Image-based Classification of Bladder State using Electrical Impedance Tomography

Eoghan Dunne[1,2]**, Adam Santorelli**[1,2]**, Brian McGinley**[1,3]**, Geraldine Leader**[4]**, Martin O'Halloran**[1,2] **and Emily Porter**[1,2]

[1]Translational Medical Device Lab, National University of Ireland Galway, Galway City, Ireland

[2]Department of Electrical & Electronic Engineering, College of Engineering & Informatics, National University of Ireland Galway, Galway City, Ireland

[3]Department of Computer Science & Applied Physics at Galway-Mayo Institute of Technology, Galway City, Ireland

[4] School of Psychology, National University of Ireland Galway, Galway City, Ireland

E-mail: e.dunne13@nuigalway.ie

Abstract.

Objective: In this study, we examine the potential of using machine learning classification to determine the bladder state ('not full', 'full') with electrical impedance tomography (EIT) images of the pelvic region. Accurate classification of these states would enable urinary incontinence (UI) monitoring to alert the patient, before involuntary voiding occurs, in a low-cost and discrete manner. *Approach:* Using both numerical and experimental data, we form datasets that contain diverse observations with varying clinical parameters such as bladder volume, urine conductivity, and the reference used for time-difference imaging. We then classify the bladder state using both pixel-wise and feature extraction-based classification techniques. We employ principal component analysis, wavelets, and image segmentation to help create features. *Main results*: The performance was compared across several classifier algorithms. The minimum accuracy was 77.50%. The highest accuracy observed was 100%, and was found by combining principal component analysis and the Gaussian radial based function kernel support vector machine. This combination also offered the best trade-off between classification performance and the costs of training time and memory space. The biggest challenge in bladder state classification is classifying volumes near the separation volume of not full and full, in which choosing the most suitable classifier combination can minimize this error. *Significance:* We performed the first machine learning classification of bladder EIT images, achieving high classification accuracies with both numerical and experimental data. This work highlights the potential of using image-based machine learning with an EIT device to support bladder monitoring for those suffering from UI.

Keywords: Urinary Incontinence, Electrical Impedance Tomography, Machine Learning, Image Classification, Feature Extraction

1. Introduction

Urinary incontinence is a condition characterized by the involuntary discharge of urine (Abrams *et al* 2010). This condition affects more than 200 million people across the world (Norton and Brubaker 2006). Urinary incontinence is particularly prevalent in women (Yip and Cardozo 2007), the elderly (Shamliyan *et al* 2009), children and adolescents with intellectual disabilities (Lottmann and Alova 2007; Gor, Fuhrer and Schober 2012), and in individuals with urinary dysfunction, diabetes or neurological conditions such as spinal cord injury and multiple sclerosis (Shamliyan *et al* 2009). Urinary incontinence can significantly and negatively impact an individual's quality of life, both physically and psychologically (Coyne *et al* 2008). Notably, in the elderly, there is an increased risk of morbidity due to falls with the condition (Gibson *et al* 2018). Overall, urinary incontinence places a considerable burden on healthcare systems (Morrison and Levy 2006).

Non-invasive methods, such as ultrasound imaging, aid in monitoring the bladder volume, primarily in clinical settings (Kelly 2004). These devices produce images in either 2D or 3D in order to determine the bladder volume (Jalbani and Ather 2014). However, ultrasound devices require medical professionals to regularly check the bladder volume with the ultrasound probe, burdening health systems (Niestoruk *et al* 2012). Wearable ultrasound devices have been investigated to continually monitor the bladder in individuals with urinary incontinence (Kristiansen, Djurhuus and Nygaard 2004; Niestoruk *et al* 2012), yet remain bulky (Cao *et al* 2011) and costly (Niu *et al* 2011). Therefore, there is a need for a tool to monitor urinary incontinence in a discrete and low-cost way.

Electrical impedance tomography (EIT) is a non-invasive, portable, and low-cost method that can be used to monitor bladder volume through interpretation of reconstructed images of the pelvic region (Leonhardt *et al* 2011). This technology involves the injection of small alternating currents into the pelvic region and the measuring of the resultant voltages. Then, the resultant voltages are used to reconstruct the conductivity distribution within the pelvis and produce an image. The reconstructed image can then be used to map the bladder shape and pixel intensities to the bladder volume (Schlebusch *et al* 2014a).

Previous studies using EIT have shown the viability of the technology for monitoring bladder volume in clinical settings (Leonhardt *et al* 2011; Li *et al* 2016). Further, more recent studies have suggested EIT as alternative to ultrasound, particularly for determining residual urine (Leonhäuser *et al* 2018). Image-based assessments, such as correlating the global impedance of the image and the area of the bladder in the image with the bladder volume, have also been investigated with simulation data (Schlebusch *et al* 2014a). However, for urinary incontinence, it is far more important to alert the incontinent individual to go to the toilet in time than it is to report the absolute bladder volume. Thus, in order to address this need, this work focuses on classifying 'not full' and 'full' bladder states. We have previously used machine learning to classify the bladder state using the raw voltage data (i.e. without reconstructing an image), which showed very promising results (Dunne *et al* 2018a). However, image reconstruction has a potential advantage over the raw voltage data, in that images also provide spatial information, which could further enhance classification of the bladder state. For example, as the bladder fills, the bladder ascends and widens within the pelvic region, and thus images may help capture this spatial information. Therefore, in this work, we perform the first machine learning classification of EIT images for bladder monitoring. We classify the bladder state directly using the pixel intensity values, and by using feature extraction (FE) and selection techniques, on images from both simulated and experiment data. We investigate the effects of bladder volume, urine conductivity, and the type of reference used in image reconstruction, on the classification outcomes. We also compare pixel-wise and FE-based methods and highlight the advantages and drawbacks of each. Furthermore, while the target population in this study is children with intellectual disabilities who are struggling with urinary incontinence and toilet training, the methodology and results are broadly applicable to all populations with urinary incontinence.

This paper is structured as follows: we discuss the simulation and experimental data collection in section 2; the types of classification performed in sections 3 and 4; the classifiers and the architecture used in section 5; the feature selection in section 6; the results in section 7; and we conclude in section 8.

2. Data Collection & Preparation

In this section, we describe the design of the simulation and experimental scenarios, the data obtained from each experiment, and the data pre-processing methods prior to the implementation of any machine learning algorithms.

2.1. Simulation Dataset Collection & Preparation

Using the open-sourced Electrical Impedance Tomography and Diffuse Optical Tomography Reconstruction Software (EIDORS) (Adler *et al* 2006), we constructed a finite element model (FEM) of the human pelvis. The pelvic FEM is an elliptical cylinder, similar to in previous bladder classification studies (Schlebusch *et al* 2014a; Dunne *et al* 2018a). The FEM is proportionally-accurate to the human computed tomography (CT) images from the Visible Human Project (Ackerman 1998), and the experimental phantom that is discussed later in subsection 2.2. The model matches the average size of a 5-year old male child (Dunne *et al* 2018b).

This cylindrical FEM has been previously implemented to compare machine learning algorithms with simulation data in this field (Schlebusch *et al* 2014a; Dunne *et al* 2018a). The electrode shape and locations were modelled to match those of our experimental setup (Swisstom 44 cm child belt, swisstom.com), as shown in figure 1. Ellipsoidal FEMs represent the bladder in the pelvis, with varying diameters based on the volume of urine being represented (40-420 ml). The diameters of the bladder ellipsoids were chosen to match the dimensions of actual human bladders (Hirahara *et al* 2006; Dunne, *et al* 2018b) and scaled to child-size. Each bladder has the same bottom location within the pelvic FEM. As the bladder volume increases, the centroid of the bladder ellipsoid increases in height, mimicking the process of the human bladder ascending in the pelvic cavity towards the abdomen.

In bladder monitoring, the main biological factors that may vary are the bladder volume and the urine conductivity. Both of these factors are variable in time, and intra- and inter-individually (Schlebusch *et al* 2014b). For example, the urine conductivity can range from between 0.59-3.22 S/m (Schlebusch *et al* 2014b). Time-difference imaging is the primary technique used in bladder EIT (Leonhardt *et al* 2011; Dunne *et al* 2018b) and it requires reference data in order to reconstruct an image. The reference data may be data from an empty bladder or from a different bladder volume than the target (i.e., full) bladder volume. Generally, it is preferable to have as large a difference between the target volume and the reference volume as possible, as this will lead to the biggest time-difference, and thus reconstructed images with higher contrast. Here, we use the bladder volume, urine conductivity, and reference data as the factors to test the ability of machine learning in identifying the bladder state with a low signal-to-noise ratio (SNR).
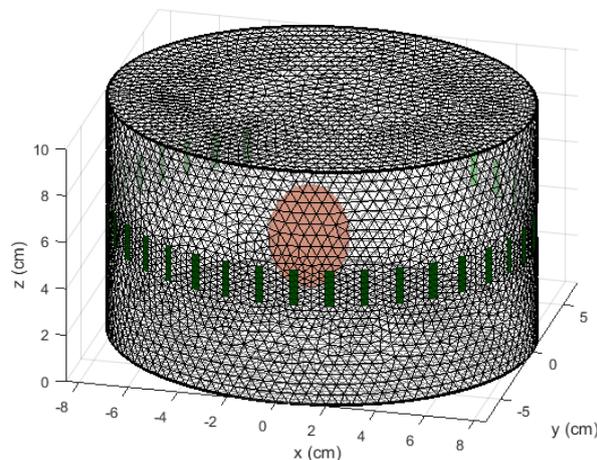

Figure 1: The FEM for simulation data generation, with 32 electrodes. The FEM is an elliptical cylinder and is proportionally-accurate to a child-sized human pelvis. The FEM has a background conductivity of 0.2 S/m, determined from a weighted average of the conductivities of the tissues in the pelvic cavity. The calculation is further discussed in (Dunne et al 2018b). The FEM contains the 240 ml bladder modelled by an ellipsoid, with dimensions based on human data as described in our previous work (Dunne *et al* 2018b). The bottom position of all bladders of different volumes is fixed to the same location, while the centroid ascends up through the cylinder. This positioning mimics how the human bladder ascends upwards within the pelvis toward the abdomen as the bladder fills.

The simulated parameters for each factor are given in table 1, and are based on current literature (Schlebusch *et al* 2014b, Kelly 2004). Each simulated scenario consisted of one bladder volume, one urine conductivity, and one particular reference (with one specific reference bladder volume and one reference urine conductivity). An SNR of 40 dB was used for all simulations, in which the additive white Gaussian noise was added to the voltage signals before image reconstruction. This type of noise has been previously used in EIT scenarios such as mimicking detached and erroneous electrodes (Asfaw and Adler 2005). However, we note that additive white Gaussian noise may not accurately model all realistic factors such as electrode movement. Further work is needed in clinical situations to determine the applicable noise models. This SNR was chosen as it was the point where most misclassification began to occur in our previous study on raw voltage data classification (Dunne *et al* 2018a). Further, this SNR level represents a more noisy scenario than is typical with EIT devices (Avery *et al* 2017), which generally have SNRs in the 60 – 80 dB range. However, the 40 dB level attempts to mimic cases in the real world where SNR degrades, such as with patient-electrode contact (Asfaw and Adler 2005). The definition of the SNR used in this study is detailed in (Dunne *et al* 2018a).

The simulated data was generated in the EIDORS v. 3.9 package using an electrode "skip of 4" pattern, an injection current of 6 mA_{p-p} and the GREIT (Graz consensus Reconstruction algorithm for EIT) time-difference image reconstruction algorithm (Adler *et al* 2009). The following parameters were used for GREIT image reconstruction: Noise Figure = 0.5; Target Size = 0.03; Target Offset = 0; 500 Targets. Each reconstructed image consisted of a 32x32 pixel array, in which the pixel intensity is indicative of the conductivity change in the region between the target and the reference.

Table 1. Simulation dataset properties.

Property	Values
SNR	40 dB
Volumes	
- Not Full	{40, 80, 120, 160, 200, 240, 280, 290} ml
- Full	{300, 310, 320, 340, 360, 380, 400, 420} ml
Urine Conductivities	{0.75, 2, 3.25} S m^{-1}
Reference Volumes	{30, 50} ml
Reference Conductivities	{0.75, 2, 3.25} S m^{-1}

For each scenario, 10 observations were obtained that differed only by the effect of the additional white gaussian noise; for each observation, a new image was formed. The total number of observations in the dataset was 2880. Example images for two specific observations are shown in figure 2 for 40 ml and 400 ml bladders, both with a conductivity of 3.25 S m^{-1} and a 30 ml reference of conductivity 0.75 S m^{-1}.
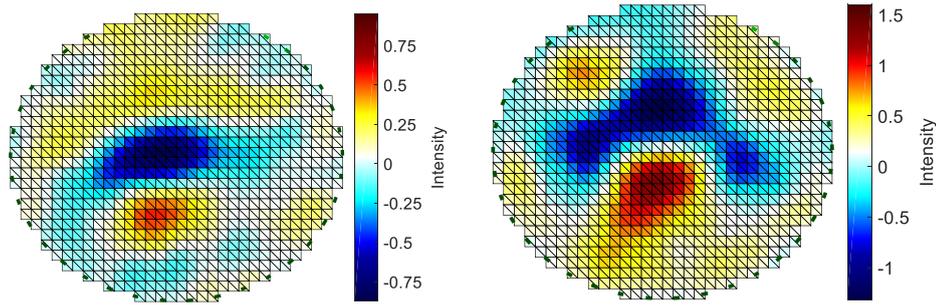
Figure 2: Example time-difference images using a 40 ml (left) and 400 ml (right) bladder volume of conductivity 3.25 S m^{-1} and a 30 ml reference of conductivity 0.75 S m^{-1}. The SNR was 40 dB. The intensity of the pixels relates to the conductivity change between the target and the reference. The full bladder is easily discernible (red). However, the full bladder image does have artefacts that become more pronounced with smaller differences between the target volume and reference volume and smaller changes in conductivity. The effects of a small contrast between the target volume and the reference volume are clearly seen in the 40 ml bladder image.

2.2. Phantom Dataset Collection & Preparation

Previously, we created the first anatomically and conductively realistic phantom for bladder volume monitoring (BVM) using EIT (Dunne *et al* 2018b). This modular phantom consisted of graphite, carbon black, and polyurethane to form the solid outer pelvic wall, with a homogenous fluid (ultrasound gel) inside that represented the weighted average of the conductivities for the organs in the pelvic cavity. Bladder phantoms were constructed using the same material ingredients as the outer wall, but with different quantities of each material in order to achieve the desired conductivities. The phantom formation process and the electrical properties of each component are explained in (Dunne *et al* 2018b). We used this phantom (as shown in figure 3) to gather experimental data. We also further expand on the bladder volumes in (Dunne *et al* 2018b) by forming additional volumes over the three conductivities {1.15, 1.68, 2.24} S m^{-1}. The 12 bladder volumes used in the study were {40, 60, 100, 160, 240, 260, 280, 300, 320, 340, 360, 380} ml. In total, thirty phantom bladders were used as target volumes and six phantom bladders as references.

Figure 3: *(left)* Conductively and anatomically-accurate male child-sized pelvic phantom for EIT as presented in (Dunne *et al* 2018b). The internal conductivities are modelled as a weighted average of all pelvic organs (0.2 S m^{-1}). This region is realized using ultrasound gel with additional sodium chloride. *(right)* The employed bladder phantoms, ranging from 40-380 ml.

A Swisstom Pioneer Set with the 32 electrode Swisstom 44 cm child belt was used to gather the data from the pelvic phantom. Raw voltage data was captured for each bladder volume. The data was captured at 15 frame s^{-1}, with a 50 kHz injection frequency, an injection current of 6 mA$_{p-p}$ and a skip of 4 (i.e., the number of electrodes that are skipped between the two injection electrodes, and similarly between the two measurement electrodes). Similar to the simulation setup, the bottom location of the bladders was maintained constant as the bladder volumes increase. For each measurement scenario, 10 frames were recorded. Different reference frames were obtained by averaging 10 frames for each scenario of 40 ml and 60 ml bladder volumes. Then, time-difference images were generated using combinations of observed bladder frames with the references. Unlike the simulation scenario, image reconstruction of the experimental phantom uses a tailored FEM (i.e. the same shape as the pelvic phantom) (Dunne *et al* 2017; Dunne *et al* 2018b) in order to minimize image reconstruction errors in the forward problem due to the ill-

posed problem of EIT. The previously outlined GREIT settings were used and each image reconstructed was a 32x32 pixel array. The dataset consisted of 1800 observations in total. The properties of the dataset are outlined in table 2.

Table 2. Phantom Dataset Properties

Property	Values
Bladder Volumes	
- Not Full	{100, 160, 240, 260, 280} ml
- Full	{300, 320, 340, 360, 380} ml
Urine Conductivities*	{1.15, 1.68, 2.24} S m^{-1}
Bladder Reference Volumes	{40, 60} ml
Bladder Reference Conductivities*	{1.15, 1.68, 2.24} S m^{-1}

*Conductivities of the bladder phantoms at 50 kHz

3. Pixel-Wise Classification

Pixel-wise classification involves using the pixel intensities of the reconstructed images (either from the entire image or a cropped portion). In this portion of the study, we use the entire image in order ensure the maximum amount of useful information is input into the classifier.

Each pixel was given to the classifier as a feature. In order to do this, the 32x32 pixel arrays were converted into 1x1024 row vectors. Columns with all zero (or NaN) entries were removed, e.g. pixels outside the region delimited by the pelvic FEM. The dataset size was then 2880x704 for the simulation dataset and 1800x525 for the experimental phantom dataset (excluding labels). The difference in the reduced number of columns between the simulation and experiment pixel-wise data is due to the different FEMs used during image reconstruction, as discussed in the previous section.

4. Feature Extraction

Feature extraction is used to reduce the dimensionality of a large dataset, and ensures that only the key information that more specifically describes the observations in the dataset is retained (Santorelli *et al* 2014a). This process can reduce training and testing computational time and saves storage space due to the reduced number of features. A number of methods to extract features exist such as exacting features with feature descriptor tools or analysing the collected data using dimensional reduction tools.

Common advanced feature descriptor algorithms include Histogram of Oriented Gradients (HOG) and Speeded-up Robust Features (SURF) (Bay *et al* 2008). However, these algorithms work best for high frequency transitions in the images, i.e. where there are sharp changes or well-defined edges, such as in finger print detection (Hany and Akter 2015). The images from EIT have low frequency transitions as can be seen in the reconstructed images in figure 2. Thus, we perform FE with features more suitable to this application that are manually crafted from the region of interest (ROI), image segmentation, and wavelets. These features are discussed in the subsequent subsections. In terms of feature reduction techniques, we use principal component analysis (PCA), which is a common technique employed in medical imaging classification (Nandi *et al* 2015; Oliveira *et al* 2015; Li *et al* 2017). Thus, for FE, we first use manual FE (MFE) and then apply PCA.

4.1. Region of Interest

The image reconstruction algorithm results in a 32x32 pixel array for each image. However, this array contains areas outside the FEM (to make a square) and areas within the pelvic FEM but outside the bladder region that did not provide additional useful information for monitoring the bladder volume. Thus, we cropped each image to a region of interest (ROI), determined from an ideal case.

The urinary bladder has, on average, the highest conductivity in the human body (Dunne *et al* 2018b). As the bladder volume increases, the average and maximum pixel intensities also increase. Using an ideal case of the maximum full bladder volume and the smallest reference bladder volume (at the same urine conductivity and a noiseless image), we determined the ROI pixel locations. We took the horizontal and vertical vectors that contained the maximum pixel intensity location in a 420 ml full bladder image (figure 4) and then calculated the 2.6 standard deviation of each of these vectors. This led to a ROI described by a 16x18 pixel array, centred around the maximum value in the image as shown in figure 4.

This ROI was extracted from each simulation and phantom image for use in defining and applying ROI-based features. The cross-sectional bladder locations in both simulation and phantom were approximately the same.
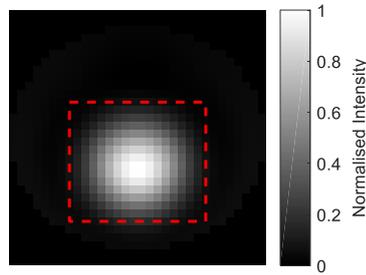

Figure 4: A grey-scale noiseless image of a 420 ml full bladder reconstructed using a noiseless 30 ml reference image. The red box is the 16x18 ROI centred around the maximum location in the image. This region, in the exact location shown, is applied to all simulated and experimental images in this portion of the study, prior to feature extraction.

4.2. Image Segmentation

To determine the bladder shape from the image, segmentation was employed. A previous imaging study for BVM using EIT (Schlebusch *et al* 2014a) used Canny's Edge Detection Algorithm (Canny 1986) to segment the bladder in order to associate the number of pixels (area) with the bladder volume. However, EIT images suffer from low-frequency transitions that make determining the exact edges difficult. Thus, a more suitable algorithm may be region growing, which involves defining the seed point for the algorithm and then allowing the algorithm to grow until a threshold is reached. Thus, we applied this algorithm on the ROI (after applying a grey-scale transformation) in order to segment the image. The region growing algorithm uses the mean intensity measure with a 15% threshold and the seed was defined as the maximum pixel intensity in the image.

4.3. Discrete Haar Wavelet Transformation

Properties of discrete wavelets such as the discrete statistics, the power, and the entropy have been previously shown to be useful features in imaging classification studies (Huang and Aviyente 2008; Rajan, Ramesh and Yuvaraj 2014). Thus, we further process the ROI using the 2D Discrete Haar Wavelet Transform (DHWT) (Ruch *et al* 2011) and interpret the coefficients returned as a signal, which is then used as the basis for some of the proposed features outlined in subsection 4.4.

4.4. MFEs

The features in table 3 are based on the ROI, the image segmentation, or the coefficient signal from the DHWT. This list is a small but a key sample of features that could possibly be extracted from EIT images of the bladder (10 features total).

Table 3. Descriptions of the MFE

Feature Name	Description
Mean of the ROI	Applied to all pixels in the ROI. The metric has been previously shown to be an effective method of classifying bladder state (Dunne *et al* 2018c).
Standard Deviation of the ROI	Applied to all pixels in the ROI. The metric can help distinguish small bladder volumes where there may be high background noise due to poor contrast in difference imaging.
Normalized Centre of Mass *(2 features)*	Centre of mass (COM) normalized by the respective length of region of interest. The COM is calculated on greyscale pixel intensities. As images are 2D, the COM provides two features: 'x' and 'y' coordinates.
Area of the Segment	As the bladder fills and widens, the cross-sectional area changes. Thus, we calculate the area by determining the number of pixels within the segmented bladder.
Area of the Equivalent Ellipse	The bladder is often modelled as an ellipsoid. Taking the cross section will give an ellipse whose area is dependent on the bladder volume. In order to calculate the elliptical area, the 'x' and 'y' lengths of the segmented bladder are first determined around the COM.
Perimeter of the Segment	As the bladder widens, the cross-sectional parameter also changes. The perimeter is determined by counting the number of pixels on the edge of the bladder segment.
Total Power of the Wavelet	The total power of the signal is defined as the coefficients from the Fast Fourier Transform multiplied by the conjugate of the coefficients.
Maximum of Wavelet	Taking the maximum of the wavelet coefficients from the DHWT of the image.
Log Energy Entropy of the Wavelet	Taking the log energy entropy of the wavelet coefficients from the DHWT of the image.

4.5. MFE-based Classification Dataset Preparation

For MFE-based classification, the image dataset was analysed in terms of the features as previously discussed, reducing the overall number of features fed into the classifier. The resulting dataset sizes for simulation and experimental classification were 2880x10 and 1800x10 respectively.

4.6. Principal Component Analysis

PCA uses the maximum variances of a dataset to reduce the dimensionality and to create principal components (PCs) that can be used as new features for the classification. Thus, in this work, we applied the technique to the pixel-wise simulation and experimental datasets to determine how PCA can aid in reducing time and memory costs as well as to determine how PCA may affect the classification accuracy.

  For each training-test set pair (further explained in section 5), we applied PCA to the training set to establish the training set PC, transformation coefficients, and scores. Then, the training transformation coefficients were applied to the corresponding test set in order to gather the test scores for assessing the classifier. The training and test sets were centred and normalized based on the training set mean and standard deviation. This ensures the test set is not contaminated by previous knowledge of the training set. The applied PCA methodology is similar to one applied by Howley *et al* 2006.

  We found that the first 100 components explained 99.997% of the total variance. Thus, we analysed the performance for first 100 components to determine the best number of PCs to use in order to maximize the generalized accuracy. Then, this best number was selected for use in the test sets.

5. Classifier & Classification Architecture

5.1. Classifier

Support-vector machines (SVMs) are a commonly used machine learning algorithm. SVMs transform the data into a high-dimensional space in order to separate the two classes of input data using a hyperplane (Schölkopf 2002). Linear SVM attempts to separate the input data by using a line to define the hyperplane that maximizes the margins between the two classes in the higher dimensional space. This machine learning algorithm has previously been shown to be effective for bladder monitoring on raw voltage data (Dunne *et al* 2018a). The Gaussian Radial Basis Function (GRBF) is another commonly used kernel for SVM in medical imaging studies (El-Naqa *et al* 2002; Santorelli *et al* 2014b). To examine the impact of kernel choice in bladder state detection, we compare linear SVM to GRBF SVM.

  Specifically, in this work, we use SVM with both kernels to classify 'not full' and 'full' bladder

states. Classification with both kernels is performed with each feature type. We examine the effect of each factor, e.g. bladder volume and conductivity, on classification accuracy. Further, we use Bayesian Optimization (MathWorks 2018) to minimize the generalization error by adjusting the hyperparameters: the box constraint and the kernel scaling factor. This process helps prevent overfitting.

5.2. Classification Architecture

To train and evaluate the classifier, we use 10-fold testing as described in (Li *et al* 2017; Dunne *et al* 2018a). The method repeats the training and testing 10 times, never testing again on the same observations. The advantage of 10-fold testing is that we get a more generalized error, which is important for BVM using EIT as the known factors such as urine conductivity vary both intra-individually and inter-individually (Schlebusch *et al* 2014b). The classification architecture and dataset splits are described in figure 5.

To ensure that the training is generalized, we perform 10-fold cross-validation with Bayesian Optimization to choose the ideal box constraint and kernel scaling factor. These parameters are chosen based on the lowest generalized error in the cross-validation and are then used for the final trained classifier with the test set. The architecture applies to both pixel-wise and FE-based classification. However, in FE-based classification, unsupervised feature selection is performed on the training data. This process is further outlined in section 6.
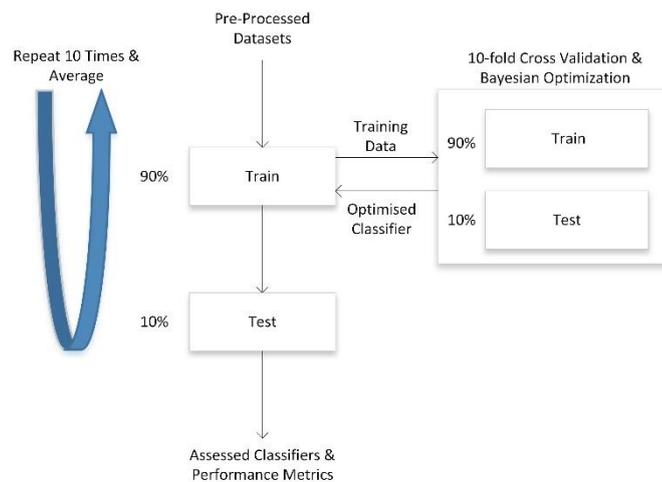

Figure 5: Classification architecture for both simulation and experimental scenarios. The architecture ensures fair testing of the dataset by never testing on the same test set during each of the 10 iterations.

6. Feature Selection

Two methods of FE are applied in this paper, MFE and PCA. Both aim to reduce the complexity of the dataset in order to improve performance in terms of time, memory and possibly accuracy. In order to choose the best performing features for each FE method and the individual classifiers, feature selection techniques were applied. These are further discussed in the following subsequent subsections.

6.1. MFE Feature Selection

In order to rank the important features, the unsupervised Laplacian Score (He, Cai and Niyogi 2005) was employed for MFE-based classification. The main assumption is that two data points close together relate to a single class (He, Cai and Niyogi 2005). The algorithm is unsupervised, thus, the algorithm can be applied on the whole training dataset regardless of the supervised processes without biasing the result.

We applied the feature selection algorithm (using Euclidean distance to form the weight matrix of the Laplacian score) to the training data and then, the 10-fold validation was run with both SVM kernels. The performance of the number of features extracted was determined by the averaged generalized accuracy from the cross-validation process. This process was carried out by first using only the highest ranked feature and then sequentially adding features until all 10 were included in the training set. The generalized accuracy for each number of ranked features is given in figure 6.

Figure 6: Analysing the optimum number of features for the MFE-based classification using the generalized accuracy gathered from the cross validation and Bayesian optimization. The features were added based on the ranking for feature selection. The highest generalized accuracy is observed at 9 features for Linear SVM and at 10 features for GRBF SVM, thus 9 and 10 features are our optimum number of features from the feature selection process for each classifier, respectively.

From figure 6, the general trend is that as features are added, the generalization accuracy increases. For linear SVM, the 2nd and 10th highly ranked features (perimeter of the segment and the 'x' normalized COM, respectively) did not increase generalization accuracy. However, in combination with lower ranked features, the perimeter of the segment did contribute to higher generalized accuracies. Thus, based on our chosen features and the ranking, the top nine features were chosen for FE-based classification with linear SVM. In other words, all features are used except for 'x' normalized COM.

For the GRBF SVM, feature 6 (the elliptical area) did not increase the generalized accuracy. However, when the features were put in combination with more features of lower ranking, the generalized accuracy was increased. Therefore, all 10 features were beneficial to the GRBF SVM and thus, were used in the FE-based classification.

6.2. PCA Component Selection

Having reduced the number of components (discussed in Section 4.6), the generalized accuracy was established for each classifier and set of components on the training set. The generalized accuracies are given in figure 7.
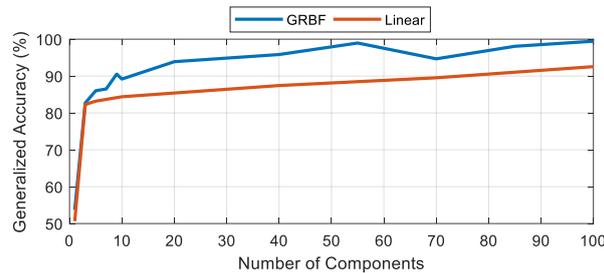

Figure 7: Determining the optimum number of input PCs for PCA-based classification. The GRBF SVM outperforms linear SVM in terms of generalized accuracy on the 10-fold validation with Bayesian Optimization. However, both kernel SVMs have their maximum generalized accuracy at 100 PCs.

For both GRBF and linear SVMs, the best generalized accuracy is when all of the first 100 components are given to the classifier. Thus, we use 100 components to test the PCA-based classification on the test sets.

7. Results & Discussion

In this section, we discuss the results of both pixel-wise and FE-based classification of simulation and experimental data. We investigate how the various parameters of bladder volume, urine conductivity, reference data, and noise impact the classification accuracy. We also examine where and why the most misclassifications occur.

7.1. Simulation Results & Discussion

For the full simulation dataset, the performance metrics of accuracy, sensitivity, and specificity were calculated from the classification outcomes with each feature type (pixel-wise, MFE-based, and PCA-based). The average value for each performance metric over 10-fold testing is given in tables 4, 5, and 6, for the different feature types respectively. For all feature types, the GRBF SVM outperforms linear SVM, showing the GRBF kernel best separates the features from the given datasets. For these scenarios, the GRBF SVM tends to prioritize the sensitivity while the linear SVM tends to prioritize the specificity. Favouring the sensitivity is preferable in this application, in order to alert the child to go to the toilet before involuntary urination occurs. To further prioritize the sensitivity, the misclassification cost may be adjusted for training in future work.

Between each of the three feature types in tables 4-6, PCA-based classification performs the best with the test set accuracy of 98.61 ± 1.88%. This result surpasses the next-best performer, pixel-wise classification, by approximately 3%. This outcome is not surprising as the use of PCA in classification with SVM has been previously shown to help improve the overall classification accuracy of high-dimensionality data (Howley *et al* 2006). The MFE-classification suffers the worst classification accuracy of the three techniques. However, only 10 features are employed compared to 100 features or more for both pixel-wise and PCA-based classification. Despite this, for this specific application (with a long-term goal of a wearable, wireless bladder monitor), the accuracy may be sacrificed if reduced memory is needed. This could make MFE-based classification more desirable.

To determine where the classifiers suffer in terms of accuracy, we analysed the misclassified bladder volumes, urine conductivities and references. For this analysis, we used the best performing kernel, GRBF. We found that approximately 91.80%, 64.04%, 95% of the bladder volumes were misclassified around 20 ml of the 300 ml bladder separation volume for the pixel-wise, MFE, and PCA-based classification, respectively. This outcome is in line with our previous study on the classification of the raw voltage data, where the biggest contributors to misclassification were the bladder volumes near the bladder state separation and low SNRs.

In terms of the urine conductivity, the lowest urine conductivity of 0.75 S m^{-1} accounted for the largest proportion of the misclassifications (approximately 42%) across the three urine conductivities and three feature types. This was followed by 2 S m^{-1} with around 30% and 3.25 S m^{-1} with around 28% of the total misclassifications. The majority of these misclassifications at 0.75 S m^{-1} occurred for the bladder volumes between 280-320 ml. The reason for the poor performance at low urine conductivities may be due to difficulty in discriminating between close bladder volumes as there is a decreased contrast in pixel intensities between the reference conductivity and the urine conductivity at 0.75 S m^{-1}.

The third and final factor that we varied was the reference image. Reference images were varied across two volumes with three conductivities each. The 50 ml reference tended to cause more misclassifications than the 30 ml reference by a maximum of 15% of the total misclassifications across results with all feature types. However, for PCA-based classification, there was no difference in the number misclassified by the 30 ml and 50 ml reference images. The difference in misclassification between the reference conductivities at 30 ml and 50 ml was small (within 13% of all misclassified) across results with all feature types. However, it is important to note that at 50 ml, the majority of misclassifications were at a conductivity of 0.75 S m^{-1}, highlighting the weaker contrast hypothesis for the most misclassifications at lower urine conductivities.

Thus, for bladder state image classification, the bladder volumes around the bladder state separation volume had the biggest role in misclassification.

Table 4. Pixel-wise simulation classification results for both classifiers

Metric	GRBF SVM	Linear SVM
Accuracy	95.76 ± 3.58%	92.64 ± 1.94%
Sensitivity	96.08 ± 3.98%	91.96 ± 2.98%
Specificity	95.51 ± 3.54%	93.35 ± 2.49%

Table 5. MFE-based simulation classification results for both classifiers

Metric	GRBF SVM	Linear SVM
Accuracy	84.17 ± 2.89%	77.50 ± 2.90%
Sensitivity	86.79 ± 2.68%	76.26 ± 3.30%
Specificity	81.70 ± 5.83%	78.83 ± 4.07%

Table 6. PCA-based simulation classification results for both classifiers with 100 components

Metric	GRBF SVM	Linear SVM
Accuracy	98.61 ± 1.88%	92.47 ± 1.19%
Sensitivity	98.72 ± 1.47%	91.93 ± 2.55%
Specificity	98.45 ± 2.53%	93.03 ± 1.42%

7.2. Experimental Results & Discussion

The performance metrics obtained with the experimental data are given in table 7 for MFE-based classification only, as both pixel-wise and PCA-based classification achieved a classification accuracy of 99.99% or greater on the test set for both SVM kernels. As in the simulation results, and as expected, pixel-wise and PCA-based classification outperformed MFE-based classification.

There are a number of factors that could lead to the superior experimental performance relative to simulation. The simulation dataset included more bladder volumes around the bladder separation volume than did the experimental data, i.e. ±20 ml in phantom experiments compared to ±10 ml for the simulation data. We observed in the simulation results that most bladder volumes were misclassified around the bladder state separation (300 ml). Thus, we can confidently classify the phantom bladder state within 20 ml volume resolution with the pixel-wise and PCA-based classification. For bladder state classification, smaller volume resolutions do not pose a problem as the classifier can be biased to prioritizing either the sensitivity or specificity.

The simulation dataset was also generated with a high level of noise (SNR of 40 dB) that could possibly mimic some of the unknown real-world factors. Whereas, in the experiment, we used a phantom setup without any movements, and used commercially available EIT hardware that has an unknown (unpublished) SNR. However, typical EIT hardware has SNRs between 60-80 dB (Avery *et al* 2017). In order to validate and confirm this result, we further ran the simulation with the same scenarios as in table 1 but with a SNR of 80 dB. With 80 dB SNR, an accuracy of 100% was achieved with both kernels for pixel-wise and PCA-based classification. Similarly, using the GRBF SVM MFE-based classification, we found an accuracy of 97.60 ± 0.87%. This result is similar to the experimental results for the same kernel, thus, indicating that the data had a high SNR.

With the experimental data, the MFE-based classification performed comparably well with both kernels, having accuracies of over 94%. Overall, the behaviours observed with the experimental data with respect to variations in bladder volume, conductivity, and the reference were similar to those observed in the simulation results.

Table 7. MFE-based experimental classification results for both classifiers

Metric	GRBF SVM	Linear SVM
Accuracy	99.83 ± 0.37%	94.78 ± 0.91%
Sensitivity	99.65 ± 0.79%	92.69 ± 1.90%
Specificity	100 ± 0%	96.96 ± 1.93%

7.3. Computational & Memory Cost

In certain systems, power and/or memory size may be limited. In this case, classification accuracy may be traded for lower training time memory costs. Feature extraction helps in reducing the number of features required, which can result in lower training times and memory space. Thus, we assess how each feature type performs under training times and size of the dataset. The factors discussed in this paper are very patient-specific and therefore, ideally, the device needs to be trained on the data collected from the user. This may be done over a specific training period. The training time was used (rather than the test time), as this is where most of the time is spent in generating a classification result. The simulation dataset (the largest and most complex) was used for the analysis of cost of each feature type. For each kernel and feature type, the training was performed 3 times on a 7th generation Intel® Core™ i7 processor with 48 GB of 2133 MHz random access memory. The Bayesian Optimization was parallelized on the 4-core system. The training time results are given in table 8 and the 10-fold dataset sizes are given in table 9.

The GRBF kernel had the fastest training times regardless of feature type, as this kernel more easily separated the data compared to the linear kernel. The MFE-based classification had the quickest training times relative to the other feature types, but the MFE-based classifiers were also given the fewest number of features (9-10 features).

Table 8. Runtime cost for each classifier and each feature type

Feature Type	GRBF SVM Training Time (Minutes)	Linear SVM Training Time (Minutes)
Pixel-wise	30.66 ± 1.07	276.66 ± 19.18
MFE	4.82 ± 0.57	131.37 ± 8.73
PCA	5.39 ± 0.22	258.68 ± 3.44

In terms of memory cost with the 10-fold dataset, the MFE-based classification with the lowest number of features had the lowest memory cost. This could be advantageous in applications where memory space is costly or not available and the accuracy can be sacrificed. Each of the datasets for MFE and PCA-based classification require less than one seventh of the memory of pixel-wise datasets, providing a large saving on hardware.

Overall, in terms of the classification performance, training time and memory, PCA-based classification with the 100 features seems to be the optimum choice for image classification with bladder EIT images.

Table 9. Memory cost for each classifier and each feature type

Feature Type	Number of Features	10-Fold Dataset Size
Pixel-wise	704	1.64 GB
MFE		
- Gaussian	10	36.91 MB
- Linear	9	34.61 MB
PCA	100	232.75 MB

7.4. Image Classification & Voltage Data Classification

In this study, we hypothesized that adding spatial information would improve classification when compared to voltage data classification that we previously studied (Dunne *et al* 2018a). In our previous work, we analysed raw voltage data classification performed with the factors of different SNRs (including 40 dB), bladder volumes, and urine conductivities, using linear SVM. Linear SVM with the raw voltage data resulted in a classification accuracy of $98.07 \pm 0.27\%$. This performance was better than linear SVM accuracy found with image data in this study. However, the GRBF kernel on the image data slightly outperformed the linear SVM on the raw data ($98.61 \pm 1.88\%$ vs. $98.07 \pm 0.27\%$, respectively). Thus, showing that image classification is a feasible option for the application. It is also important to note that image classification requires another factor that the raw data does not, i.e., a reference image. This factor could explain the decrease in performance between the linear SVMs raw and image data performance along with the linear approximation of the image reconstruction algorithm GREIT.

Further work is needed in both a direct comparison with the raw and image data and to explore how clinical factors (posture, movement, contact, etc.) could impact image classification accuracy and to further improve classification with pre-processing to reduce noise. However, overall, image classification shows promise for aiding in bladder EIT applications.

8. Conclusion

In this study, we performed the first machine learning classification on bladder EIT images. We used SVM classifiers to classify bladder state with complex datasets with varying bladder volume, urine conductivity, and reference data. The classifiers were tested on both simulation and experimental image data. In simulation, we applied high levels of noise (SNR of 40 dB) in order to mimic the effects of unknown real-world factors. We also investigated using GRBF and linear SVM kernels with three different datasets extracted from the image datasets: pixel-wise, MFE and PCA-based image classification.

We achieved high test set accuracies between 77-100% in our study. The biggest challenge in bladder state image classification is classifying the bladder volumes within ± 20 ml of the bladder state separation volume (300 ml in this study). However, in order to suit the needs of those with urinary incontinence, the sensitivity can be prioritized in the training of the classifier.

We found the best overall classification method in terms of classification performance was the combination of PCA with GRBF SVM (with an accuracy of $100 \pm 0\%$ in experiment and $98.61 \pm 1.88\%$ in simulation). Also, the classification performance from this combination was comparable to our previous raw voltage classification study, indicating viability of bladder EIT image classification. PCA-based classification was found to offer the best trade-off between classification performance and the costs of training time and memory. With the experimental data, we have shown that the bladder state can be confidently classified within 20 ml volume resolution using the classification techniques in this paper. Future work will include testing classification performance with varying body mass, posture, electrode contact and movement, changing of belt locations, and altering the image reconstruction setting such as the hyperparameters and the number of pixels in the images. Failing electrodes and compensation algorithms should be analysed for the application in the future, such as in (Mamatjan *et al* 2013). Also, other future work includes investigation of the using machine learning for not full and full classification compared to simple thresholding after transformation of image data to bladder volumes.

Overall, this work demonstrates the promise of EIT as a bladder monitoring support tool for those suffering from urinary incontinence.

Acknowledgements

References

Abrams P *et al* 2010 Fourth international consultation on incontinence recommendations of the international scientific committee: Evaluation and treatment of urinary incontinence, pelvic organ prolapse, and fecal incontinence *Neurourology and Urodynamics*, **29**(1) 213–240

Ackerman M J 1998 The visible human project *Proceedings of the IEEE*, **86**(3), 504–511

Adler A *et al* 2006 Uses and abuses of EIDORS: an extensible software base for EIT, *Physiological Measurement*. IOP

Publishing, **27**(5) S25–S42

Adler A *et al* 2009 GREIT: a unified approach to 2D linear EIT reconstruction of lung images *Physiological Measurement* **30**(6), S35–S55

Asfaw, Y and Adler, A 2005 Automatic detection of detached and erroneous electrodes in electrical impedance tomography *Physiological Measurement* **26**(2) S175-83

Avery J *et al* 2017 A versatile and reproducible multi-frequency electrical impedance tomography system *Sensors* **17**(2) 1-20

Bay H *et al* 2008 Speeded-Up Robust Features (SURF) *Computer Vision and Image Understanding* **110**(3) 346–359

Canny J 1986 A computational approach to edge detection. *IEEE transactions on pattern analysis and machine intelligence*, **8**(6) 679–698

Cao H *et al* 2011 A wireless strain sensor system for bladder volume monitoring in *2011 IEEE MTT-S International Microwave Symposium* 1–4

Coyne K S *et al* 2008 The impact of overactive bladder, incontinence and other lower urinary tract symptoms on quality of life, work productivity, sexuality and emotional well-being in men and women: results from the EPIC study *BJU International* **101**(11) 1388–1395

Dunne E, *et al* 2017 Realistic 3D finite element mesh of the adult human pelvis for electrical impedance tomography *Proc. 18th Intl. Conf. Biomedical Applications of Electrical Impedance Tomography* (Hanover, NH: Zenodo) p 65

Dunne, E *et al* 2018a Supervised Learning Classifiers for Electrical Impedance-based Bladder State Detection *Scientific Reports* **8**(1) 5363

Dunne E *et al* 2018b A realistic pelvic phantom for electrical impedance measurement *Physiological Measurement* **39**(3) 034001

Dunne E *et al* 2018c EIT Image-Based Bladder State Classification for Nocturnal Enuresis *19th Intl. Conf. Biomedical Applications of Electrical Impedance Tomography* (Edinburgh: Zenodo) p 40

El-Naqa I *et al* 2002 A support vector machine approach for detection of microcalcifications *IEEE Transactions on Medical Imaging* **21**(12) 1552–1563

Gibson W *et al* 2018 The association between lower urinary tract symptoms and falls: Forming a theoretical model for a research agenda *Neurourology and Urodynamics* **37**(1) 501–509

Gor R A, Fuhrer J and Schober J M 2012 A retrospective observational study of enuresis, daytime voiding symptoms, and response to medical therapy in children with attention deficit hyperactivity disorder and autism spectrum disorder *Journal of Pediatric Urology* **8**(3) 314–317

Hany U and Akter L 2015 Speeded-Up Robust Feature extraction and matching for fingerprint recognition *2nd International Conference on Electrical Engineering and Information and Communication Technology, iCEEiCT 2015*, pp. 21–23

He X, Cai D and Niyogi P 2005 Laplacian score for feature selection *Proc. of the 18th Int. Conf. on Neural Information Processing Systems* ed Weiss Y *et al* (Vancouver: MIT Press Cambridge) 507-514

Hirahara N *et al* 2006 Four-dimensional ultrasonography for dynamic bladder shape visualization and analysis during voiding *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine* **25**(3) 307–13

Howley T *et al* 2006 The effect of principal component analysis on machine learning accuracy with high-dimensional spectral data *Knowledge-Based Systems* **19**(5) 363–370

Huang K and Aviyente S 2008 Wavelet feature selection for image classification *IEEE Transactions on Image Processing* **17**(9) 1709–1720

Jalbani I K and Ather M H 2014 The accuracy of three-dimensional bladder ultrasonography in determining the residual urinary volume compared with conventional catheterisation *Arab Journal of Urology*. Arab Association of Urology **12**(3) 209–213

Kelly C E 2004 Evaluation of voiding dysfunction and measurement of bladder volume *Reviews in urology* **6** S32-7

Kristiansen N K, Djurhuus J C and Nygaard H 2004 Design and evaluation of an ultrasound-based bladder volume monitor *Medical and Biological Engineering and Computing* **42**(6) 762–769

Leonhäuser, D *et al* 2018 Evaluation of electrical impedance tomography for determination of urinary bladder volume: Comparison with standard ultrasound methods in healthy volunteers, *BioMedical Engineering Online*. BioMed Central, **17**(1) 1–13

Leonhardt S *et al* 2011 Electric impedance tomography for monitoring volume and size of the urinary bladder *Biomedizinische Technik/Biomedical Engineering* **56**(6) 301–307

Li R *et al* 2016 Preliminary Study of Assessing Bladder Urinary Volume Using Electrical Impedance Tomography *Journal of Medical and Biological Engineering* **36**(1) 71–79

Li Y *et al* 2017 Microwave breast cancer detection via cost-sensitive ensemble classifiers: Phantom and patient investigation *Biomedical Signal Processing and Control* **31** 366–376

Lottmann H B and Alova I 2007 Primary monosymptomatic nocturnal enuresis in children and adolescents *International Journal of Clinical Practice* **61** 8–16

Mamatjan, Y *et al* 2013 Compensating Electrode Errors Due to Electrode Detachment in Electrical Impedance Tomography *CMBES Conference* Ottawa

MathWorks 2018 *Bayesian Optimization Algorithm*, *- MATLAB & Simulink - MathWorks United Kingdom* Available at: https://uk.mathworks.com/help/stats/bayesian-optimization-algorithm.html (Accessed: 1 May 2018).

Morrison A and Levy R 2006 Fraction of nursing home admissions attributable to urinary incontinence *Value in Health* International Society for Pharmacoeconomics and Outcomes Research (ISPOR) **9**(4) 272–274

Nandi D *et al* 2015 Principal component analysis in medical image processing: a study *Int. J. Image Mining*, **1**(1) 65–86

Niestoruk L *et al* 2012 A concept for wearable long-term urinary bladder monitoring with ultrasound. Feasibility study *EDERC 2012 - Proceedings of the 5th European DSP in Education and Research Conference* 134–138

Niu, H *et al* 2011 Design of an ultrasound bladder volume measurement and alarm system *5th International Conference on Bioinformatics and Biomedical Engineering, iCBBE 2011* **M**(1) 3–6

Norton P and Brubaker L 2006 Urinary incontinence in women *Lancet* **367**(9504) 57–67

Oliveira B L *et al* 2015 Combined Breast Microwave Imaging and Diagnosis System in *Proceedings of Progress In Electromagnetics Research Symposium* 274–278.

Rajan A, Ramesh G P and Yuvaraj J 2014 Glaucomatous image Classification using Wavelet Transform in *2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies* 1398–1402

Ruch D K and Van Fleet P J 2011 The discrete haar wavelet transform and applications *Wavelet Theory* (Hoboken, NJ: Wiley)

Santorelli A, Li Y *et al* 2014a Investigation of classification algorithms for a prototype microwave breast cancer monitor *8th European Conference on Antennas and Propagation, EuCAP 2014* (EuCAP) 320–324

Santorelli A, Porter E *et al* 2014b Investigation of Classifiers for Tumor Detection With an Experimental Time-Domain Breast Screening System *Progress In Electromagnetics Research* **144** 45–57

Schlebusch T *et al* 2014a Bladder volume estimation from electrical impedance tomography *Physiological Measurement* **35**(9) 1813–1823

Schlebusch T *et al* 2014b Impedance ratio method for urine conductivity-invariant estimation of bladder volume *Journal of Electrical Bioimpedance* **5**(1) 48–54

Schlkopf B and Smola A 2001 Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Cambridge, MA: The MIT Press) 1-21

Shamliyan T A *et al* 2009 Male urinary incontinence: prevalence, risk factors, and preventive interventions *Reviews in urology* **11**(3) 145–165

Yip S K and Cardozo L 2007 Psychological morbidity and female urinary incontinence *Best Practice and Research: Clinical Obstetrics and Gynaecology* **21**(2) 321–329