Real-time Insertion Depth Tracking of Cochlear Implant Electrode Array with Bipolar Complex Impedance and Machine Intelligence

Nauman Hafeez, Nikolaos Boulgouris, Philip Begg, Richard Irving, Chris Coulson, Hao Wu, Huan Jia, Xinli Du

***Abstract*—Cochlear implants have significantly improved hearing for many as the most successful prosthesis, however, hearing outcomes vary. Uncertainty during electrode array (EA) insertion, including trauma and depth control, is one factor. To minimize radiation exposure from imaging methods like CT scans, this in-vitro study investigates the use of bipolar electrode impedance and artificial intelligent models to determine EA insertion depth. Complex impedance data was collected by inserting a commercial EA into a scaled-up 2D scala tympani model using a robotic feeder system. A support vector machine model produced a 98% classification accuracy for final insertion depth estimation. A CNN-LSTM hybrid model yielded 0.85 R-squared and 1.72 mm mean absolute error in depth estimation at each millimeter during a 25 mm insertion. This approach to depth assessment based on impedance may help with cochlear implant procedures and find use in other medical implant applications.**

***Index Terms*—cochlear implant, electrode array, bipolar impedance, insertion depth, machine intelligence**

I. INTRODUCTION

Cochlear Implant (CI) with all its shortcomings is still the most advanced prosthesis that is providing hearing treatment to people with severe to profound hearing loss [1]. One of the shortcomings is variability in post-surgical hearing outcomes that may be due to different factors. One such factor is trauma induced during the electrode array insertion that may erode residual hearing capabilities [2]. Another factor affecting CI performance is electrode array (EA) insertion depth [3]. The depth at which the electrodes are inserted into the cochlea determines their proximity to the auditory nerve fibers, which is an important factor in the ability of the implant to stimulate the auditory nerve and provide sound perception. Recently, Cochlear and MED-EL introduced their continuous impedance magnitude measuring tools (SmartNav and eIFT respectively) for better electrode placement.

In general, a deeper insertion depth is associated with better performance in terms of speech recognition and sound quality [4]–[6]. This is because the deeper electrodes can reach more auditory nerve fibers, which can improve the ability of the implant to stimulate the auditory nerve and provide a more robust and accurate representation of sound. However, deeper insertion depths can also be associated with a higher risk of surgical complications and may not be suitable for all patients [2].

Electrode array insertion depth can be assessed by medical imaging modalities e.g., Computed Tomography (CT), either at the end of the procedure or sometimes during the surgery if an imaging device is available in the operating room [7]. However, those imaging techniques come with the risk of radiation exposure and are mostly avoided especially during the insertion process [8]. It also increases the time and cost of the surgical procedure.

Several studies have used different methods to predict EA insertion depth before implantation as part of the surgical planning process [9]. In one of the studies, basal turn diameter and width were used to predict the angular insertion depth using a planning software [10]. The results were compared with the routine prediction method using 2D CT images. [11], [12] also presented methods to predict insertion depth.

In other biomedical applications, an ultrasound-based percutaneous needle insertion guidance robotic system has been developed which is used for real-time monitoring of positioning and orientation [13], [14]. Real-time monitoring of the insertion depth of an optical fiber-based needle array was carried out for cancer treatment [15].

We have seen complex impedance measurements vary concerning the EA position in the cochlear model during insertion [16]–[18]. Giardina et al. [19] briefly touched upon the relationship of response magnitude with insertion depth in their research with evidence from the in-vitro experiments. It was concluded that magnitude response (in mV) increases with the insertion depth, however, there was no further comment on whether it's possible to estimate insertion depth with impedance measurements. The magnitude response of most apical electrodes was compared at different insertion depths as well as magnitude response of most deepest and shallowest electrodes was also compared after the full insertion. Electrocochleography (ECochG) has also been used intraoperatively for monitoring EA insertion and there is a strong correlation between response to ECochG and EA scalar position during implantation [20]–[22]. Therefore, ECochG has the potential to estimate insertion depth, however, only in patients with residual hearing which is not always the case [23].

One of the studies [24], recently found a strong correlation between intra-operative impedances and insertion depth.

Nauman Hafeez and Xinli Du are with Department of Mechanical and Aerospace Engineering, Brunel University London, UB8 3PH, Uxbridge, UK. Nikolaos Boulgouris is with Department of Electronic and Electrical Engineering, Brunel University London, UB8 3PH, Uxbridge, UK. Philip Begg, Richard Irving and Chris Coulson are with University Hospitals Birmingham NHS Foundation Trust, B15 2GW, Birmingham, UK. Hao Hu and Huan Jia are with Department of Otolaryngology Head and Neck Surgery Shanghai Ninth People's Hospital, School of Medicine, Shanghai Jiaotong University, No 639 Zhizaoju Road, 200011, Shangai, China (Corresponding Author: Xinli Du (xinli.du@brunel.ac.uk))

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI: 10.1109/TMRB.2024.3407355, 2024 IEEE Transactions on Medical Robotics and Bionics

Tissue resistances were extrapolated from trans-impedance recordings and their relationship with insertion depth was statistically validated. Final electrode positions were acquired from CT images of 20 implanted subjects. In all subjects the same lateral wall electrode array was inserted. The proposed model can estimate the linear insertion depths by a margin of 0.76 ± 0.53 mm (mean and variance). In another study from the same group, postoperative insertion depths were estimated and tracked for up to 60 months using monopolar impedance recordings [25]. The insertion depth of all electrodes was estimated with an absolute error of $0.9mm \pm 0.6mm$ or $22^{\circ} \pm 18^{\circ}$ angle ($mean \pm SD$).

Post-surgery hearing outcomes dependency on EA insertion depth is very much debatable among researchers. For example, according to [5], it is concluded that deep electrode array insertion gives better speech perception whereas in another study it is the opposite due to potential effects of trauma because of deeper insertion [26]. However, CI surgery needs objective measures to improve the insertion process and assess surgical outcomes. These measures could also be used for real-time monitoring of the EA insertion process.

In this work, we have chosen to utilize complex bipolar electrical impedance of electrodes to track the linear insertion depth of EA. It is because impedance magnitude measuring capability is built into all commercial cochlear implants and it is viable to introduce implants with complex impedance measuring capabilities. Furthermore, previous studies [27], [28] have demonstrated that complex impedance has efficacy for EA localization. This work has employed two strategies to track linear insertion depth; the first one is based on classification that tries to distinguish partial insertion depths in the artificial ST whereas the other based on regression tries to capture each millimeter insertion depth using a recording of a single pair of electrodes. For classification, classical machine learning models are trained using complex impedance data of an electrode pair, and a hybrid model of convolution neural network (CNN) and long short-term memory (LSTM) network is used for regression.

II. Methodology

A. Data Recording Setup

An overall block diagram of the experimental setup is shown in Fig. 1. A 2D plastic ST model filled with 0.9% saline solution was used to mimic the inner ear structure. EVO® electrode array from Oticon Medical was utilized that has 20 equally spaced platinum-iridium electrodes having an active length of 24mm and insertion length of 25mm. It is a medium-sized EA compared to MED-EL FlexSoft/Flex34 (31.5/34mm) and Cochlear Slim20 (20mm) [29]. The electrode array was placed on a 3-degree-of-freedom (DoF) actuation system (Physik Instrumente (PI) GmbH & Co.) with a holder. There were x-y axes linear actuators and a rotational actuator. The x-axis linear actuator and the rotational actuator were used to adjust the electrode array insertion trajectory whereas the y-axis linear actuator was meant for electrode array insertion into the ST model. These actuators were connected to the PC via controllers to deliver instructions to them.

A custom-built complex impedance measuring meter was designed using data acquisition devices (DAQ) (National Instruments, Austin, TX) and a basic 2-element series circuit. One element of this circuit is a fixed 1 $k\Omega$ resistor R and the other is the electrode pair (EP). Impedance magnitude $|Z_{EP}|$ of an electrode pair is measured by applying a voltage signal V_{in} (generated by DAQ's voltage write port) of 1V (peak-peak) on this series circuit, measuring the current in the circuit by Ohm's law ($I_c = V_{in}/R$) and eventually $|Z_{EP}| = V_{EP}/I_c$ (V_{EP} is recorded by the DAQ's voltage read port). The impedance phase θ is measured as the difference between the current phase and voltage phase across EP. Impedance real (resistance, R) and imaginary (reactance, X) parts are calculated using the polar coordinates ($|Z|$ and θ) as $R = |Z|cos\theta$ and $X = |Z|sin\theta$. These are the 4 time-series data collected during the insertion process. Multiple electrode pairs can be recorded using a multiplexer controlled by the digital ports of a DAQ device.

A custom-built MATLAB software with a graphical user interface (GUI) was used for sending commands to both the actuation system and the impedance meter. Linear insertion depth and velocity of the insertion used to be chosen before the insertion as well as the number of electrode pairs for complex impedance measurement. As the EA is inserted in the ST model, the impedance of EPs was sequentially recorded one by one.

Complex bipolar impedance data recording of the electrodes during the insertion process was carried out by a system comprising of a 3-degree of freedom actuation system, a saline-filled plastic cochlea model, an impedance meter, and an electrode array attached to the actuation system as shown in Fig. 1.

B. Equivalent Circuit

The impedance model is based on Tykocinski et al. [30], and consists of three components, 1) resistance between electrodes due to bulk medium (saline), 2) polarization resistance, and 3) capacitive reactance due to electrode-electrolyte interface. Components (2) and (3) combined are called polarization impedance and can be modeled as a parallel circuit with (1) in series with them. According to this model, we not only need to measure impedance magnitude but also the phase, and impedance resistive and reactive components to fully comprehend the model and use these features for prediction. Since we are measuring bipolar impedance, the total impedance includes resistive and reactive components as shown in Fig. 2. R_3 is the bulk resistance between the 2 electrodes. On either side of it, we have polarization circuits for each electrode due to electrochemical reactions.

C. Data Collection

There are two datasets collected to analyze the hypothesis that complex bipolar impedance can be used to track the linear insertion depth of the electrode array during insertion. **Dataset-1** consists of complex impedance measurements of most apical electrode pair at different insertion depths. The electrode array was inserted 10 times for each insertion depth.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI: 10.1109/TMRB.2024.3407355, 2024 IEEE Transactions on Medical Robotics and Bionics

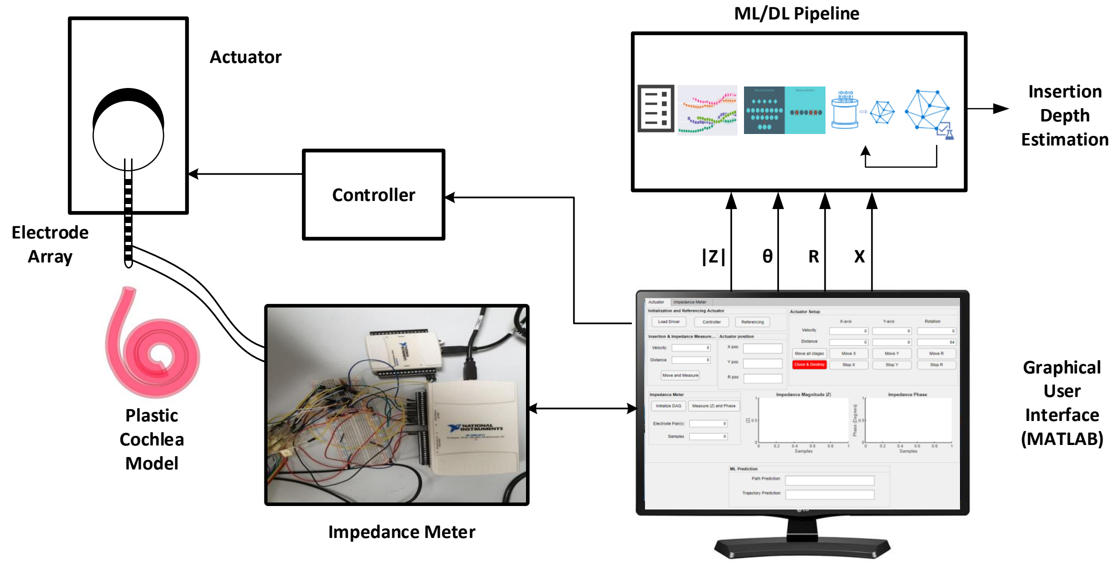

Fig. 1. The experimental setup for the automated electrode array insertion, impedance data collection, and analysis pipeline.

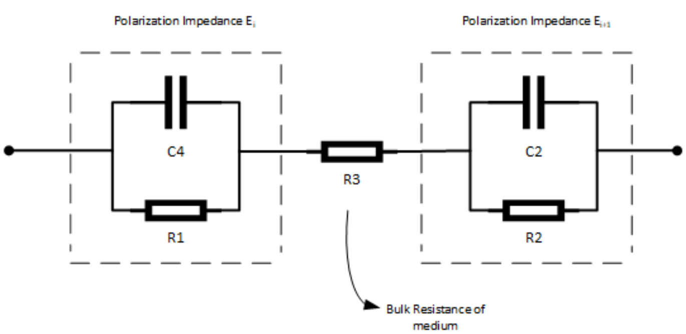

Fig. 2. Bipolar Impedance equivalent circuit model for the stimulation electrodes

Insertion depths we tried were 5 mm, 10 mm, 15 mm, 20 mm, and 25 mm as shown in Fig. 3. Therefore, we have 50 examples in this dataset. Each example has 4 time-series features (Z, θ, R, X). However, we collected 25 samples for a time series each even if the EA is inserted for less than 25 mm. For example, the system would measure 5 samples each for a time series of 5 mm and then the rest of the 20 samples would be measured when EA is static at 5 mm depth. This is to ensure, we have equal length time series for training and testing. The electrode array is inserted from the middle of the ST model in all experiments of this dataset. This dataset is used for the classification of different insertion depths of the electrode array.

Fig. 4 shows a graphical representation of the complex impedance measurements taken during the experiments. The graph shows impedance magnitude, phase, resistive, and reactive parts of the first electrode pair EP1 at different insertion depths. We have electrical measurements row-wise and different insertion depths column-wise. Graphs are plotted for 25 samples, however, measurements are taken both while the electrode array is continuously inserted and at rest for depths lower than 25mm. Each time sample corresponds to a 1mm insertion depth for 25mm depth. For other insertion depths, partial measurements are taken when the array is static after being inserted to the corresponding depth. The graphs shown in the figure are the mean and standard deviation of the respective insertion depth experimental data. Looking at the first column, when EA is inserted for 5 mm, impedance magnitude decreases even though EA remains at the center of the ST model. However, the decrease is small and the decrease is continuous during both instances when EA is moving (inserted) for 5mm and then static for the rest of the 20 time-samples. On the other hand, the impedance phase is getting less negative but the most change is observed when the electrode array is moving and not static. The change is also minimal, amounting to less than $0.5°$. In the same way, a decrease in resistance and reactance can also be seen. In the second column, more changes in all electrical properties are observed when the EA has touched the lateral wall and rests there. This is due to the proximity of EP1 to the plastic material which is a higher impedance material than the saline solution. We can see there is 100Ω increase in $|Z|$ and 1 change in phase θ. Similarly, for insertion depths of 15mm, 20mm, and 25mm, there are more pronounced changes in electrode pair complex impedance after it touches the lateral and slides along it or rests there.

For regression analysis, **dataset-2** with examples of insertions from different directions will be used for maximum coverage of the ST model. In dataset-2, the electrode array is inserted from three different angles (medial, middle, and lateral) where the middle position was the center of the plastic model and medial and right at $0.5°$ to either side from the center. The electrode array is inserted into the ST model

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information:
DOI: 10.1109/TMRB.2024.3407355, 2024 IEEE Transactions on Medical Robotics and Bionics

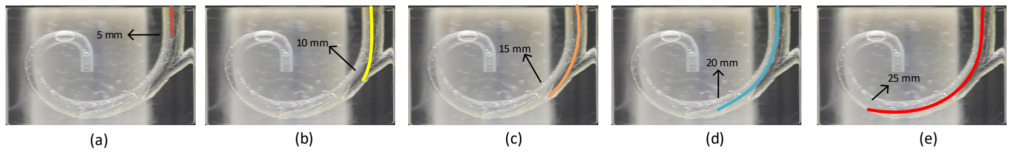

Fig. 3. Electrode array insertion at different insertion depths. (a)5 mm, (b)10 mm, (c)15 mm, (d)20 mm, (e)25 mm

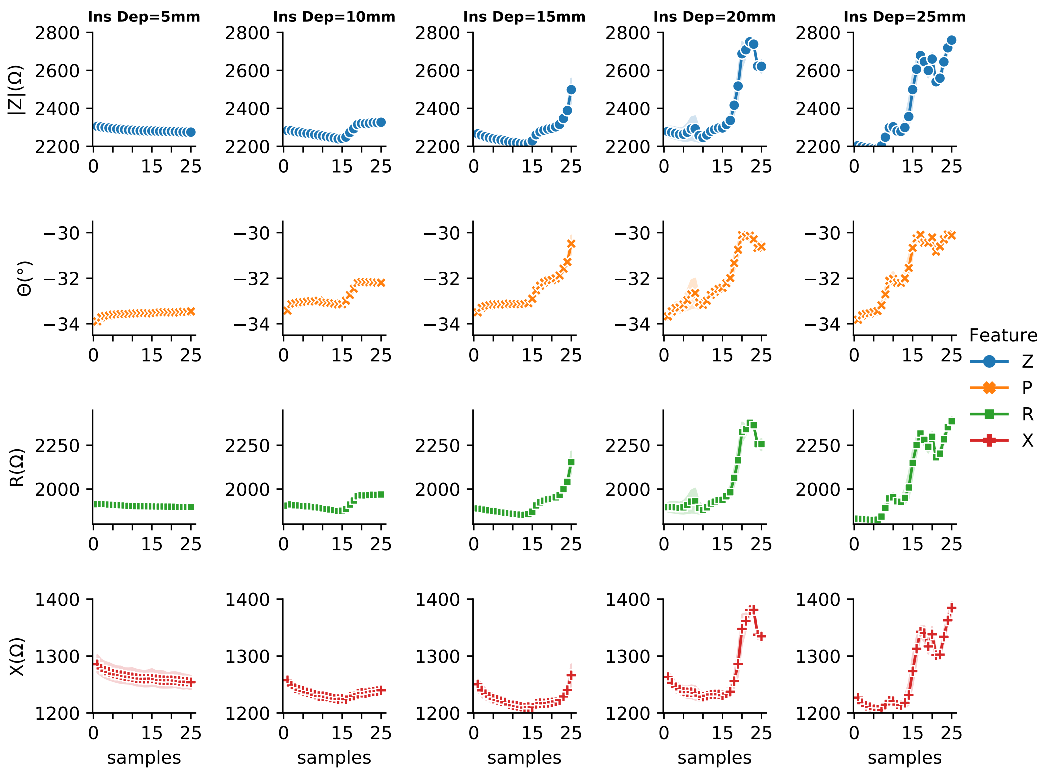

Fig. 4. Electrical impedance visualization of EP1 at different insertion depths. Columns[left-right]: 5 mm, 10 mm, 15 mm, 20 mm, 25 mm. Rows[top-bottom]: Impedance magnitude, phase, resistance, and reactance

for 25 mm at a speed of 0.08 mm/s by moving the vertical linear actuator (EA is placed on it with a holder). Impedance recordings of 8 electrode pairs (EP1:E1-E2, EP2:E3-E4, ... EP8:E15-E16) were taken with a sampling time of ≈1.5s. Therefore, each pair has 25 samples during complete insertion (1 sample/pair/mm). The most apical electrode pair EP1 is used in the analysis to predict mm-insertion depth.

D. Data and Models for Classification

Dataset-1 is used for the classification of different depths of insertion and it is defined as $D = \{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}) ...(x^{(i)}, y^{(i)})\}$ where $i = 50$, x is the input vector and y is the target label. The input x is presented as $x = \{\mathbf{Z}, \Theta, \mathbf{R}, \mathbf{X}\}$ where $\mathbf{Z} = \{|Z|_1, |Z|_2, ...|Z|_m\}$, $\Theta = \{\theta_1, \theta_2, ...\theta_m\}$, $\mathbf{R} = \{R_1, R_2, ...R_m\}$, and $\mathbf{X} = \{X_1, X_2, ...X_m\}$ where $m = 25$. The target label $y \in \{1, 2, 3, 4, 5\}$ represents 5 classes for different insertions depths (5mm, 10mm, 15mm, 20mm, and 25mm respectively). The data is first standardized by subtracting the mean and dividing it by the standard deviation. Then the machine learning models are trained and tested using 5-fold cross-validation scheme.

Supervised machine learning algorithms are used for classification namely, support vector machines (SVM), k-nearest neighbors (kNN), random forest (RF), and (shallow) artificial neural network (ANN).

Support Vector Machines: Support Vector Machines (SVM) were initially designed as discriminative functions and were mainly used for large-margin classification. Simply put, given a set of training examples, the SVM classifier learns the position of its support vectors and then subsequently the optimal hyperplane to categorize unseen samples. Owing to their impressive generalization ability, SVMs have been extensively studied to solve several classification tasks (see e.g. [31]–[33]). More recently, their use cases have expanded to include function approximation, regression, and time series prediction by introducing an appropriate loss function [34].

k**-Nearest Neighbours**: k-Nearest Neighbor (kNN) classifier is a fairly simple supervised learning algorithm that classifies a data point based on how its neighbors are classified.

kNN algorithm is based on feature similarity, that is, choosing the right k (number of neighbors) value is important for better accuracy. Another parameter is the distance measure that induces the relationship between neighbors [35], for example, Euclidean distance, Dynamic Time Warping (DTW), Manhatten distance, etc. [36] found that a simple 1-NN classifier with DTW distance measure generally produces better results than more complex classification algorithms on time series data.

Artificial Neural Network: ANNs are complex computational models that learn to solve a given task by extracting meaningful information from the training data. ANNs were first introduced in 1943 [37] and the authors presented a simplified version of the computational model based on how neurons in the brain perform complex tasks. An ANN is composed of simple computational nodes, called neurons, typically arranged in the form of layers. Each layer performs a simple linear transformation on the incoming data, followed by a point-wise non-linear function. These seemingly simple models have proven to be immensely powerful in solving problems that have resisted the efforts of the scientific community for many years.

Random Forest: Random Forest (RF) classifier first introduced by Breiman [38]. It is an ensemble of relatively uncorrelated decision trees where they vote for the most popular class and the class with the most votes becomes the model prediction for a particular sample. A decision tree is a set of nodes and leaves that are constructed based on a training set constituted of a feature collection. The purpose is to divide the training set into smaller subsets by carrying out a sequence of tests. When a subset contains only the cases belonging to a single class, the process terminates. These types of tests are called the nodes and subsets are known as the leaves. Predicting the class for a new data sample, the process starts at the tree root and finishes up at one of the leaves.

The k-fold cross-validation method is employed for training and testing the dataset to avoid any over-fitting and k=5 is chosen. The model was trained and tested 5 times in our case and produced the accuracy of the model in each iteration. During each iteration, the dataset was divided into 5 folds (portions); 4 folds were assigned as a training set and one fold as a test set so the training and test sets are different in each iteration. The final accuracy was the average of accuracies of all iterations.

There are four models (SVM, ANN, kNN, RF) that are cross-validated on specific and optimized hyperparameters. For ANN, we used a shallow multilevel perceptron (MLP) with two hidden layers of size layer1=100 neurons and layer2=10 neurons, and finally a classification layer with 3 neurons. The hidden layers had an associated non-linearity of rectified linear units (ReLU). A softmax layer was used to convert the output of the MLP into probabilities of the respective classes. The network was trained for 1000 epochs using an Adam optimizer at a learning rate of 0.001. For SVM, the Radial Basis Function (RBF) kernel function was used to train the data, and Dynamic Time Warping (DTW) was used as a distance metric and 5 number of neighbors for the kNN classifier model. RF is trained with 5 trees in the forest with Gini impurity to measure the quality of the tree split. We have not used any weight class metric as our dataset is balanced. All models are trained using the Python Sklearn library.

E. Data and Models for Regression

In the regression problem, the input X is mapped to the output Y by the algorithm $M : X \rightarrow Y$ where X is an input sample set $X = \{x^{(1)}, x^{(2)}, ...x^{(m)}\}$ and corresponding output target set $Y = \{y^{(1)}, y^{(2)}, ...y^{(m)}\}$ where m in our dataset is 3425. One of the examples of input x corresponds to $\{|Z|, \theta, R, X\}$ having a time sample each for impedance magnitude, phase, resistance, and reactance. The output target set Y values range from 1-25 mm as shown in Fig. 5. In this, we have used our insertion trajectory dataset to cover the dynamics of the ST model and different insertion paths. Linear regression algorithm is used to fit the data, however, after getting unsatisfactory results advanced convolution and recurrent neural networks are applied for this regression problem.

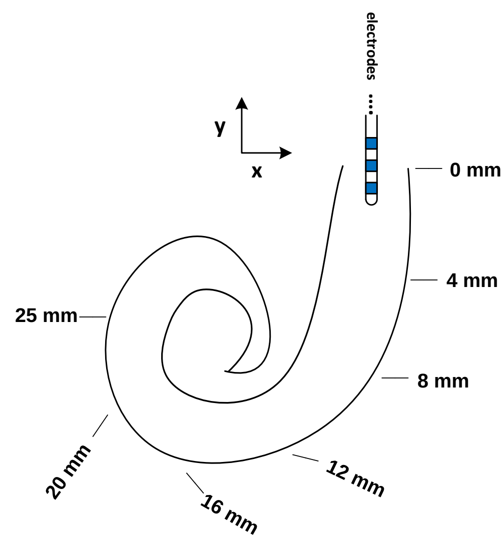

Fig. 5. Insertion depth blueprint of a 3:1 scaled-up ST model. EA is inserted to 25mm depth.

Convolution Neural Network (CNN): CNNs are commonly used for computer vision, but are also used for natural language processing and time series analysis [39]. They detect patterns in data and are useful for 2 and 3-dimensional data such as images and video frames. Recently, 1-dimensional CNNs have been introduced for 1-D data in the biomedical field [40]. Convolutional layers are the backbone of detecting patterns and extracting features from input data. A filter slides across the data and convolving is the process of taking the dot product of the filter values with the data. Each layer has different filters called kernels, and the data matrix is convolved with each kernel separately.

With the success of 2D and 3D CNNs, there are also 1D CNNs for time series or 1-dimensional data. So instead of 2D filters, there are 1D filters to slide over the 1D data and perform convolution. They produce 1D output for further processing in the next layers. Time series exhibit one dimensional data (time) instead of 2-D (width and height). Filters provide non-linear

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI: 10.1109/TMRB.2024.3407355, 2024 IEEE Transactions on Medical Robotics and Bionics

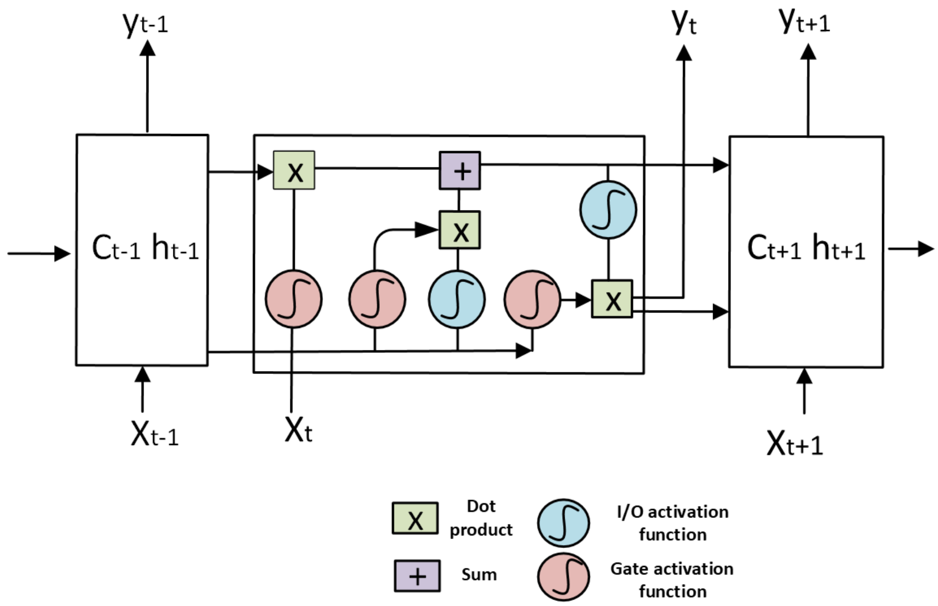

Fig. 6. Long Short Term Memory network block with key operations

transformation of 1D data. The convolution to the time series data can be presented as

$$C_t = f(w * X_{t-l/2:t+l/2} + b) \tag{1}$$

where C presents the outcome of the convolution process (dot-product) resulting from the application of a function f to a univariate time series X with filter w of length l. There is an additional bias factor b. f represents the final non-linear function applied to the convolution result for example a ReLU or sigmoid function. Convolving 1D data with multiple filters in a layer produces a multivariate output whose size is equal to the number of filters. Filter values are initialized randomly and they are learned automatically as the process goes according to the dataset. A pooling layer can be added after the convolution layer to reduce the size of the convolutional layer output by aggregating. Another type of layer called the normalization layer helps the network converge faster. The biggest advantage of CNN training is its speed compared to the conventional fully connected networks due to the weight-sharing mechanism.

Recurrent Neural Network (RNN): RNN is recommended for data with temporal information, but suffers from vanishing or exploding gradients during training [41]. LSTM, a type of RNN, solves this problem using memory blocks and gates. The gates control data flow and decide which information to keep or discard, similar to human memory.

Fig. 6 shows the inner structure of an LSTM block/cell and the operations involved in it. These operations allow the LSTM network to keep or let information during the training process. There are three LSTM blocks in the figure and the center one is elaborated. The other two blocks will help understand the flow of information from the previous to the next block. The memory block mentioned above is also called cell state. Cell state is like a memory that carries the information along the sequence of data. Gates decide which information to keep in the cell state. In Fig. 6, X is the input sequence, and C is the cell state, h is the hidden state. We have three neurons with sigmoid activation functions in pink (from left to right) first sigmoid is the forget gate, the second sigmoid is the input gate, and third sigmoid is the output gate. The sigmoid activation function outputs the value between 0 and 1. Any value close to 0 will be forgotten and values close to 1 will be kept in the cell. In the forget gate, current input X_t and previous hidden state h_{t-1} values are combined to form a vector and multiplied with the weights before passing on to the sigmoid function, the resulting value will decide whether to keep or forget it according to the rule mentioned earlier. Output of the forget gate for time sample t can be presented as

$$f_t = \delta(w_f[h_{t-1}, X_t] + b_f) \tag{2}$$

where f_t is the output of the forget gate, δ is the sigmoid function, w_f and b_f are the weight matrix and bias associated with the forget gate neuron, h_{t-1} is the hidden state of the previous cell, and X_t is the input. Before getting into the input gate, it's important to mention another activation function involved which is tanh in the blue color neurons. These are input/output activation functions. To update the cell state, h_{t-1} and X_t are combined to form a vector and multiplied with the weight matrix associated with the input gate before being forwarded to the input gate sigmoid function and the same information is passed on to tanh function. The input gate output and tanh function are presented as in 3 and 4

$$i_t = \delta(w_i[h_{t-1}, X_t] + b_i) \tag{3}$$

$$ti_t = tanh(w_{ti}[h_{t-1}, X_t] + b_{ti}) \tag{4}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI: 10.1109/TMRB.2024.3407355, 2024 IEEE Transactions on Medical Robotics and Bionics

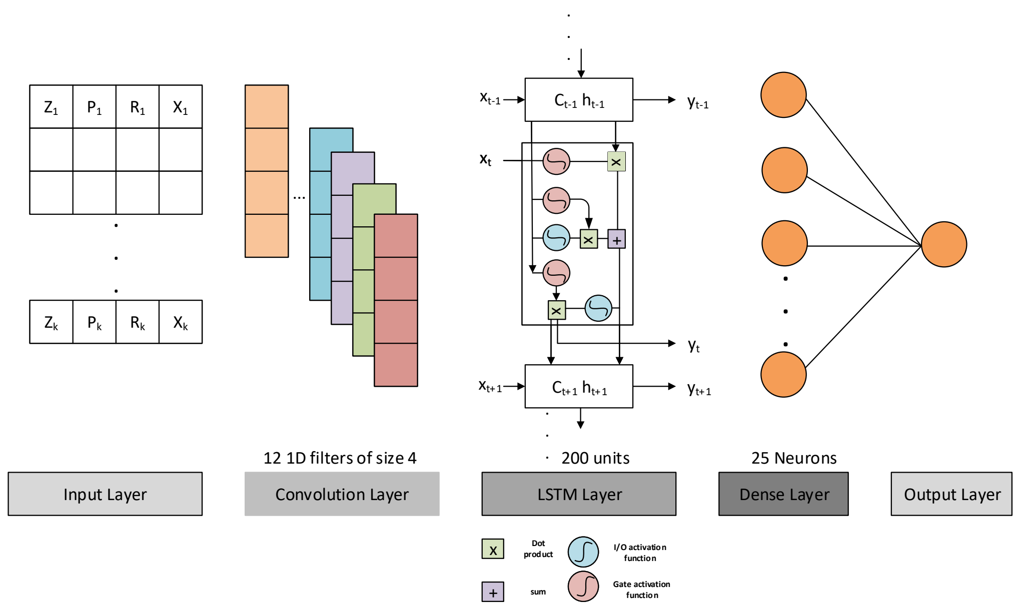

Fig. 7. Hybrid CNN-LSTM model for insertion depth prediction

Tanh function outputs the value between -1 and 1 to regulate the network. The dot product of the output from the tanh and sigmoid function is taken and this product would help decide which information from the tanh function to keep or leave out. To find the cell state, the forget state output is pointwise multiplied by the previous cell's cell state and added to the input gate output. This process updates the current cell state value and is presented as in 5.

$$c_t = (i_t \cdot ti_t) + (f_t \cdot c_{t-1}) \tag{5}$$

In the output gate, it calculates the hidden state that keeps the information of the previous inputs. First, the previous cell's hidden state and current input are combined to form a vector, multiplied by the weight matrix associated with the output neuron, and fed into the sigmoid function. It can be presented as in 6

$$o_t = \delta(w_o[h_{t-1}, X_t] + b_o) \tag{6}$$

On the other hand, the cell state is processed through a tanh function. The dot product of output from the sigmoid function and the output from the tanh output are taken to decide what information would be carried in the hidden state. In this way, both the cell state and hidden state are carried to the next LSTM cell. The current cell hidden state and output can be presented as in 7 and 8

$$h_t = o_t \cdot tanh(c_t) \tag{7}$$

$$y_t = w_y h_t + by \tag{8}$$

CNN-LSTM Hybrid Model: A hybrid 1D CNN and LSTM model (as shown in Fig. 7) is used for regression to predict linear insertion depth using complex bipolar impedance features. The model consists of an input layer, convolutional layer, LSTM layer, dense layer, and output layer. The convolution layer extracts spatial features, while the LSTM layer extracts temporal features. The input layer arranges the data for the convolutional layer. The model is trained on EP1 data ($|Z|$, θ, R and X) and can be applied to other electrode pairs. The dataset has 3425 examples, each with four recorded values and a target label between 1-25mm according to the depth of the most apical electrode during the insertion process.

Once we have the input layer setup, there is a 1-D convolutional layer. The reason for using the 1-D CNN layer is its low computational complexity compared to 2-D CNN and also 1-D CNN even with shallow architectures can learn from complex 1-D signals [40]. There are 12 1D filters of size 1x4. The selection of the number of filters is carried out as hyperparameter tuning starts from 4 filters. Each filter learns different features from the input 1-D signal. The size of the filter has been kept as the size of the input signal. The convolution layer produces the output (feature map) of the size 4x12 for each input training example. The feature map is fed into the LSTM layer with 200 cells. The number of LSTM cells is selected according to the size of its input from the CNN layer. Each cell works as explained in Fig.6 and carries the information to the next LSTM cell. The LSTM layer produces an output of the size 1x12 that is fed into the fully connected layer with 25 neurons and onto a single neuron output layer. The output layer finally gives the insertion depth prediction.

III. Results

A. Classification Results

Cross-validated accuracies of four machine learning algorithms are given in Table I with their standard deviation. Support vector machine gave the highest accuracy with 98% with an SD of 4%, whereas Shallow neural network and k-nearest

TABLE I
TEST ACCURACIES OF FIVE CLASS INSERTION DEPTH DATASET

EP	Accuracy % (SD)			
	ANN	SVM	kNN	RF
1	96.0(4.9)	98.0(4.0)	96.0(4.9)	92.0(9.8)

neighbors were equally good at giving the same accuracy of 96% with an SD of 4.9%. Random forest algorithm not only gave the lowest accuracy of 92% but also had the highest SD of 9.8%. It is important to mention that insertion depths are estimated/classified using 100% of the most apical EP data. The same procedure can be repeated with other EPs data or can also use a combination of EPs data.

Further analysis is carried out by looking into the confusion matrix that gives us a summary of the predictions of a classification problem. It gives information about the true positive, true negative, false positive, and false negative predictions with count values broken down by each class. Fig. 8 shows the confusion matrix of all 4 learning models. In each matrix, x-labels and y-labels are representing the five classes. The confusion matrix for SVM is at the top left corner, where we have only 2 incorrect predictions of insertion depth 15mm as 10mm insertion depth. All other insertion depth classes have a 100% true positive rate. With the ANN case on the right top corner, only one incorrect prediction of the 10mm insertion depth class is seen as 5mm insertion depth. All other classes have true positive predictions. The confusion matrix of kNN (left down) is similar to the matrix of SVM. For the random forest machine learning model, we can see three incorrect predictions; 10mm insertion depth is incorrectly predicted as 5mm, and 15mm insertion is once incorrectly predicted as 10mm and once as 5mm. All other 3 classes have 100% true positive rate.

With confusion matrix analysis, it is evident that the insertion depths of 10mm and 15mm have incorrect predictions whereas the rest of the classes have no incorrect predictions in all trained models. The reason may be explained that these insertion depths are at the crossroads of straight and curved insertion paths, especially 10mm depth. Having said that, with such a small dataset, the models are trained well and prediction accuracies are above 95%.

B. Regression Results

The learning model is developed in Python 3.6 using the Keras library with tensorflow on the backend. The dataset is divided into 85/15 ratios for training and testing the regression model. The activation function for the convolutional layer and dense layer is Rectified Linear Unit (ReLU) which suppresses all negative values and is a linear function for values above 0. The activation function for the LSTM layer is tanh and recurrent activation is the sigmoid function.

The model is trained for 1000 epochs with a batch size of 30. Adaptive moment estimation (ADAM) optimizer is used with a learning rate of 0.001 and mean squared error (MSE) as the loss function. To avoid overfitting, earlystopping has been used with a patience value of 200. This helps to stop the training process if the model is overfitting. Fig. 9 shows the progress of model training where the x-axis is the number of epochs and the y-axis is the MSE value. It is evident that the model is not overfitting as the MSE of the validation set is a little higher than the MSE of the train set and there is not a big difference between those RMSE values.

Linear insertion depth estimation was carried out by a hybrid learning model. The model is trained and tested on a dataset collected in vitro in a plastic ST model. Out of 3425 data examples, 2911 are used for training the model while the rest of 514 are used for testing the model performance. The performance measures of our model are depicted in Table II. The mean absolute error for the test dataset is 1.62mm which indicates good model performance. The MSE is 7.36 mm^2 which is a bit high concerning our data range of 1-25 mm. The reason for this high value is the inherited property of MSE to penalize large errors compared to small errors. Better performance measures of regression models are RMSE and R2 values that turned out to be 2.71 and 0.85 (85%) respectively.

TABLE II
REGRESSION PERFORMANCE METRICS FOR VALIDATION DATASET USING CNN-LSTM MODEL

EP	Dataset	Model: CNN+LSTM			
		MAE	MSE	$RMSE$	R^2
1	Test	1.62	7.36	2.71	0.85

To get a clearer visual representation of the predictive performance of the model, Fig.10 presents a fitted linear regression line. Actual targets are along the x-axis and model predicted values are along the y-axis. This graph also shows the distribution of the actual and predicted insertion depths on the top and right sides of it in green color. It can be seen, that there is even distribution of our test set in terms of linear insertion depth 1-25mm data samples. If we see the distribution of the predicted depths, the performance is good in the middle part of the insertion depths whereas prediction performance deteriorates for the deeper depths estimation, for example, from 22-25 mm depths. The insertion depth estimation is also not up to the mark at 1 mm. The reason for this may be the stabilization of impedance values at the beginning of the insertion.

IV. DISCUSSION

The results of the current study suggest that the complex impedance of the electrode pairs can be utilized for final and partial insertion depths (down to 1mm) estimation of the CI electrode array. Our machine-learning models correctly predicted different linear insertion depths with a true positive rate of up to 100%. The only exception is when we tried to predict a depth of 10-15mm, it is the time when the array starts to touch the first ST curve. The difference in time, when EA touches the curve when inserted from different angles, could be the cause of a slightly lower true positive rate in this region. For mm-depth prediction, the reason for using the deep learning model was a lower performance on linear regression

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI: 10.1109/TMRB.2024.3407355, 2024 IEEE Transactions on Medical Robotics and Bionics

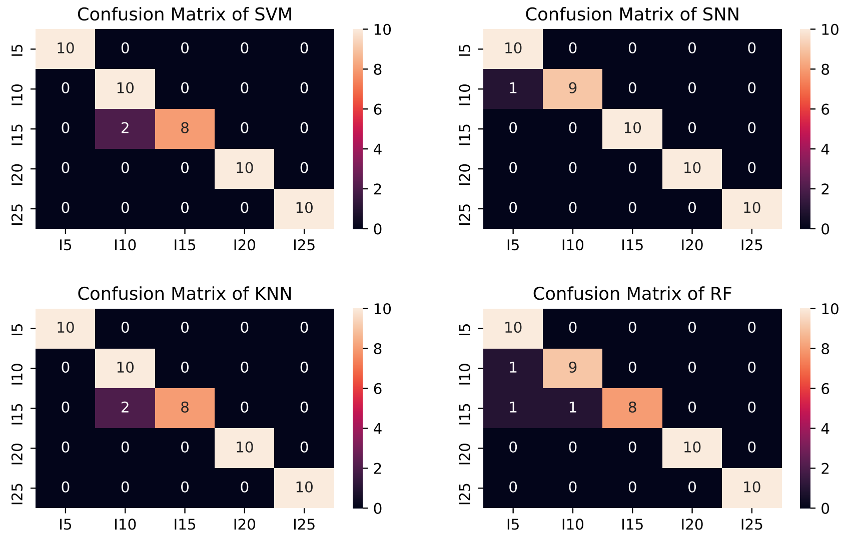

Fig. 8. Confusion matrix of classification

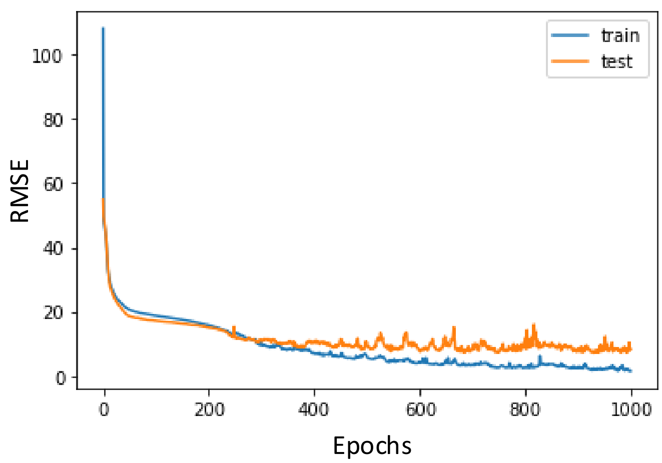

Fig. 9. CNN-LSTM hybrid model training for insertion depth regression

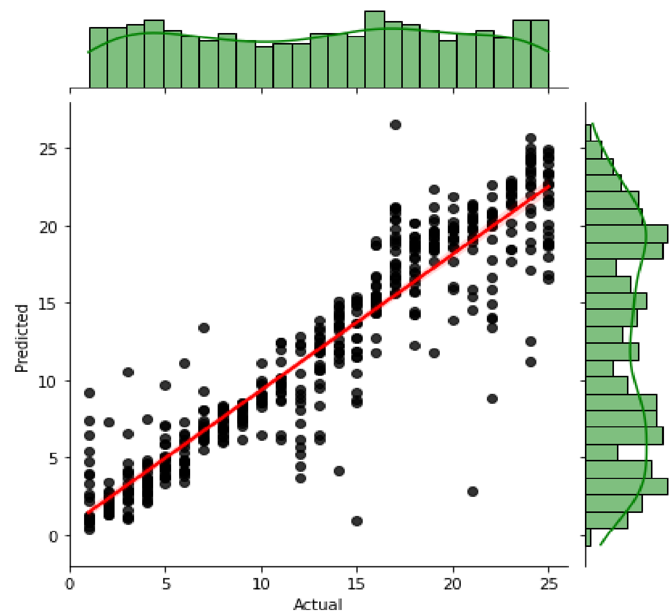

Fig. 10. Regression line with actual/predicted values distribution in the test set

and other machine learning models. A possible reason for better performance by our hybrid CNN-LSTM model is its ability to learn both spatial and temporal features.

There are certain limitations of this study, for example, we have used only most apical electrode pairs to train our models, however, these models should be tested on other electrode pairs. Having said that, it is easy to ascertain the depth of other electrodes as the distance between electrodes on the array is known. As we recorded the data on a fixed-size ST model, it is straightforward to convert our linear insertion depths to angular depths, it will be a different scenario when we have different size ST models like humans do. Another consideration is the use of a 3:1 scaled-up model where EA is inserted up to the second turn. Insertion depth estimation could be better if a normal adult-sized ST model is used, allowing the EA to be inserted for additional turns, which may enhance discrimination. The size and stiffness of the EA are also important factors while evaluating the insertion pattern using impedance measurement. The stiffness of the EA contributes heavily to the trauma caused during the insertion. Moreover, variation in stiffness among EAs may contribute to different impedance profiles along the path (e.g., we have observed impedance increase after the first turn even when insertion was stopped).

Future endeavors also include validating the developed algorithm on 3D cochlear models, potentially followed by animal and cadaver trials. This step aims to further refine and assess the algorithm's accuracy and reliability in more complex, realistic scenarios, thereby advancing its potential clinical application for enhancing cochlear implant procedures.

It is also important to mention, unlike other studies that use impedance recording built into the CI, we used a different method. Commercial CI only records impedance magnitude and several studies have used different impedance equivalent models and estimation algorithms (e.g., bivariate spline extrapolation [24]) to derive polarization and resistive impedance components. Studies have shown that this extra data help aid better electrode placement. Our method directly records impedance magnitude, phase, and resistive and reactive components that have the potential to be easily applicable to track insertion depth on run time during the insertion process.

To accurately track the insertion depth or trajectory of EA intraoperatively, it may be useful to devise a system that integrates different sensing modalities such as ECochG, monopolar/bipolar complex impedance recordings for machine learning model to predict better.

V. Conclusion

The paper concentrates on leveraging the complex bipolar electrical impedance of electrodes to monitor the linear insertion depth of Electrode Arrays (EA) within commercial cochlear implants. It employs two main methodologies: (1) classification, aimed at recognizing partial insertion depths, and (2) regression, intended to capture varying depths using recordings from electrode pairs. The study utilizes classical machine learning models and a hybrid Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) network for classification and regression, respectively. The research achieved a remarkable 98% classification accuracy using the Support Vector Machine (SVM) algorithm. Additionally, it introduced a CNN-LSTM hybrid regression model designed for precise tracking of EA insertion depths in millimeter increments, accomplishing a mean absolute error of 1.72 and an R-squared value of 0.85 during a 25mm insertion depth. Notably, this investigation utilized a 2D cochlear model. Real-time recording of complex impedance of electrodes may likely pave the way for objective feedback for surgeons when manually inserting the electrode array within the cochlea. This technique could also be handy as a sensing modality of a robotic system for the automated insertion of an electrode array that is designed to preserve residual hearing by reducing induced trauma.

Acknowledgments

We are thankful to Oticon Medical for providing electrode arrays for experimentation. We are also thankful to the Royal National Institute for Deaf People (RNID) and Edith Evelyn Wali Mohammed Trust for the funding support.

References

[1] B. J. Copeland and H. C. Pillsbury, "Cochlear implantation for the treatment of deafness," *Annu. Rev. Med.*, vol. 55, pp. 157–167, 2004.

[2] O. Adunka and J. Kiefer, "Impact of electrode insertion depth on intracochlear trauma," *Otolaryngology-head and neck surgery*, vol. 135, no. 3, pp. 374–382, 2006.

[3] S. Chakravorti, J. H. Noble, R. H. Gifford, B. M. Dawant, B. O'Connell, J. Wang, and R. F. Labadie, "Further evidence of the relationship between cochlear implant electrode positioning and hearing outcomes," *Otology & neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, vol. 40, no. 5, p. 617, 2019.

[4] M. W. Canfarotta, M. T. Dillon, K. D. Brown, H. C. Pillsbury, M. M. Dedmon, and B. P. O'Connell, "Insertion depth and cochlear implant speech recognition outcomes: A comparative study of 28-and 31.5-mm lateral wall arrays," *Otology & Neurotology*, vol. 43, no. 2, pp. 183–189, 2022.

[5] C. A. Buchman, M. T. Dillon, E. R. King, M. C. Adunka, O. F. Adunka, and H. C. Pillsbury, "Influence of cochlear implant insertion depth on performance: a prospective randomized trial," *Otology & Neurotology*, vol. 35, no. 10, pp. 1773–1779, 2014.

[6] K. Yukawa, L. Cohen, P. Blamey, B. Pyman, V. Tungvachirakul, and S. O'Leary, "Effects of insertion depth of cochlear implant electrodes upon speech perception," *Audiology and Neurotology*, vol. 9, no. 3, pp. 163–172, 2004.

[7] S. Appachi, S. Schwartz, S. Ishman, and S. Anne, "Utility of intraoperative imaging in cochlear implantation: A systematic review," *The Laryngoscope*, vol. 128, no. 8, pp. 1914–1921, 2018.

[8] R. Kloeckner, D. P. dos Santos, J. Schneider, L. Kara, C. Dueber, and M. B. Pitton, "Radiation exposure in ct-guided interventions," *European journal of radiology*, vol. 82, no. 12, pp. 2253–2257, 2013.

[9] M. M. Khan, R. F. Labadie, and J. H. Noble, "Preoperative prediction of angular insertion depth of lateral wall cochlear implant electrode arrays," *Journal of Medical Imaging*, vol. 7, no. 3, p. 031504, 2020.

[10] G. Mertens, V. Van Rompaey, P. Van de Heyning, E. Gorris, and V. Topsakal, "Prediction of the cochlear implant electrode insertion depth: clinical applicability of two analytical cochlear models," *Scientific reports*, vol. 10, no. 1, pp. 1–10, 2020.

[11] C. Rathgeb, M. Demattè, A. Yacoub, L. Anschuetz, F. Wagner, G. Mantokoudis, M. Caversaccio, and W. Wimmer, "Clinical applicability of a preoperative angular insertion depth prediction method for cochlear implantation," *Otology & neurotology*, vol. 40, no. 8, pp. 1011–1017, 2019.

[12] K. S. van der Marel, J. J. Briaire, R. Wolterbeek, B. M. Verbist, and J. H. Frijns, "Development of insertion models predicting cochlear implant electrode position," *Ear and hearing*, vol. 37, no. 4, pp. 473–482, 2016.

[13] S. Chen, F. Wang, Y. Lin, Q. Shi, and Y. Wang, "Ultrasound-guided needle insertion robotic system for percutaneous puncture," *International Journal of Computer Assisted Radiology and Surgery*, vol. 16, no. 3, pp. 475–484, 2021.

[14] J. Chevrie, N. Shahriari, M. Babel, A. Krupa, and S. Misra, "Flexible needle steering in moving biological tissue with motion compensation using ultrasound and force feedback," *IEEE Robotics and automation letters*, vol. 3, no. 3, pp. 2338–2345, 2018.

[15] T. D. Yang, K. Park, H.-J. Kim, N.-R. Im, B. Kim, T. Kim, S. Seo, J.-S. Lee, B.-M. Kim, Y. Choi *et al.*, "In vivo photothermal treatment with real-time monitoring by optical fiber-needle array," *Biomedical optics express*, vol. 8, no. 7, pp. 3482–3492, 2017.

[16] J. Pile, A. D. Sweeney, S. Kumar, N. Simaan, and G. B. Wanna, "Detection of modiolar proximity through bipolar impedance measurements," *The Laryngoscope*, vol. 127, no. 6, pp. 1413–1419, 2017.

[17] C.-T. Tan, M. Svirsky, A. Anwar, S. Kumar, B. Caessens, P. Carter, C. Treaba, and J. T. Roland Jr, "Real-time measurement of electrode impedance during intracochlear electrode insertion," *The Laryngoscope*, vol. 123, no. 4, pp. 1028–1032, 2013.

[18] L. Sijgers, A. Huber, S. Tabibi, J. Grosse, C. Röösli, P. Boyle, K. Koka, N. Dillier, F. Pfiffner, and A. Dalbert, "Predicting cochlear implant electrode placement using monopolar, three-point and four-point impedance measurements," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 8, pp. 2533–2544, 2022.

[19] C. K. Giardina, E. S. Krause, K. Koka, and D. C. Fitzpatrick, "Impedance measures during in vitro cochlear implantation predict array positioning," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 2, pp. 327–335, 2018.

[20] K. Koka, W. J. Riggs, R. Dwyer, J. T. Holder, J. H. Noble, B. M. Dawant, A. Ortmann, C. Valenzuela, J. K. Mattingly, M. Harris *et al.*, "Intra-cochlear electrocochleography during cochear implant electrode insertion is predictive of final scalar location," *Otology & neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, vol. 39, no. 8, p. e654, 2018.

[21] M. S. Harris, W. J. Riggs, C. K. Giardina, B. P. O'Connell, J. T. Holder, R. T. Dwyer, K. Koka, R. F. Labadie, D. C. Fitzpatrick, and O. F. Adunka, "Patterns seen during electrode insertion using intracochlear electrocochleography obtained directly through a cochlear implant," *Otology & neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, vol. 38, no. 10, p. 1415, 2017.

[22] C. K. Giardina, T. E. Khan, S. H. Pulver, O. F. Adunka, C. A. Buchman, K. D. Brown, H. C. Pillsbury, and D. C. Fitzpatrick, "Response changes during insertion of a cochlear implant using extracochlear electrocochleography," *Ear and hearing*, vol. 39, no. 6, p. 1146, 2018.

[23] S. Weder, C. Bester, A. Collins, C. Shaul, R. J. Briggs, and S. O'Leary, "Toward a better understanding of electrocochleography: analysis of real-time recordings," *Ear and hearing*, vol. 41, no. 6, pp. 1560–1567, 2020.

[24] P. Aebischer, S. Meyer, M. Caversaccio, and W. Wimmer, "Intraoperative impedance-based estimation of cochlear implant electrode array insertion depth," *IEEE transactions on biomedical engineering*, vol. 68, no. 2, pp. 545–555, 2020.

[25] S. Schraivogel, P. Aebischer, F. Wagner, S. Weder, G. Mantokoudis, M. Caversaccio, and W. Wimmer, "Postoperative impedance-based estimation of cochlear implant electrode insertion depth," *Ear and hearing*, vol. 44, no. 6, pp. 1379–1388, 2023.

[26] B. P. O'Connell, J. B. Hunter, D. S. Haynes, J. T. Holder, M. M. Dedmon, J. H. Noble, B. M. Dawant, and G. B. Wanna, "Insertion depth impacts speech perception and hearing preservation for lateral wall electrodes," *The Laryngoscope*, vol. 127, no. 10, pp. 2352–2357, 2017.

[27] N. Hafeez, X. Du, N. Boulgouris, P. Begg, R. Irving, C. Coulson, and G. Tourrel, "Electrical impedance guides electrode array in cochlear implantation using machine learning and robotic feeder," *Hearing research*, vol. 412, p. 108371, 2021.

[28] N. Hafeez, X. Du, N. Boulgouris, P. Begg, R. Irving, C. Coulsan, and G. Tourrels, "Towards unblinding the surgeons: Complex electrical impedance for electrode array insertion guidance in cochlear implantation," in *Proceedings of the International Symposium on Auditory and Audiological Research*, vol. 7, 2019, pp. 29–36.

[29] Y. N. Ertas, D. Ozpolat, S. N. Karasu, and N. Ashammakhi, "Recent advances in cochlear implant electrode array design parameters," *Micromachines*, vol. 13, no. 7, p. 1081, 2022.

[30] M. Tykocinski, L. T. Cohen, and R. S. Cowan, "Measurement and analysis of access resistance and polarization impedance in cochlear implant recipients," *Otology & Neurotology*, vol. 26, no. 5, pp. 948–956, 2005.

[31] V. N. Vapnik, "The nature of statistical learning theory," *New York: Springer2V erlag*, 1995.

[32] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*, 1992, pp. 144–152.

[33] P. Schiilkop, C. Burgest, and V. Vapnik, "Extracting support data for a given task," in *Proceedings of the 1st international conference on knowledge discovery & data mining*, 1995, pp. 252–257.

[34] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data mining and knowledge discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[35] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21–27, 1967.

[36] Z. Geler, V. Kurbalija, M. Radovanović, and M. Ivanović, "Comparison of different weighting schemes for the k nn classifier on time-series data," *Knowledge and Information Systems*, vol. 48, no. 2, pp. 331–378, 2016.

[37] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[38] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[39] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, "Deep learning for time series classification: a review," *Data mining and knowledge discovery*, vol. 33, no. 4, pp. 917–963, 2019.

[40] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1d convolutional neural networks and applications: A survey," *Mechanical systems and signal processing*, vol. 151, p. 107398, 2021.

[41] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.