartesian axes but have the electrode ring in the same position (as described in Sect. 12.2.3). These anatomies are used to generate measurement frames with and without the equivalent bleeds present. Noise is added to all measurement frames, leading to a 60 dB SNR. The results indicate that the classifier struggles with unseen anatomy; the sensitivity and specificity were below 0.60



for both raw and sorted measurement frames, a decrease in over 0.40 from the classifier performance with known anatomies. Further analysis showed that an excess of brain tissue or lack of outer tissue in a test model compared to the training model was often misclassified as a bleed. Conversely, lack of brain tissue or excess outer tissue compared to the training model was often misclassified as normal.

## 12.4 SVM Applied to EIT Processed Measurement Frames

Section 12.3 examined the use of a linear SVM classifier to classify FEM models of the head and brain as having a bleed or no bleed. The emphasis was on the effect of individual variables such as noise, bleed location and size, electrode positioning, and head anatomy on classifier performance. The section constituted an initial exploratory study with minimal attempt to intelligently select features for input to the classifier or indeed in selection of the best type of SVM classifier. In this section, research into these areas is reported, starting with the effect of a change of kernel on performance. Then, the effect of pre-processing and selecting input features is examined.

In all the experiments in this section, all 243 normal models and 1944 bleed models are used to generate measurement frames. As described in Sect. 12.2.3 (and elaborated on in [15]), the starting STL files of the head and brain are each distorted by  $\pm 5\%$  in each Cartesian axis as well as in all three axes simultaneously, giving nine distinct head and nine distinct brain anatomies. FEM models of all combinations of these brain anatomies as well as the electrode ring in one of three heights resulted in 243 normal models. Bleed models were based on every combination of these normal head models combined with one of either the 30 ml or 60 ml bleed in one of the four locations, leading to a total of 1944 bleed models. An equal number of frames from the normal head set and bleed head set were used to generate 155,520 measurement frames.

A consistent method is applied in this section to optimise the performance of the SVM classifiers. First, the data is separated into five separate folds, each with a unique training data set and testing data set that is made up of 80% and 20% of the original data set, respectively. The training data set is used to optimise the SVM classifier hyper-parameters, namely the box constraint and kernel scaling factor. A Bayesian optimisation procedure is implemented to identify the hyper-parameters that lead to the greatest generalised accuracy across fivefold cross-validation. Once identified, a final trained SVM classifier is created with these optimised hyper-parameters. The excluded testing data set is then used to obtain performance metrics for the final classifier. This procedure is then repeated for all five of the unique training-testing data pairs, and final classifier performance is presented as the mean and standard deviation (STD) across these five iterations. This nested testing methodology, which has been used previously in the literature [26, 53], provides a more generalised and robust indication of classifier performance.

### ***12.4.1 Radial Basis Function Kernel Compared to Linear Kernel***

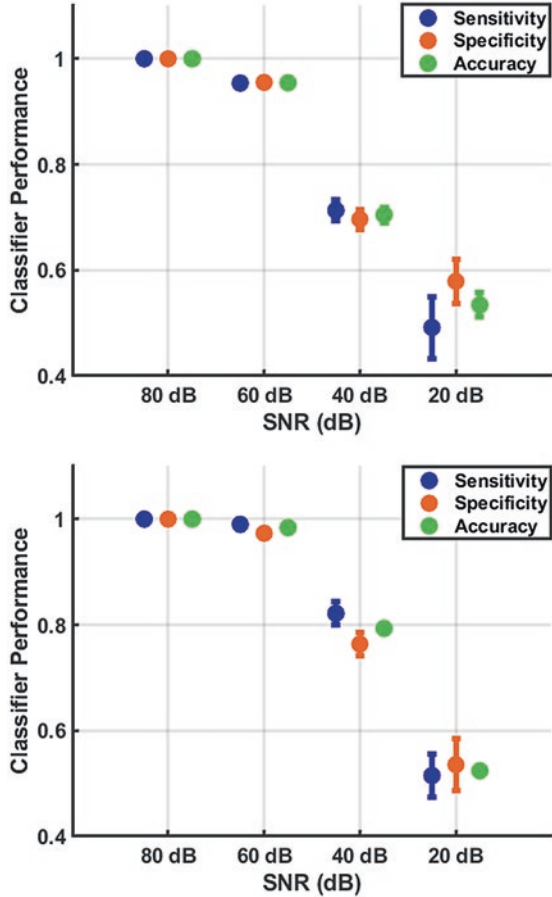
The RBF kernel can be used for SVMs when the relationship between the features and labels is non-linear, has less hyperparameters than a polynomial kernel, and has less numerical difficulties [54]. The RBF can be conceptualised as a flexible membrane that fits through sample points while minimising the curvature. Hence, the hyperplane is a “gently varying surface” and is suitable for scenarios where the data points (measurement values) do not change dramatically within a short distance in the  $n$ -dimensional hyperspace.

The first investigation of this section involves comparing the use of the linear and RBF kernels with a SVM classifier trained and optimised across all four SNR levels (80 dB, 60 dB, 40 dB, and 20 dB). In Fig. 12.12, the classifier performance, in terms of the sensitivity, specificity, and accuracy, for both the linear-SVM (top) and the RBF-SVM (bottom), is shown. Each dot on the plot denotes the mean classifier performance across the fivefold testing, with error bars representing the standard deviation range. While perfect classifier performance ( $1.00 \pm 0.00$  in all metrics) is achieved by both kernel types at 80 dB, it is observed from this figure that use of the RBF kernel can improve the classifier performance, notably at the 60 dB and 40 dB SNR levels; there is an increase in the mean accuracy between approximately 0.03 (3%) and 0.09 (9%), respectively, at these SNR levels when using the RBF kernel. When the SNR decreases to 20 dB, the performance of both classifiers approaches that of guesswork, with the mean accuracy only slightly above 50%, indicating that the changes in impedance due to the presence of the bleed are embedded within the noise. This finding suggests that hardware should guarantee an SNR well above 20 dB. From Fig. 12.12, we can in fact infer that the SNR for a hardware system should be on the order of 60 dB to expect accurate detection of brain bleeds. The improvement with the use of the RBF kernel over the linear kernel provided the motivation for the use of this kernel in all the following sections of Sect. 12.4.

### ***12.4.2 Frame Pre-processing***

In the previous sections, the classifier input features were the unprocessed EIT measurement frames, with the injection channels removed. This section will explore the use of various pre-processing techniques, ranging from manually chosen feature-extraction methods, such as taking the mean of sub-frames, to using electrode pair proximity to decide input features, to variance-based methods such as Laplacian scores and PCA. These feature extraction methods are carried out on data at all four SNR levels (80 dB, 60 dB, 40 dB, and 20 dB), with the RBF-SVM classifier optimised as described in Sect. 12.4.1. As before, classifier performance is presented as the results across fivefold testing.

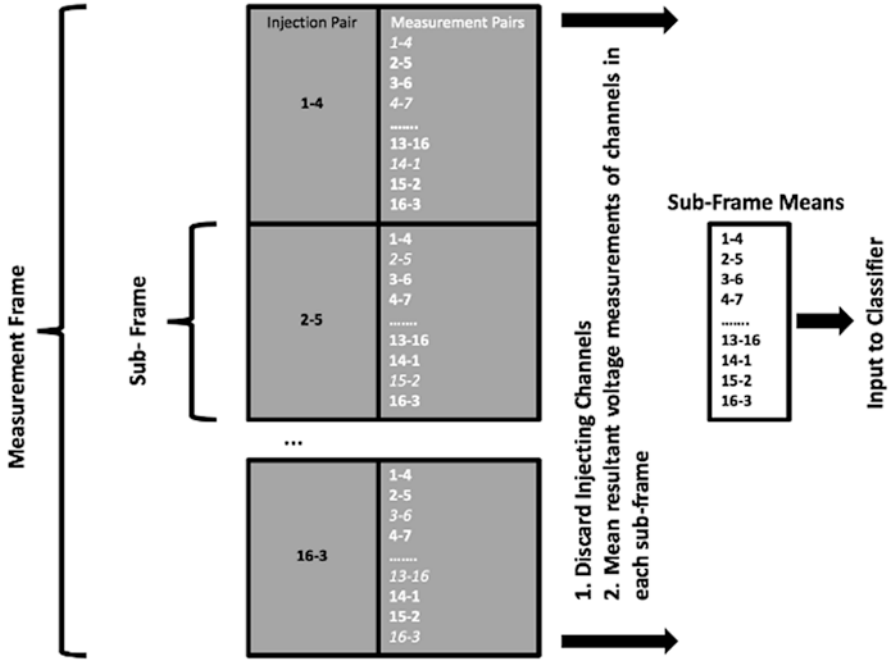
**Fig. 12.12** Comparison of classifier performance using the linear-SVM (top) and RBF-SVM (bottom). Each dot on the plot denotes the mean classifier performance across the fivefold testing, with the error bars representing the standard deviation range at the respective SNR level. There is a significant improvement in performance when using the RBF kernel, notably at the 60 dB and 40 dB SNR levels



### Sub-frame Means

A sub-frame is defined as the set of measurement channels associated with a given injection pair. A measurement frame from a 16-electrode array using a skip 2 pattern will have 16 such sub-frames, each with 13 channels (three channels are removed as they use either of the injecting electrodes). The 13 voltage measurements in each of the 16 sub-frames are averaged, with the resulting 16 mean-values used as the input features to the classifier. This reduces the dimensionality of the input from 208 features to 16 features. The pre-processing work-flow is shown in Fig. 12.13 below.

The performance of the RBF-SVM classifier using the sub-frame means as inputs is reported in Fig. 12.14 at each SNR level as the mean  $\pm$  standard deviation of the sensitivity, specificity, and accuracy after fivefold cross validation and Bayesian optimisation. As seen, the performance at 80 dB is excellent, being near  $1 \pm 0$  for all metrics, with a fall off at lower SNRs with, for example, sensitivity at

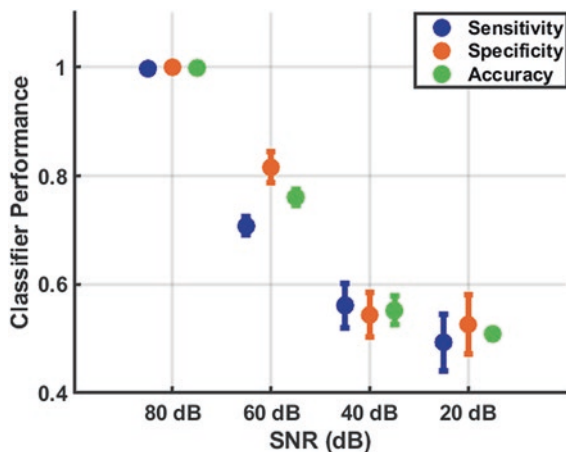


**Fig. 12.13** Generating the mean of each sub-frame. Each sub-frame is made up of the 16 channels associated with a given injection pair. Removal of channels involving either of the injection pair electrodes gives 16 sub-frames each with 13 channels. The mean of the voltage measurements from each set of 13 channels in a given sub-frame is used, leaving 16 values, the sub-frame means, which are used as inputs to the classifier

approximately  $0.71 \pm 0.02$  at 60 dB and all metrics at approximately 0.5 at 40 dB and 20 dB. It is noteworthy however that near identical performance is achieved at 80 dB relative to that of using full measurement frames (with a difference of  $<0.01$  (1%) in all metrics), despite the significant drop in the number of features. Such a reduction in dimensionality, with nearly no effect on performance, would result in a less computationally expensive algorithm.

### Near and Far Sub-frame Channels

In this section we explore using selected channels of each measurement sub-frame based on the physical locations of the recording electrodes relative to the injection pairs. Specifically, we analyse classifier performance when using “near” sub-frame channels and “far” sub-frame channels. The “near” sub-frame channels are defined as the seven channels nearer in physical location to the injecting pair of a given sub-frame. The “far” sub-frame channels are defined as the six channels further in location from the injecting pair. The complete set of near channels from each sub-frame are amalgamated and used as the input to the classifier with the same process performed to the far channels. This process reduces the input feature size to 112



**Fig. 12.14** Performance of the RBF-SVM using sub-frame means as input features. Each dot on the plot denotes the mean classifier performance across the fivefold testing, with the error bars representing the standard deviation range at the respective noise level. The performance at 80 dB SNR is near the ideal of  $1 \pm 0$  for all metrics, comparable to the performance achieved when using the complete frames. However, performance falls off quickly at lower SNRs, with all metrics below 0.85 at 60 dB and at approximately 0.5 at 40 dB and 20 dB

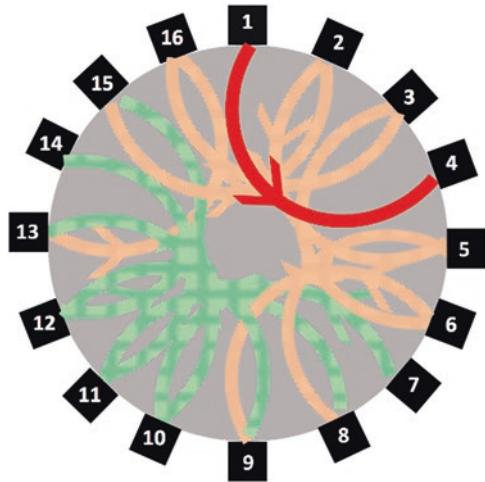
features for the near sub-frame channels and to 96 features when using the far sub-frame channels, as compared to 208 for a full measurement frame. It is anticipated that the near sub-frame channels are more informative due to their proximity to the injecting pairs. The near and far sub-frame channels, for one sub-frame (that of the 1–4 injection pair), are shown in Fig. 12.15. The injecting electrode pair is denoted by the red arrow, with the near sub-frame channels shown in orange, and the far sub-frame channels shown in green.

The performance of the RBF-SVM classifier using the near and far sub-frame channels are again reported at each SNR level as the mean  $\pm$  standard deviation of the sensitivity, specificity, and accuracy. These results are given in Fig. 12.16. Both the near and far sub-frame channels offer perfect performance (sensitivity, specificity, and accuracy of  $1.00 \pm 0.00$ ) at 80 dB SNR, with a slight drop in performance at 60 dB SNR (but all values are  $\geq 0.99 \pm 0.01$ ) before further drops at the 40 dB and 20 dB SNR levels. The near sub-frame channels result in better performance than the far sub-frame channels. Performance at all SNR levels for the near sub-frame channels in particular is equivalent to that of using complete frames despite an almost 50% reduction in dimensionality.

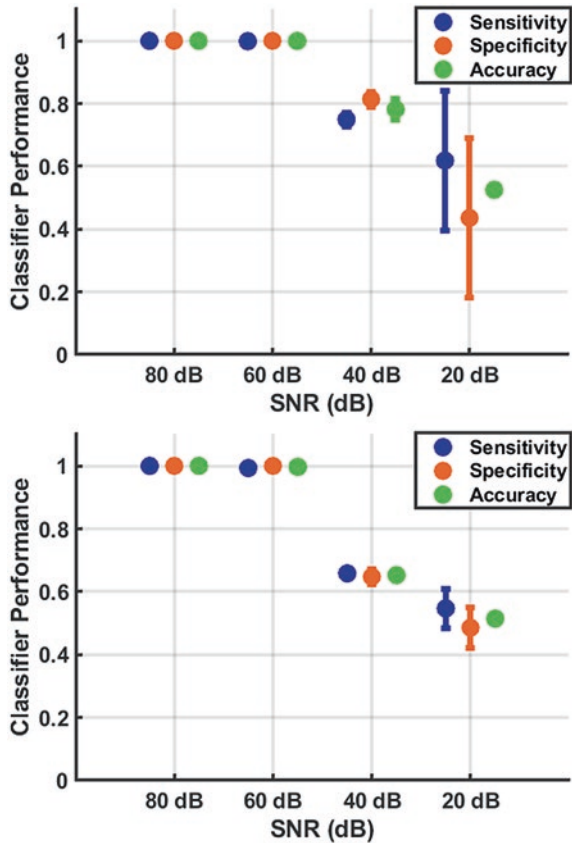
## Laplacian Scores

A type of feature selection method is filter-based methods. Filter methods work by analysing the data before classification, giving a ranking to each feature. Then, the number of ranked features that optimises performance can be chosen by the user. In

**Fig. 12.15** Near and far sub-frame channels. Here the injection pair of 1–4 is shown (red). The 7 nearest channels are shown in orange, with the 6 far channels shown in green. Channels involving the measurement pair are not considered. These near sub-frames and far sub-frames channels are then used as inputs to the classifier



**Fig. 12.16** Comparison of RBF-SVM classifier performance with using the near (top) and far (bottom) sub-frame channels as input features. Both the near and far sub-frame channels results in perfect ( $1.00 \pm 0.00$ ) performance at 80 dB SNR and near perfect ( $\geq 0.99 \pm 0.01$ ) at 60 dB SNR. At lower SNR levels of 40 dB and 20 dB, the near sub-frames outperform the far sub-frames. Of note, the near sub-frames result in equivalent performance to using full measurement frames at all SNR points

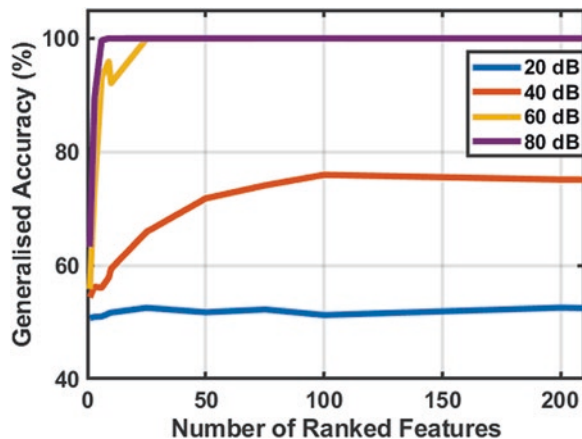


the context of this work, features correspond to the measurement channels. Filter methods can be implemented as either supervised or unsupervised methods. Supervised filter methods require both the observations (inputs) and classes (labels) in order to rank the features. In order to avoid any bias or data contamination, it is important to carefully choose a subset of the entire data set for the feature selection process when using supervised filter methods. Alternatively, unsupervised filter methods can use the entire dataset in order to rank the features, without biasing the classification result. An unsupervised feature selection algorithm, the Laplacian Score algorithm [55, 56], was used in this work to rank the features on the measurement sets (datasets). Specifically, the Laplacian Score algorithm works on the assumption that if two data points are close, then the data points most likely share a label [55]. Further detail on the algorithm can be found in [55]. The distance metric used in this work to define the weight matrix of the algorithm was the Euclidian distance. The advantage of using the filter-based feature selection is that after determination of the optimal number of ranked features, the original data can be used as input for the classification, with only the additional computational cost of removal of unnecessary features.

After first standardising the data, the Laplacian score is applied to each data set corresponding to each of the four SNR levels (80 dB, 60 dB, 40 dB, and 20 dB) to obtain a ranking of the 208 features at each SNR level. The optimal number of ranked features is then chosen through finding the number of features that lead to greatest generalised accuracy in the cross-validation training of the SVM classifier. In Fig. 12.17, the generalised accuracy is presented, at each of the four SNR levels, as the number of Laplacian score ranked features is increased. Based on Fig. 12.17, we can determine the optimal number of features, i.e. the best combination between the number of features and the best generalised accuracy; these optimal points are tabulated in Table 12.1.

The performance of the classifier at each SNR level is assessed with the pre-determined number of ranked features as given in Table 12.1. The results are shown

**Fig. 12.17** The performance of the RBF-SVM classifier using a different number of ranked features, measured by the generalised accuracy. The ranked features were determined using the Laplacian score. The optimal point at a given SNR the one offering the highest accuracy with the lowest number of features



**Table 12.1** The optimum number of ranked features at each SNR level (Maximal accuracy with fewest number of features)

SNR point	Number of ranked features	Generalised accuracy (%)
80 dB	25	100
60 dB	75	100
40 dB	100	75.96
20 dB	208	52.55

in Fig. 12.18. The accuracy, sensitivity, and specificity are perfect ( $1.00 \pm 0.00$ ) at 80 dB SNR, and all are better than  $0.97 \pm 0.01$  at 60 dB SNR. Thus, classification performance is preserved while significantly reducing the input feature size from 208 to 25 and 75 features for the 80 dB and 60 dB SNR levels, respectively. Even at 40 dB SNR, classifier performance was essentially unchanged (compared to using full measurement frames) while reducing the input data set to only 100 features. As with all previous analyses, as the SNR level decreased to 20 dB, classifier performance approaches that of a random guess (metric scores of 0.5).

While unsupervised filter-based feature selection does allow preservation of the captured data to be used as inputs to the classifier in a reduced form, transforming the data with variance techniques such as PCA may enhance the results. The PCA approach is considered next.

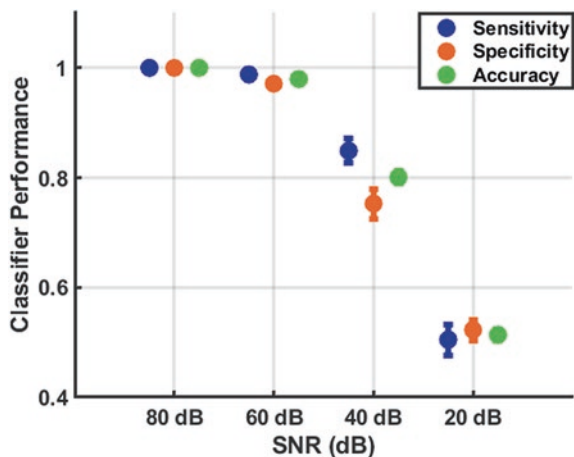
## Principal Component Analysis

A commonly implemented feature extraction method is PCA [24, 25]. PCA is used to reduce the dimensionality of data by generating new variables that represent the original data. These new variables, referred to as the principal components, are created from a linear combination of the original variables, with each successive component defining an orthogonal axis to the previous components. Thus, the entire set of principal components form an orthogonal basis for the space defined by the original data set. The data set can then be projected onto this new orthogonal basis in such a way that the variance in each axis is maximised, allowing data to be, potentially, better discriminated [57], and only a select few principal components can be used to accurately represent the data. Thus, PCA is used to both extract specific features and reduce the dimensionality of the data.

The projection of the original data on specific principal components can be referred to as the “scores”. For every observation, it is these scores that will be used as input features to the RBF-SVM classifier. As PCA is a variance based feature extraction algorithm, it is important to prevent any data contamination; when performing PCA, it is necessary that there is no knowledge of the test data set. In this work, PCA is performed on only the training data, with the transformative coefficients stored and then applied to the test-set data to obtain the projection onto the principal components. Thus, we can ensure that there is no knowledge of the test-set data when performing PCA.

Similar to the previous section, a search for the optimal number of principal components is completed prior to assessing the classifier performance. The opti-



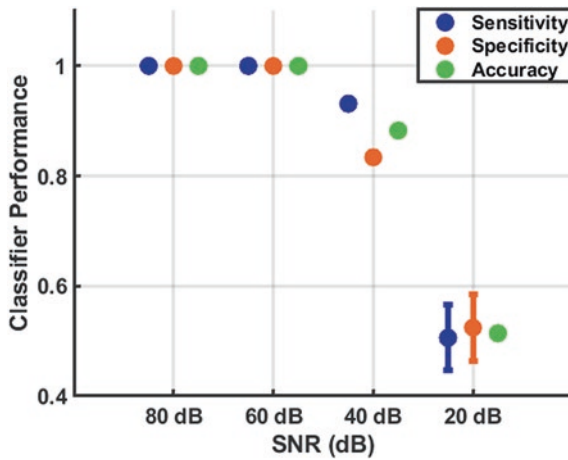
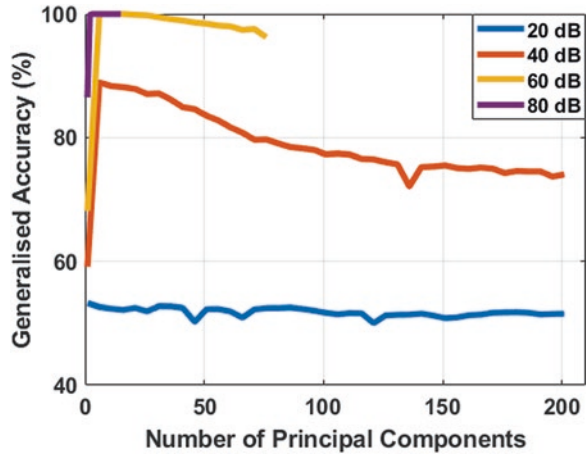


**Fig. 12.18** Performance of the RBF-SVM Classifier at each SNR level using features based on Laplacian scores. The number of ranked features offering maximal accuracy is pre-determined with this feature set (selection of channels) used to train and set the classifier. Perfect performance is achieved at 80 dB SNR with  $1.00 \pm 0.00$  in all metrics, with the use of only 25 features, with accuracy  $>0.97$  at 60 dB SNR using 75 features

mal number of principal components is found by finding the best generalised accuracy, for each of the four SNR levels, across the cross-validation training. In Fig. 12.19, a comparison of the generalised accuracy and the number of principal components, for each of the four SNR levels, is shown. From this graph it becomes clear that for each SNR level, there is a range of principal components when performance is maximised prior to a decrease of performance as more principal components are added. This is explained by the fact that each successive principal component explains less and less variance of the original data. Therefore, those final components are simply expressing the noise in the data set, with no meaningful information contained. The optimal number of components chosen for the 80 dB, 60 dB, 40 dB, and 20 dB SNR levels is 10, 10, 11, and 31 principal components, respectively.

The classifier performance is then assessed by projecting the test data set onto the principal components using the stored projection coefficients found in training. In Fig. 12.20, the performance of the classifier is compared at all four of the SNR levels. The use of PCA leads to a marked improvement in comparison to using the entire raw data set (complete measurement frames), while also significantly reducing the input data set to at most 31 features. Most notably, at 40 dB SNR, there is an increase of almost 10% in the mean accuracy compared to using the complete measurement frames, while decreasing the input feature size from 208 features to only 11 features. Also, significantly at 60 dB SNR perfect performance is achieved using only 10 components. However, as in all previous analyses, the classifier is no better than random guesswork at 20 dB SNR.

**Fig. 12.19** A comparison of the generalised accuracy at the four SNR levels as the number of principal components is increased. The optimal number of principal components at each SNR level is that number giving the highest generalised accuracy which is 10, 10, 11, and 31 principal components for 80 dB, 60 dB, 40 dB, and 20 dB SNR levels, respectively



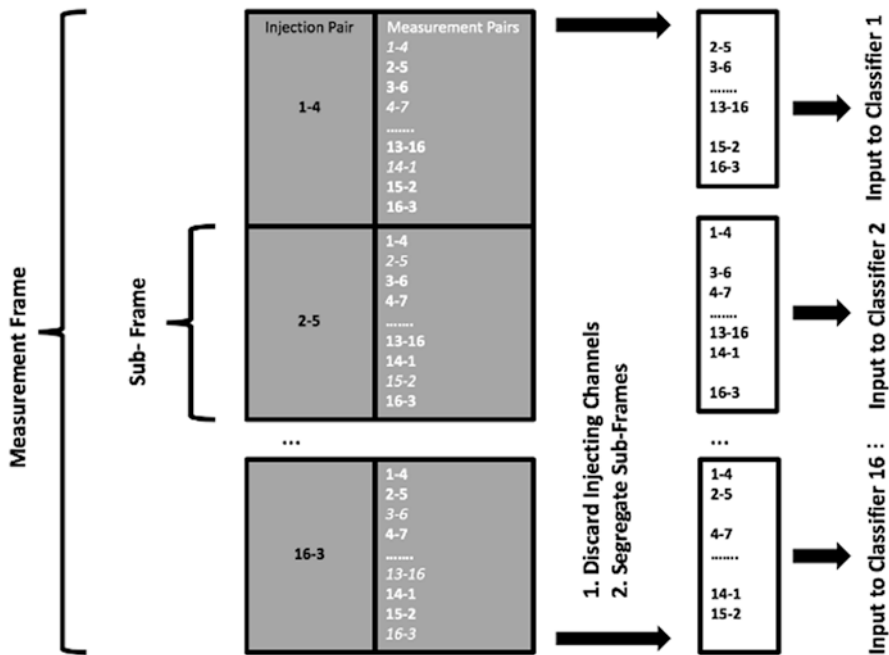
**Fig. 12.20** Comparison of the performance at each of the four SNR levels for the classifier after performing PCA. Perfect performance ( $1.00 \pm 0.00$  in all metrics) is given at the 80 dB and 60 dB SNR levels despite using only 10 components at each point. A near 0.1 (10%) improvement in accuracy is seen at the 40 dB SNR level compared to the full measurement frames, but performance is approximately 0.5 in all metrics at 20 dB SNR, essentially representing a random classifier

### 12.4.3 Ensemble Classifier

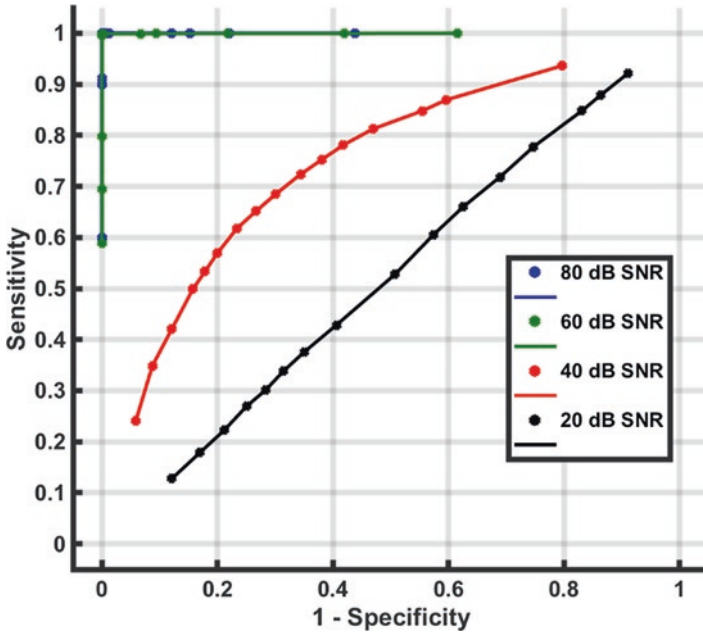
An ensemble classifier aims to make use of multiple classifiers to make an informed decision. Additionally, these classifiers allow for better control of the sensitivity and specificity of the classifier performance [58]. In this work, an ensemble classifier was created by assigning a classifier to each of the 16 sub-frames for a given complete measurement frame. A voting scheme from each of the 16 classifiers was then used for the final classification decision. The design and implementation of this ensemble classifier is shown in Fig. 12.21.

For each observation, each of the 16 classifiers separately classified the case as  $\pm 1$  (bleed or normal). Next, the sensitivity, specificity, and accuracy of the ensemble classifier at different threshold points were calculated. A threshold was the minimum number of separate classifiers needed to classify a case as a bleed for it to be classified as such; if the number was below this threshold, then the case was classified as not bleed. The threshold was adjusted from 1 to 16 in steps of 1. This control on the sensitivity and specificity allowed for the generation of a ROC curve. In Fig. 12.22, a comparison of the ROC curve, at each of the four SNR levels, for the ensemble classifier is shown.

For a low threshold (for example 1), the general trend is that the FP (1 – Specificity) rate will be high as the ensemble classifier is very sensitive to bleeds. This translates as a high sensitivity at a cost to specificity if the system is not robust. At a high threshold (for example 16), sensitivity is lost but specificity is maximised as the FN is high, with more classifiers needing to agree on labelling a case as a bleed before it is classified as a bleed. The accuracy will lie in between these two values of specificity and sensitivity at all threshold points. The trade-off in sensitivity and specificity is best illustrated at the lower SNR levels of 40 dB and 20 dB. For the higher SNR values of 80 dB and 60 dB, there is a threshold (or set of thresholds)



**Fig. 12.21** Example of the design and implementation of the ensemble classifier. The measurement frame for a given case can be divided into sub-frames with the channels from each sub-frame used as the input for a separate classifier. The complete set of frames are segregated in this way with 16 classifiers trained and tested. Each classifier separately labels a case as  $\pm 1$  with the aggregate result calculated according to a threshold which can be adjusted



**Fig. 12.22** ROC curves for the ensemble classifier at each SNR level. The points on each curve correspond to each discrete threshold value, between 1 and 16 (from right to left), with the corresponding line interpolated between the points. The curves illustrate the trade-offs between sensitivity and specificity possible at each SNR level by changing the operating points. Both the 80 dB and 60 dB plots offer an operating point of perfect performance (0,1). Performance is reduced at 40 dB and is worst at 20 dB SNR as expected. The 20 dB line is approximately that of a random classifier, being a diagonal line passing through the points (0,0) and [1]

in the intermediate area where sensitivity, specificity, and accuracy all are  $1 \pm 0$ . For both the 80 dB and 60 dB SNR levels, this area is centred at a threshold of 10. The ROC curve allows the user to select the operating point offering optimal performance, which for the proposed application of bleed detection is maximal sensitivity as justified in Sect. 12.2.2. As shown in Fig. 12.22, the 80 dB and 60 dB SNR levels result in an operating point offering the perfect combination of sensitivity and specificity both equal to 1. At 40 dB SNR, for example, a maximal sensitivity of just over 0.9 is achieved with a reduction in specificity to 0.2, with a worse performance given at 20 dB SNR, which has the performance of a random classifier.

## 12.5 Discussion and Conclusions

This chapter illustrates the important role that computational modelling tools have in exploring both the feasibility and the challenges in developing technologies that tackle important medical problems such as brain bleed detection. Brain haemorrhages are a medical emergency that require a prompt and accurate diagnosis prior

to any appropriate treatment being administered. An ideal technological solution would be portable, non-invasive, cost effective, and crucially feature a sensitivity to the presence of a bleed (with ideally simultaneous high specificity) in the brain. Such a technology may be found in EIT coupled with modern machine learning algorithms. This work examined the feasibility of EIT coupled with ML to develop a bleed/ normal classifier based on EIT measurement frames. The approach removes the image reconstruction steps that are challenging to EIT. Further, it is EIT applied to a static scene where the most successful EIT modality, time difference EIT, cannot be applied. The chapter builds on the material presented in earlier works, including [14] and particularly [15] where, to our knowledge, such an approach with a static scene was investigated for the first time.

The effect of individual variables on performance such as the effect of noise in measurement frames, bleed location, bleed size, electrode positioning, and variations in anatomy was initially summarised in Sect. 12.3. The conclusions drawn from this section are: good performance (sensitivity, specificity, and accuracy at or near 1) is achievable particularly at 80 dB SNR; the technique is sensitive to new bleed locations not seen in the training data (although the simple pre-processing step of sorting the measurement values can improve this); the technique robustly detects bleeds larger than those trained on, but struggles with those smaller; the technique is robust to small changes in electrode positioning; and the technique struggles with unseen anatomies, in this case modelled as deviations in the morphology of the head and brain FEM models.

The simple replacement of the linear kernel with a Gaussian RBF kernel resulted in improved performance. Although both resulted in perfect sensitivity, specificity, and accuracy of  $1 \pm 0$  at 80 dB SNR, the benefit of the RBF kernel is seen at 60 dB and 40 dB SNR levels with an increase in the mean accuracy between approximately 3% and 9%, respectively. This significant improvement in classifier performance highlights the need to explore options related to classifier choice and also the input feature selection process.

The final part of this work examined methods that moved the nature of the classifier input away from raw or minimally processed measurement frames with a view to increasing computational efficiency through intelligent feature selection that reduced dimensionality. Approaches used included processing of the measurement frames to create sub-frame means, near and far sub-frame channels, using Laplacian scores and PCA to extract specific features, and examining an ensemble classifier with thresholding to control the sensitivity to bleeds. A summary of the performance of these different classifiers, at the 60 dB and 40 dB SNR levels, where performance was mostly impacted, is shown in Tables 12.2 and 12.3, respectively. For all classifiers, the 80 dB SNR level yielded perfect classification results, whereas at 20 dB SNR all classifiers performed at essentially a guess level.

Each of the methods described in Sect. 12.4 significantly reduced the dimensionality of the input data to the classifier. The sub-frame means approach reduced the input data size to only 16 features, however suffered from poor performance when the SNR levels dropped below 80 dB, with a decrease in the mean accuracy of almost 25% in comparison to using all 208 features even at 60 dB SNR.

**Table 12.2** Summary of different classifier performance at 60 dB SNR (all metrics reported as the mean  $\pm$  standard deviation of the sensitivity (Sens.), specificity (Spec.), and accuracy (Acc.) with a perfect score being  $1.00 \pm 0.00$ )

		Sens.	Spec.	Acc.
Classifier type	Lin.	$0.95 \pm 0.01$	$0.96 \pm 0.01$	$0.95 \pm 0.01$
	RBF	$0.99 \pm 0.00$	$0.97 \pm 0.00$	$0.98 \pm 0.00$
	Mean	$0.71 \pm 0.02$	$0.82 \pm 0.03$	$0.76 \pm 0.01$
	Near	$1.00 \pm 0.00$	$1.00 \pm 0.00$	$1.00 \pm 0.00$
	Far	$0.99 \pm 0.01$	$1.00 \pm 0.00$	$1.00 \pm 0.00$
	Laplac.	$0.99 \pm 0.01$	$0.97 \pm 0.01$	$0.98 \pm 0.00$
	PCA	$1.00 \pm 0.00$	$1.00 \pm 0.00$	$1.00 \pm 0.00$
	Ensemb.	$0.99 \pm 0.00$	$1.00 \pm 0.00$	$1.00 \pm 0.00$

All classifiers used RBF kernel except when labelled 'Linear'. Linear (Lin.): Linear kernel with full measurement frames as the classifier input; RBF: RBF kernel with full measurement frames as the classifier input; Mean: Sub-frame means as classifier input; Near: Near sub-frame channels as input; Far: Far sub-frame channels as classifier input; Laplacian (Laplac.): Optimal number of ranked features as determined by Laplacian filtering used as classifier input; PCA: Optimal number of principal components used as classifier input; Ensemble (Ensemb.): Results correspond to the threshold offering maximal sensitivity

**Table 12.3** Summary of different classifier performance at 40 dB SNR (all metrics reported as the mean  $\pm$  standard deviation of the sensitivity (Sens.), specificity (Spec.), and accuracy (Acc) with a perfect score being  $1.00 \pm 0.00$ )

		Sens.	Spec.	Acc.
Classifier type	Lin.	$0.71 \pm 0.02$	$0.70 \pm 0.02$	$0.70 \pm 0.01$
	RBF	$0.82 \pm 0.02$	$0.76 \pm 0.02$	$0.79 \pm 0.00$
	Mean	$0.56 \pm 0.04$	$0.54 \pm 0.04$	$0.55 \pm 0.03$
	Near	$0.75 \pm 0.02$	$0.81 \pm 0.02$	$0.78 \pm 0.03$
	Far	$0.66 \pm 0.01$	$0.65 \pm 0.02$	$0.65 \pm 0.01$
	Laplac.	$0.85 \pm 0.02$	$0.75 \pm 0.03$	$0.80 \pm 0.01$
	PCA	$0.93 \pm 0.00$	$0.83 \pm 0.01$	$0.88 \pm 0.00$
	Ensemb.	$0.61 \pm 0.04$	$0.77 \pm 0.05$	$0.69 \pm 0.01$

All classifiers used RBF kernel except Linear. Abbreviations of the classifier type are consistent with Table 12.3

The near and far sub-frame channels gave an approximate 50% reduction in dimensionality. Using the near sub-frame channels preserved the classifier performance when in comparison to the full data set, whereas the far channels led to a reduction in the mean accuracy of almost 15% at 40 dB SNR. These results imply, as was hypothesised, that the near sub-frame channels are more important for classifier performance.

Using the Laplacian scores to rank and choose features led to similar classifier performance using all 208 features at all SNR levels. However, at 80, 60, and 40 dB, the input features were reduced to only 25, 75, and 100 features respectively.

The use of PCA to extract and select features, in combination with the RBF-SVM classifier, lead to the best overall results, with mean accuracy values of 100%

and 88.26% at the 60 dB and 40 dB SNR levels. This marks a 1.25% and 8.91% improvement over using all 208 features, while only needing the first 10 and 11 components, at 60 dB and 40 dB SNR, respectively.

The ensemble classifier approach offered a trade-off between sensitivity and specificity depending on the threshold used. At 80 dB and 60 dB, a wide region centred around a threshold of 10 offered perfect sensitivity, specificity, and accuracy. However, this method fails to match the performance of using all the input features at 40 dB.

This work has demonstrated promise in the approach of using EIT measurement frames coupled with ML for bleed detection. Careful consideration and experimentation in regard to measurement frame processing, choice of ML algorithm, and parameters can significantly improve performance. These areas alone merit further study as well as the testing with a more realistic multi-layered computational model and physical phantom. Encouragingly, EIT hardware with SNR levels at or near 80 dB exist, which adds to the hope that computational results can be translated into real world models [59]. EIT is already a valuable imaging tool in time changing scenes but has the potential to be a valuable modality in cases with static pathologies such as brain bleeds with innovative methods such as those presented in this set of studies. We encourage researchers to further build on and develop these ideas and paradigms in order to make a measurable impact in tackling important medical problems and improving patient outcomes.

**Acknowledgements** The research leading to these results has received funding from the European Research Council under the European Union's Horizon 2020 Programme/ERC Grant Agreement BioElecPro n.637780, Science Foundation Ireland (SFI) grant number 15/ERC/S/3276, the Hardiman Research Scholarship from NUIG, the charity RESPECT, the Irish Research Council GOIPD/2017/854 fund, and the People Programme (Marie Curie Action) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA Grant Agreement no. PCOFUND-GA-2013-608728.

## References

1. Velayudhan, V. Stroke imaging: Overview, computed tomography, magnetic resonance imaging [Internet]. Medscape. [cited 2016 Oct 19]. Available from: <http://emedicine.medscape.com/article/338385-overview>
2. Ovbiagele, B., & Nguyen-Huynh, M. N. (2011). Stroke epidemiology: Advancing our understanding of disease mechanism and therapy. *Neurotherapeutics*, 8(3), 319–329.
3. Donnan, G. A., Fisher, M., Macleod, M., & Davis, S. M. (2008). Stroke. *The Lancet*, 371(9624), 1612–1623.
4. Birenbaum, D., Bancroft, L. W., & Felsberg, G. J. (2011). Imaging in acute stroke. *The Western Journal of Emergency Medicine*, 12(1), 67–76.
5. Faul, M., & Coronado, V. (2015). Epidemiology of traumatic brain injury. In *Handbook of clinical neurology [Internet]* (pp. 3–13). Elsevier. [cited 2018 Sep 19]. Available from: <http://linkinghub.elsevier.com/retrieve/pii/B9780444528926000015>.
6. NICE (National Institute for Health and Care Excellence). Head injury overview [Internet]. nice.org.uk. [cited 2016 Oct 19]. Available from: <https://pathways.nice.org.uk/pathways/head-injury>

7. Kim, J. J., & Gean, A. D. (2011). Imaging for the diagnosis and management of traumatic brain injury. *Neurotherapeutics*, 8(1), 39–53.
8. Lee, B., & Newberg, A. (2005). Neuroimaging in traumatic brain imaging. *NeuroRx*, 2(2), 372–383.
9. Shen, D., Zhang, D., Young, A., & Parvin, B. (2015). Editorial: Machine learning and data mining in medical imaging. *IEEE Journal of Biomedical and Health Informatics*, 19(5), 1587–1588.
10. Giger, M. L. (2018). Machine learning in medical imaging. *Journal of the American College of Radiology*, 15(3), 512–520.
11. Brattain, L. J., Telfer, B. A., Dhyani, M., Grajo, J. R., & Samir, A. E. (2018). Machine learning for medical ultrasound: Status, methods, and future opportunities. *Abdominal Radiology (NY)*, 43(4), 786–799.
12. Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19(1), 221–248.
13. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., et al. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke Vascular Neurology*, 2(4), 230–243.
14. McDermott, B., O'Halloran, M., Porter, E., & Santorelli, A. (2018). Brain haemorrhage detection through SVM classification of impedance measurements. In *2018 40th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Honolulu, Hawaii, United States: IEEE.
15. McDermott, B., O'Halloran, M., Porter, E., & Santorelli, A. (2018). Brain haemorrhage detection using a SVM classifier with electrical impedance tomography measurement frames. Stoean R, editor. *PLoS One*, 13(7), e0200469.
16. Brown, B. (2003). Electrical impedance tomography (EIT): A review. *Journal of Medical Engineering & Technology*, 27(3), 97–108.
17. Alberts, B. (Ed.). (2002). *Molecular biology of the cell* (4th ed.). New York: Garland Science. 1548 p.
18. Holder, D., & Institute of Physics (Great Britain) (Eds.). (2005). *Electrical impedance tomography: methods, history, and applications*. Bristol/Philadelphia: Institute of Physics Pub. 456 p. (Series in medical physics and biomedical engineering).
19. Adler, A., & Boyle, A. (2017). Electrical impedance tomography: Tissue properties to image measures. *IEEE Transactions on Biomedical Engineering*, 64(11), 2494–2504.
20. Adler, A., Grychtol, B., & Bayford, R. (2015). Why is EIT so hard, and what are we doing about it? *Physiological Measurement*, 36(6), 1067–1073.
21. Horesh, L., Gilad, O., Romsauerova, A., Arridge, S., & Holder, D. (2005). Stroke type differentiation by multi-frequency electrical impedance tomography – a feasibility study. In *Proc IFMBE* (pp. 1252–1256).
22. Dowrick, T., Blochet, C., & Holder, D. (2015). *In vivo* bioimpedance measurement of healthy and ischaemic rat brain: Implications for stroke imaging using electrical impedance tomography. *Physiological Measurement*, 36(6), 1273–1282.
23. Mitchell, T. M. (1997). *Machine learning*. New York: McGraw-Hill. 414 p. (McGraw-Hill series in computer science).
24. Santorelli, A., Porter, E., Kirshin, E., Liu, Y. J., & Popovic, M. (2014). Investigation of classifiers for tumour detection with an experimental time-domain breast screening system. *Progress In Electromagnetics Research*, 144, 45–57.
25. Conceicao, R. C., O'Halloran, M., Glavin, M., & Jones, E. (2010). Support vector machines for the classification of early-stage breast cancer based on radar target signatures. *Progress In Electromagnetics Research B*, 23, 311–327.
26. Oliveira, B., Godinho, D., O'Halloran, M., Glavin, M., Jones, E., & Conceição, R. (2018). Diagnosing Breast Cancer with Microwave Technology: Remaining challenges and potential solutions with machine learning. *Diagnostics (Basel)*, 8(2), 36.
27. Golnaraghi, F., & Grewal, P. K. (2014). Pilot study: Electrical impedance based tissue classification using support vector machine classifier. *IET Science, Measurement and Technology*, 8(6), 579–587.



28. Gur, D., Zheng, B., Lederman, D., Dhurjaty, S., Sumkin, J., Zuley, M. (2010). A support vector machine designed to identify breasts at high risk using multi-probe generated REIS signals: A preliminary assessment. In: Manning DJ, Abbey CK, editors. [cited 2018 Jan 18]. p. 76271B. Available from: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.844452>.
29. Laufer, S., & Rubinsky, B. (2009). Tissue characterization with an electrical spectroscopy SVM classifier. *IEEE Transactions on Biomedical Engineering*, 56(2), 525–528.
30. Shini, M. A., Laufer, S., & Rubinsky, B. (2011). SVM for prostate cancer using electrical impedance measurements. *Physiological Measurement*, 32(9), 1373–1387.
31. Schlebusch, T., Nienke, S., Leonhardt, S., & Walter, M. (2014). Bladder volume estimation from electrical impedance tomography. *Physiological Measurement*, 35(9), 1813–1823.
32. Dunne, E., Santorelli, A., McGinley, B., Leader, G., O’Halloran, M., & Porter, E. (2018). Supervised learning classifiers for electrical impedance-based bladder state detection. *Scientific Reports*, 8(1), 5363.
33. McDermott, B., O’Halloran, M., Santorelli, A., McGinley, B., & Porter, E. (2018). Classification applied to brain haemorrhage detection: Initial phantom studies using electrical impedance measurements. In *Proceeding of the 19th international conference on biomedical applications of electrical impedance tomography*. Edinburgh.
34. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
35. Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines: And other kernel-based learning methods*. Cambridge; New York: Cambridge University Press. 189 p.
36. Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory – COLT ’92 [Internet]* (pp. 144–152). Pittsburgh: ACM Press. [cited 2018 Oct 4]. Available from: <http://portal.acm.org/citation.cfm?doi=130385.130401>.
37. Standing, S., Ananad, N., & Gray, H. (Eds.). (2016). *Gray’s anatomy: The anatomical basis of clinical practice* (41st ed.). Philadelphia: Elsevier. 1562 p.
38. Zhang, J., Yang, B., Li, H., Fu, F., Shi, X., Dong, X., et al. (2017). A novel 3D-printed head phantom with anatomically realistic geometry and continuously varying skull resistivity distribution for electrical impedance tomography. *Scientific Reports [Internet]*, 7(1). Available from: <http://www.nature.com/articles/s41598-017-05006-8>.
39. Avery, J., Aristovich, K., Low, B., & Holder, D. (2017). Reproducible 3D printed head tanks for electrical impedance tomography with realistic shape and conductivity distribution. *Physiological Measurement*, 38(6), 1116–1131.
40. Grozny. Thingiverse – Human Head [Internet]. [cited 2017 Feb 15]. Available from: <http://www.thingiverse.com/thing:172348>
41. Dilmen, N. NIH 3D print exchange- brain MRI [Internet]. [cited 2017 Feb 15]. Available from: <https://3dprint.nih.gov/discover/3DPX-002739>
42. Adler, A., & Lionheart, W. R. B. (2006). Uses and abuses of EIDORS: An extensible software base for EIT. *Physiological Measurement*, 27(5), S25–S42.
43. Schoeberl, J. Netgen [Internet]. Vienna: Vienna University of Technology; Available from: <https://ngsolve.org/>
44. Geuzaine, C., & Remacle, J.-F. (2009). Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11), 1309–1331.
45. MATLAB 2017A [Internet]. Natick: The MathWorks Inc. Available from: <https://uk.mathworks.com/>
46. Eaton, J.W., Bateman, D., Hauberg, S., Wehbring, R. GNU Octave version 4.2.2 manual: A high-level interactive language for numerical computations [Internet]. 2018. Available from: <https://www.gnu.org/software/octave/doc/v4.2.2/>
47. Grychtol, B., Adler, A. FEM electrode refinement for electrical impedance tomography. In: 2013 35th annual international conference of the IEEE Engineering in Medicine and Biology

- Society (EMBC) [Internet]. IEEE; 2013. p. 6429–6432. Available from: <http://ieeexplore.ieee.org/document/6611026/>
48. Autodesk. Fusion 360 [Internet]. Mill Valley: Autodesk. Available from: <https://www.autodesk.com/products/fusion-360>
  49. Broderick, J. P., Brott, T. G., Duldner, J. E., Tomsick, T., & Huster, G. (1993). Volume of intracerebral hemorrhage. A powerful and easy-to-use predictor of 30-day mortality. *Stroke*, 24(7), 987–993.
  50. Hemphill, J. C., Bonovich, D. C., Besmertis, L., Manley, G. T., Johnston, S. C., & Tuhim, S. (2001). The ICH score: A simple, reliable grading scale for intracerebral hemorrhage editorial comment: A simple, reliable grading scale for intracerebral hemorrhage. *Stroke*, 32(4), 891–897.
  51. Hun Wi, Sohal, H., McEwan, A. L., Eung Je Woo, & Tong In Oh. (2014). Multi-frequency electrical impedance tomography system with automatic self-calibration for long-term monitoring. *IEEE Transactions on Biomedical Circuits and Systems*, 8(1), 119–128.
  52. Jehl, M., Avery, J., Malone, E., Holder, D., & Betcke, T. (2015). Correcting electrode modelling errors in EIT on realistic 3D head models. *Physiological Measurement*, 36(12), 2423–2442.
  53. Li, Y., Santorelli, A., Laforest, O., & Coates, M. (2015). Cost-sensitive ensemble classifiers for microwave breast cancer detection. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) [Internet]* (pp. 952–956). South Brisbane: IEEE. [cited 2018 Oct 9]. Available from: <http://ieeexplore.ieee.org/document/7178110/>.
  54. Hsu, C.-W., Chang, C.-C., Lin, C.-J. (2010). A practical guide to support vector classification. p. 16.
  55. He, X., Cai, D., & Niyogi, P. (2005). Laplacian Score for Feature Selection. In *NIPS'05 Proceedings of the 18th International Conference Neural Information Process System* (pp. 507–514). Vancouver.
  56. Dunne, E., Santorelli, A., McGinley, B., Leader, G., O'Halloran, M., & Porter, E. (2018). Image-based classification of bladder state using electrical impedance tomography. *Physiological Measurement*, 39(12), 124001
  57. Conceição, R. C., O'Halloran, M., Glavin, M., & Jones, E. (2011). Evaluation of features and classifiers for classification of early-stage breast cancer. *Journal of Electromagnetic Waves and Applications*, 25(1), 1–14.
  58. Li, Y., Porter, E., Santorelli, A., Popović, M., & Coates, M. (2017). Microwave breast cancer detection via cost-sensitive ensemble classifiers: Phantom and patient investigation. *Biomedical Signal Processing and Control*, 31, 366–376.
  59. Avery, J., Dowrick, T., Faulkner, M., Goren, N., & Holder, D. (2017). A versatile and reproducible multi-frequency electrical impedance tomography system. *Sensors*, 17(2), 280–280.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

