MRI Safety Assessment of Implantable Medical Devices using Neural Networks

by

Qianlong Lan

A dissertation submitted to the Department of Electrical and Computer Engineering,
Cullen College of Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Electrical Engineering

Chair of Committee: Dr. Ji Chen

Committee Member: Dr. David Jackson

Committee Member: Dr. Jiefu Chen

Committee Member: Dr. Benhaddou Driss

Committee Member: Dr. Wolfgang Kainz

University of Houston
May 2021

ACKNOWLEDGEMENTS

In fall 2017, I enrolled in the electrical and computer engineering MS-PhD program at the University of Houston (UH). Pursuing the Ph.D. program at UH is such a challenging journey, and I am surrounded by all talented colleagues which inspires me a lot in the research area. I express my sincere appreciation to the following people: my advisor, Dr. Ji Chen, who is a strict tutor, and always pushed me to dive deep into research, and supports me so much in my research. Dr. Ji Chen is a professional scholar on applied electromagnetics with distinct industry experience and always provided me guidance and support.

I also would like to appreciate Dr. David R. Jackson, Dr. Jiefu Chen, and Dr. Stuart A. Long who are professors at the electrical and computer engineering department of UH. I did learn a lot from their lectures and always provide me guidance when I first joined the research group. Furthermore, Dr. Driss Benhaddou, a professor at computer engineering technology of UH, and Dr. Wolfgang Kainz who a research biomedical engineer at the U.S Food and Drug Administration (FDA) in the center for devices and radiological health. They also gave me very strong support in my research. I am also grateful to my colleagues, Dr. Jianfeng Zheng, Dr. Ran Guo, Jiajun Chang, Wei Hu, Shuo Song, and Yu Wang who are the teammates at UH, for the helps they provided in my research.

Finally, I especially appreciate my family's strong support for my study and research, which greatly strengthens my heart and provides me the impetus to move forward and become an expert in my research area.

Abstract

Magnetic resonance imaging (MRI) is one of the most effective and widely used noninvasive imaging techniques for disease diagnosis, due to its superior performance in soft tissue imaging without harmful ionizing radiation. However, the radiofrequency (RF)-induced heating which causes temperature rises and tissue burns, is a major hazard for patients with implantable medical devices to have an MRI. Recently, the RF-induced heating for passive implantable medical devices (PIMDs) has been carefully assessed to clarify the safety conditions for MRI examination in standard and fully controlled environments. However, these assessment needs costly measurements or numerical simulations which can take a relatively long time. Therefore, it is not applicable in providing the RF-induced heating of all available configurations for diverse configurations, for fast predicting the RF-induced heating in the design stage, or estimating the potential risks for patients with unlabeled implantable devices in emergency situations, etc. It is necessary to provide a fast prediction method of RF-induced heating in standard and fully controlled conditions or environments for different kinds of implantable medical devices.

Numerical modeling and simulations are conducted to study the RF-induced heating for general PIMDs, such as commonly used plate systems, and external fixators, in a 1.5 T or 3 T magnetic resonance (MR) environment. RF-induced heating for various configurations of the implantable medical devices in the phantom that covers possible clinical scenarios is investigated to be the ground truth data. Then, the neural networks (NNs) can be used as the surrogate model to train and predict the RF-induced heating against various configurations for different kinds of devices. To get accurate prediction performance, different architectures of NNs are applied to predict the RF-induced heating

of the implantable medical devices based on numerical or measure results. To validate the NNs, part of the ground truth data was used for training, while the rest were used to test the performance. Once the NNs had been trained, the possible hazard of the new implantable medical devices with predefined configurations would be clarified.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 Introduction

Implantable medical devices are widely used in clinical treatments to diagnose damaged biological tissues or structures. Typically, the general passive implantable medical devices (PIMDs) are composed of plate and serval screws, which could represent the widely used implantable devices for bone trauma and fracture treatments. Different configurations with variation in dimension could be applied to treat certain specific bone fractures based on the patient's conditions. For patients with severe bone fracture conditions, implantable devices with complex shapes are needed to support bone repairment.

Magnetic resonance imaging (MRI) is one of the state-of-the-art imaging techniques that have superior performance in disease diagnosis due to its benefits of non-invasion and capability to discriminate soft tissues. However, PIMDs in patients may lead to localized radiofrequency (RF) energy deposition in tissues during the MRI procedure. The metallic parts of these implantable medical devices will interact with the electromagnetic field from the RF coil. During such interactions, a surface current is induced on the devices to generate a scattering field to satisfy the boundary condition. Therefore, a strong total electromagnetic field could be occurred and induced heating if the current flows into the human tissue and concentrates at the tip of devices. Such RF-induced heating can be high enough to cause unintentional tissue damage and device malfunctions. Thus, RF-induced heating is a major hazard for patients with implantable medical devices during MRI procedures. It is necessary to fully evaluate the RF-induced heating for the implantable medical devices to clarify the safe conditions so that patients can proceed to the MRI examination.

Nowadays, both experiment measurement and numerical simulation approaches can be performed to evaluate the RF-induced heating [1-6]. For the PIMDs, the RF-induced heating is assessed in a standard experiment measurement or numerical simulation approach per the American Society for Testing and Materials (ASTM) standard F2182-11a [7]. For the experiment measurement approach, many efforts have been focused on the measurement of RF-induced heating inside phantom [1-3], since *in-vivo* RF-induced heating evaluation is very expensive and sometimes can only be performed in animals [8-10]. For the numerical simulation approach [4-6], anatomical human models or phantom models were used to assess the RF-induced heating. However, accurately full-wave modeling the submillimeter structure of the implantable medical devices is very challenging and time-consuming. Consequently, neither experiment measurement nor numerical simulation is efficient to evaluate the RF-induced heating for implantable medical devices.

To ensure the safety of patients, the safe conditions for some individual implantable devices have been carefully studied following these conditions [11-13]. However, due to inefficient experiment measurements or numerical simulations, it is difficult to use a simple reference studied case to represent the possible situations or to evaluate the RF-induced heating for each situation. For example, to fully investigate the RF-induced heating of orthopedic devices under different situations, various configurations of the devices with variation in dimensions were evaluated in the heterogeneous human body under different MRI environments [14]. The recent research [15] evaluates RF-induced heating for a typical cochlear implant system by investigating thousands of configurations with a different combination of factors, such as lead trajectory, lead type, human model, and MRI landmark. Therefore, it is not applicable in many scenarios, such as providing the MRI RF exposure of all available configurations for diverse device configurations, for fast predicting the MRI RF exposure in the design stage, or estimating the potential risks for patients with unlabeled implantable devices in emergency situations, etc. A commonly used multi-configuration implantable device can have thousands to millions of different configurations to cater to clinical requirements, thus brute-force one-by-one RF-induced heating assessment is too costly and almost infeasible. A complexity-reduced method was suggested in [16], but the method can't provide MRI RF exposure for each configuration and may underestimate the possible worst-case situation [17, 18].

Preliminary studies have indicated that neural networks (NNs) could be used to predict the MRI RF-induced heating for the simple passive implantable devices in an *in-vitro* way per the American Society for Testing and Materials (ASTM) F2182 standard [19, 20]. To parametrize the medical devices and obtain the ground-truth training data, the step size for a specific geometrical feature determines the training set. However, the criterion to determine the minimum training data set was not identified which is critical for fast prediction by using NNs. Moreover, the fine 3D spatial information of the complex shape implantable medical devices that are related to MRI RF-induced heating has not been taken into consideration. Some of the unnoted geometrical changes such as sharpness of screw-tip, kurtosis or skewness of plate-end, small changes on the surface, translation, and rotation, etc. play important roles in MRI RF-induced heating evaluation.

It is essential to fully validate the NNs approach which can provide RF-induced heating fast evaluation or prediction for the implantable medical devices by using minimum training data set. The RF-induced heating for patients who have been implanted

PIMDs could be underestimated if the devices were modeled by simplified geometrical structure or measured inside the homogeneous ASTM phantom. This is dangerous for patients implanted with these kinds of devices, thus should prevent the patients from MRI examination. This can also be alleviated if NNs can provide the feasibility of fast and accurate RF-induced heating evaluation for these PIMDs. Therefore, the applicability of NNs to fast evaluate the *in-vivo* RF-induced heating, even the possibility of fast prediction for complex geometrical shape implantable medical devices should be validated.

2 Problem Statement

With the widespread application of MRI, it is important to ensure the safety of patients due to RF-induced heating under different clinically relevant conditions. The factors or conditions, such as different kinds of devices with different configurations or components, different implantation positions, different imaging conditions, different coil settings, different stature, and different landmarks of patients will contribute to different levels of RF-induced heating. To conservatively evaluate the RF-induced heating these conditions need to be carefully studied by experimental measurement or numerical simulation based on standard procedures as shown in Figure 1. However, both experimental measurement and numerical simulation are very time-consuming or need high computational power to support the numerical calculation. Thus, the benefits of MRI are limited in many scenarios, such as providing the MRI RF-induced heating of all available configurations for diverse device configurations, fast predicting the MRI RF exposure in the design stage, or estimating the potential risks for patients with unlabeled implantable devices in emergencies, etc. NNs can work as a surrogate model to replace time-consuming experimental measurements or numerical simulations. The safety labeling of the implants (MR safe, MR conditional, or MR unsafe) can be fast decided by the predicted specific absorption rate (SAR) or heating from the NNs from a wide range of clinically relevant factors.

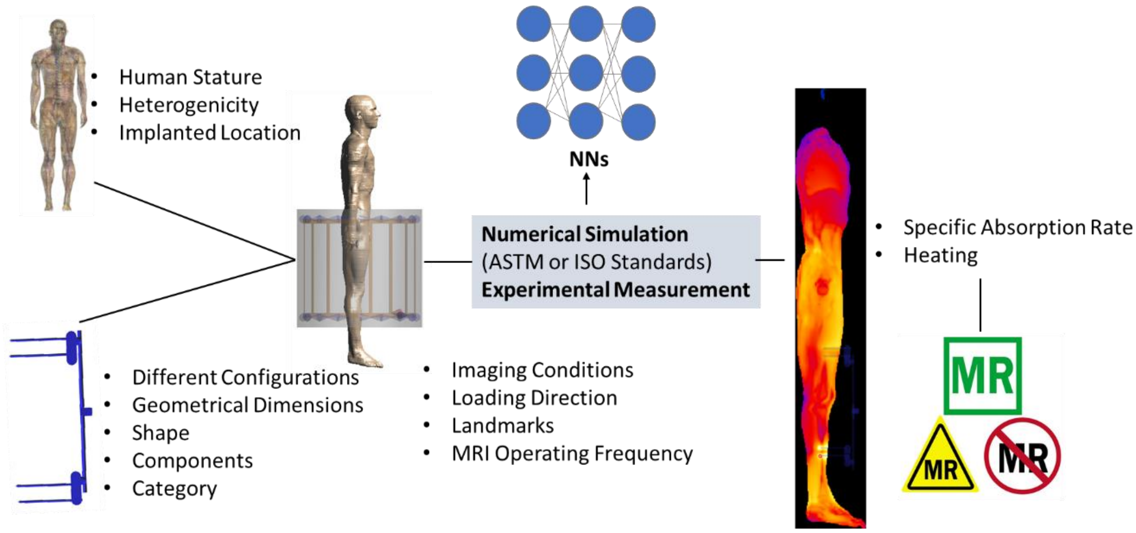


Figure 1. RF-induced heating related to various clinically relevant factors and evaluated by standard measurement or simulation.

Preliminary studies have indicated that neural networks (NNs) could be used to predict the MRI RF-induced heating for the simple passive implantable devices within predefined clinically relevant conditions [19, 20]. Numerical simulations are still needed to get the training data of the NNs for the multi-configuration devices. For a multi-configuration system with N parameters and M_n optional values for the nth parameter, the total number of configurations is $\prod_N M_n$. Thus, it's computationally expensive and practically prohibitive to get all the simulations. Furthermore, some optional values inside the parameters can be redundant for NNs. Therefore, to reduce the computational burden of numerical simulations and training time of the NNs by using the minimum number of configurations, the proper criterion to determine the training data should be proposed to address the issue.

Although the RF-induced heating for simple-shape implantable medical devices inside ASTM phantom has been used to validate the ANN, the complex shape devices are not considered in previous studies. Complex shape implantable medical devices cannot be parameterized by simple geometrical features (length, width, depth, etc.). Thus, it's not feasible to predict the RF-induced heating for complex shape implants using the NNs approach with simple geometrical features. The notable detail three-dimensional (3D) information will be lost if only using simple geometrical features which is important for accurate RF-induced heating evaluation. Many structures cannot be described using simple parameters such as shown in Figure 2.

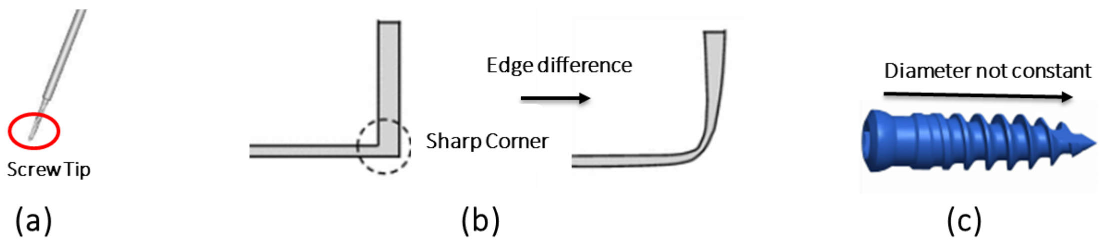


Figure 2. Examples of devices that cannot be parameterized (a) screw tip, (b) sharp corner, and (c) diameter not constant.

The edges and corners of the main structure of the complex shape medical devices also may contain slightly small variations. Based on the idea that electromagnetic simulations will use meshes in the simulations, we study the possibility of predicting the results of RF-induced heating use the device meshes with the CNN network.

Based on the analysis and instruction above, to fast evaluate the RF-induced heating for various implantable medical devices, the NNs approach still needs to be validated:

i. How to select the minimum number of configurations and which configurations should be selected for a given multi-configuration device?

ii. Can NNs adapt to the RF-induced heating fast prediction by using the minimum number of configurations? If so, how is the performance? What are the major factors affecting the *in-vivo* RF-induced heating?

iii. Whether the NNs can be used to fast evaluate or predict the complex shape of medical implantable devices?

3 Preliminary Literature Review

U.S. Food Drug Administration (FDA) recommended standard procedures in ASTM F2182-11a to evaluate the RF-induced heating of PIMDs [7]. The RF-induced heating among all possible configurations for the specific PIMDs was exhaustively studied by numerical simulation in [12, 16, 21]. These previous studies [22-28] have demonstrated that the RF-induced heating was related to multiple factors such as device geometrical dimension, patient orientation, landmark positions, etc. To reduce the computation complexity, Zheng. el [29] simplified the problem by searching for the worst-case situation by considering different combinations of configurations. However, it is still limited in the RF-induced heating efficient evaluation for each configuration under different situations. NNs have been widely applied in radiology as the modeling tool for medical diagnosis and have demonstrated superior outcomes. Automated fracture detection and classification using the neural network were achieved very high sensitivities [30, 31]. Automation of liver biometry across different imaging modalities and detection of myocardial delayed enhancement patterns can be facilitated by deep neural networks [32, 33]. Pulmonary nodule, urinary stone, anterior cruciate ligament tear within the knee joint, and acute ischemic large vessel occlusion stroke can be diagnosed and assessed by innovative models [34-37]. The ANN was first proposed in [19] to fast predict the MRI RF-induced heating for simple implantable plate devices and fully validated in [20]. The RF-induced heating of a generic stent with arbitrary orientation could also be fast predicted by using ANN [35]. However, the ASTM phantom environment is different from clinical situations, many researchers have studied the difference of RF-induced heating between the ASTM phantom and the human body [4, 7, 10]. Thus, the validity of ANN needs carefully studied or adapt to *in-vivo* RF-induced heating. Besides, the fast evaluation of RF-induced heating using ANN only studied the simple-shape PIMDs in-phantom environment. In clinical testing, multi-configuration devices may result in a very large number of possible device configurations and complex shape components [16].

Although the present literature already shows the potential applicability of NNs to facility the fast evaluation of RF-induced heating of simple-shape PIMDs. The proper criterion to determine the training data should be addressed to reduce the computational burden of numerical simulations and training time of the NNs by using the minimum number of configurations. The variation of the device itself and different combinations of configurations are contributed to the difference in RF-induced heating, and thus different safe conditions. Hence, the NNs must be improved to be capable of handling more complex cases, such as fast prediction or evaluation of RF-induced heating for complex shape implantable medical devices.

4 Objectives

Preliminary studies have shown that ANN was one feasible solution to provide a fast prediction of the MRI *in-vitro* RF-induced heating for simple PIMDs. However, the criterion to determine the minimum training data set was not identified. Moreover, the fine 3D spatial information related to RF-induced heating has not been taken into consideration, which is critical for standard RF-induced heating evaluation. Some of the small geometrical changes play important roles in MRI RF-induced heating evaluation.

In this study, numerical simulations were first performed for specific PIMD to get the ground truth data which can be used to validate the NN. To select the minimum number of configurations and reduce the computational burden of multi-configuration devices, a training data selection criterion was first determined. The determining criterion was validated by different types of implantable medical devices. For the fast evaluation or prediction of RF-induced heating for the complex shape implantable medical devices, the generic complex shape of multi-configurational devices was used to validate NN. With these thorough validations, the advantages of the NNs for RF-induced heating fast evaluation will be identified. Once the NNs were validated, the safe conditions for various kinds of implantable medical in different clinical scenarios can also be clarified and prevent patients from having potential undesired heating hazards undergoing MRI examination.

5 Mechanism Behind Neural Network and RF-induced Heating

5.1 Mechanism Behind Neural Network

The neural networks learning task needs to define a sample space which consists of input space $\boldsymbol{X}$ and output space $\boldsymbol{Y}$. For a sample $(\boldsymbol{x}, y) \in \boldsymbol{X} \times \boldsymbol{Y}$ in the sample space, suppose the relation between $\boldsymbol{x}$ and y can be described by an unknown real mapping function $y = g(\boldsymbol{x})$. The objective of the neural networks is to find a model that is similar or close to $g(\boldsymbol{x})$.

The actual form of the mapping function $g(\boldsymbol{x})$ is unknown. Therefore, we can only hypothesize a set of functions $\mathscr{F}$ based on experience, which is called hypothesis space. Then, select an ideal hypothesis $f^* \in \mathscr{F}$ by observing its characteristics on the training set D. The hypothesis space can be expressed as

$$\mathscr{F} = \left\{ f(x;\theta) | \theta \in \mathbb{R}^D \right\}, \tag{1}$$

where $f(\boldsymbol{x};\theta)$ is the function or model with a parameter of θ, D is the number of parameters. Generally, the hypothesis space can be classified into two categories, linear and non-linear hypothesis space. If the model is in linear hypothesis space,

$$f(x;\theta) = w^T x + b, \tag{2}$$

where θ is the parameter of the model, and the parameter θ contains the weight vector $\boldsymbol{w}$ and bias b. Furthermore, the model is commonly seen in a non-linear hypothesis space, and consist of non-linear basic functions $\phi(\boldsymbol{x})$,

$$f(x;\theta) = w^T \phi(x) + b, \tag{3}$$

where $\phi(\boldsymbol{x}) = [\phi_1(\boldsymbol{x}), \phi_2(\boldsymbol{x}), \ldots, \phi_L(\boldsymbol{x})]^T$ is a vector that consists of L non-linear basic functions. If $\phi(\boldsymbol{x})$ itself is a learnable basic function,

$$\phi_k(x) = \hbar(w_k^T \phi^{'}(x) + b_k), \forall 1 \le k \le K, \tag{4}$$

where $h(.)$ is non-linear function or activation function, $\phi'(\boldsymbol{x})$ is another basic function, $\boldsymbol{w}_k$ and b_k are learnable parameters, then $f(\boldsymbol{x};\theta)$ will be equivalent to a neural network. In this study, the rectified linear units (relu) non-linear function [38] was used.

In the case of multi-channel and multi-dimensional input data, a neural network can convert the composite function of each layer into the sum of several identical non-linear basic functions, i.e. the convolution kernels in a convolutional neural network [39].

An ideal neural network model $f(\boldsymbol{x};\theta)$ should get all the same $(\boldsymbol{x},y)$ as the real mapping function $y = g(\boldsymbol{x})$,

$$|f(x;\theta) - y| < \varepsilon, \forall (x,y) \in X \times Y, \tag{5}$$

where ϵ is a very small positive real value. To quantitively measure the difference between the predicted value $\hat{y} = f(\boldsymbol{x};\theta)$ and the actual value y, one common cost function, mean absolute percentage error (*MAPE*) could be used,

$$MAPE(y,\hat{y}) = \frac{1}{n}\sum_{i=1}^{n}(|\hat{y}_i - y_i| / y_i), \tag{6}$$

where n is the number of measured or predict data points in the hypothesis space $\mathcal{F}$.

It is an optimization problem to find the best model $f(\boldsymbol{x};\theta)$ as the training process of the neural network is to find the best model parameters θ^* to make the *MAPE* as small as possible,

$$\theta^* = \arg_\theta \min MAPE_{\mathcal{F}}(\theta). \tag{7}$$

A general neural network usually contains many hidden layers, different weights, and biases for each layer. Then, the training process of the neural networks becomes a non-convex optimization problem. The simplest and commonly used optimization algorithm for neural networks is gradient descent,

$$\theta_{t+1} = \theta_t - \alpha \frac{\partial MAPE_{\mathcal{F}}(\theta)}{\partial \theta}, \tag{8}$$

where θ_t is the parameter at t-th iteration, α is the searching step size or learning rate. In this study, the neural network was optimized with Adam's stochastic optimization algorithm [40].

5.2 Mechanism Behind RF-induced Heating

The RF heating of the medical implants under MRI environments is related to various parameters, such as RF coil type, subject type, subject loading position, implantation location, the geometry, and properties of medical implants, etc. In the field of electromagnetics, the energy dissipated in the lossy biological tissues is described via the SAR,

$$SAR(r) = \frac{\sigma}{2\rho}\left|\underline{E}^{t}(r)\right|^{2} = \frac{\sigma}{2\rho}\underline{E}^{t}(r)\underline{E}^{t*}(r), \tag{9}$$

where $E^{t}(r)$ is the total electric field, and σ, ρ is the conductivity and density of biological tissue respectively. Thus, the total SAR is related to the total electric field, as well as the conductive current density on the surface of the device.

The RF-induced heating near implantable devices can be simplified as a scattering problem of implanted structures buried in the lossy medium under specific sources. The incident field radiating from the RF coil penetrates the ASTM phantom or human body, interacts with the metallic devices. A surface current is induced on the device during such interactions to generate a scattering field that satisfies the boundary condition. The surface current density can be calculated as

$$\underline{J}_{s}(r) = \nabla \times \underline{H}(r) - j\omega\underline{D}(r), \tag{10}$$

where $\underline{D}(r)$ is the electric flux density and $\underline{H}(r)$ is the magnetic field. Due to the limit of boundary conditions, the current density flowing on the surface of the perfect electric conductor would be related to the tangential electric field or perpendicular magnetic field.

The current on the implants would flow into the human tissue at the tip of the device so that a strong total electric field could be observed. Due to this intense power flow, a hotspot would probably show up at the end of the device. Thus, the total electric field at the tip of the device could be separated from the vector incident electric field $\underline{E}^{i}(r_{tip})$ and scattering electric field $\underline{E}^{s}(r_{tip})$, which could be described as

$$\underline{E}^{t}(r_{tip}) = \underline{E}^{i}(r_{tip}) + \underline{E}^{s}(r_{tip}). \tag{11}$$

6 Parameterized NN for in-vitro Condition of Fully Implanted Devices

In this study, fully implanted devices that in in-vitro conditions are used as the study subjects. One simplest rod was created as a study example to determine the criterion for the training data set. Then three representative fully implanted devices were used to validate the criterion for the training data set. To obtain the ground-truth data, the RF-induced heating for these PIMDs was numerically investigated based on ASTM 2182 11-a inside the ASTM phantom. Numerical simulations were conducted in SEMCAD X (SPEAG, V14.8.6.1) software package based on the finite-difference time-domain (FDTD) method. The NNs architecture was constructed based on the validating cases which contain different inputs. At last, the NNs were trained by the training data set determined by the criterion, while the rest were used for fast evaluation of RF-induced heating and testing performance of the NNs.

6.1 Neural Network

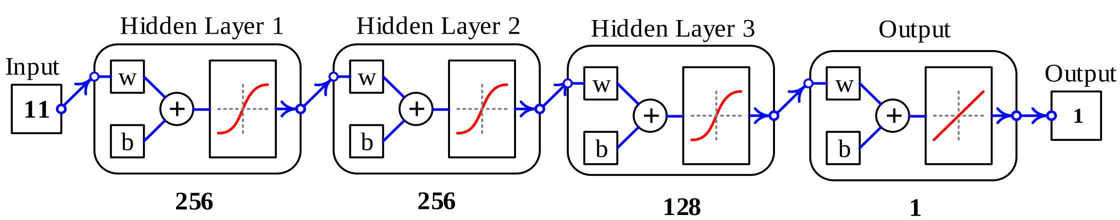


Figure 3. The architecture of the NN.

A simple three-layer feed-forward network was used to fit the SAR [19, 20], which had 3 hidden layers and 1 output layer as shown in Figure 3. The training set with different step sizes was used for training. The NN was used to predict the RF-induced heating for all cases and the worst-case. The number of inputs was 11 (random parameters) and the number of the output was 1 (the SAR value). Non-uniform numbers of hidden layers were used in this study. The first hidden layer was added with 256 neurons to capture the non-linear relationship between the input layer and the hidden layer in the high dimensional space. Then the number of neurons was decreased gradually to map the high dimensional features to lower-dimensional features between the hidden layers. Thus, the final output layer was able to learn the linear relationship from the low dimensional features in the last hidden layer.

6.2 The Criterion for Training Data Set

The simplest rod was created as a study example as shown in Figure 4. Rod length varies from 80 mm to 500 mm. The diameter was set to 3 mm. Numerical simulations were conducted at both 1.5 T and 3 T to get the RF-induced heating, in terms of peak 1/10 gram (g) averaged specific absorption rate ($SAR_{1/10g}$) in the ASTM phantom. The rod was placed at the vertical center on the right side 2 cm away from the phantom wall and at the center along the bore direction. All the results were normalized to a whole-body SAR of 2 W/kg.

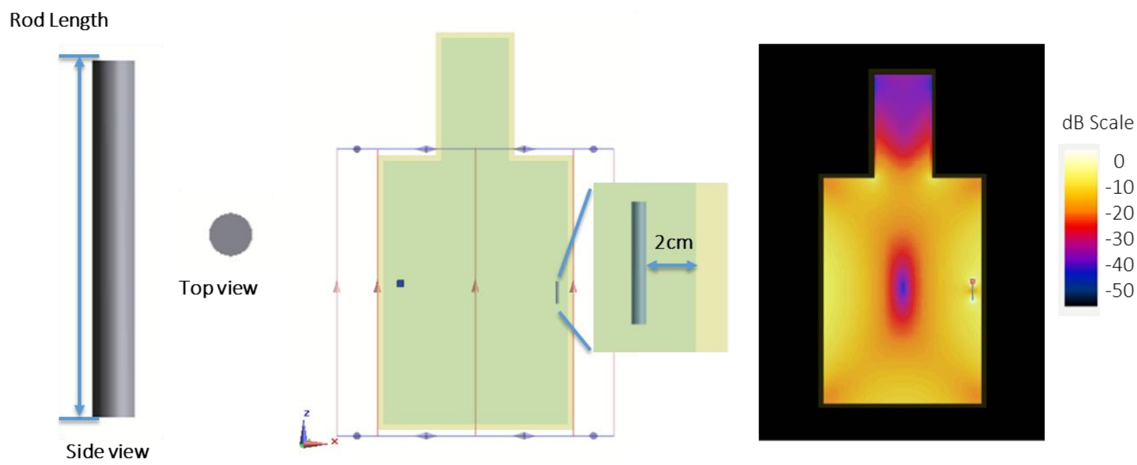


Figure 4. (a) rod, (b)simulation in phantom, (c) simulation results in 1.5T.

6.2.1 Training Set

A total number of 86 cases were studied as shown in Table 1. For each MRI operating frequency, the training set was selected as step size of 10 mm, 20 mm, 30 mm, 40 mm, 50 mm, and 60 mm. The worst-case rod length was not included in the training set. The worst-case rod length at 1.5 T is 190 mm with a SAR_{10g}=76.44 W/kg and SAR_{1g} = 268.71 W/kg. The worst-case rod length at 3 T is 100 mm with a SAR_{10g} = 33.35 W/kg and SAR_{1g} = 120.10 W/kg.

Table 1. Study of Rod Length.

Parameters	*Rod Length*	*Step Size*	No. of Values
1.5 T	[80~500] mm	10 mm	43
3T	[80~500] mm	10 mm	43

6.2.2 Results

The ANN results at a 1.5 T system with a different step size of rod length were shown in Figure 5 and Figure 6. The neural network was first trained with a small step size of 10 mm and the worst-case configuration was excluded from training. Then, the step size of rod length will be increased by 10 mm each time to validate the predicted results of the network. The Worst-case error (Ew) denotes the error between true worst-case SAR and predicted worst-case SAR. The overall error (Ea) denotes the mean error between all SAR and all predicted SAR. The prediction errors were small when the step size of rod length was less than 50 mm. However, the prediction errors of SAR_{10g} will larger than 4.41% when the step size of rod length was larger than 50 mm. Furthermore, the predicted worst-case SAR_{10g} will be lower than the true worst-case SAR_{10g} and the predicted worst-case error will larger than 3.26%. Although the SAR_{1g} was much higher than SAR_{10g} under the same frequency, the prediction errors were still small when the step size of rod length was smaller than 50 mm. Therefore, the step size used to select the training set of the neural network should be less than 50 mm in a 1.5 T system.

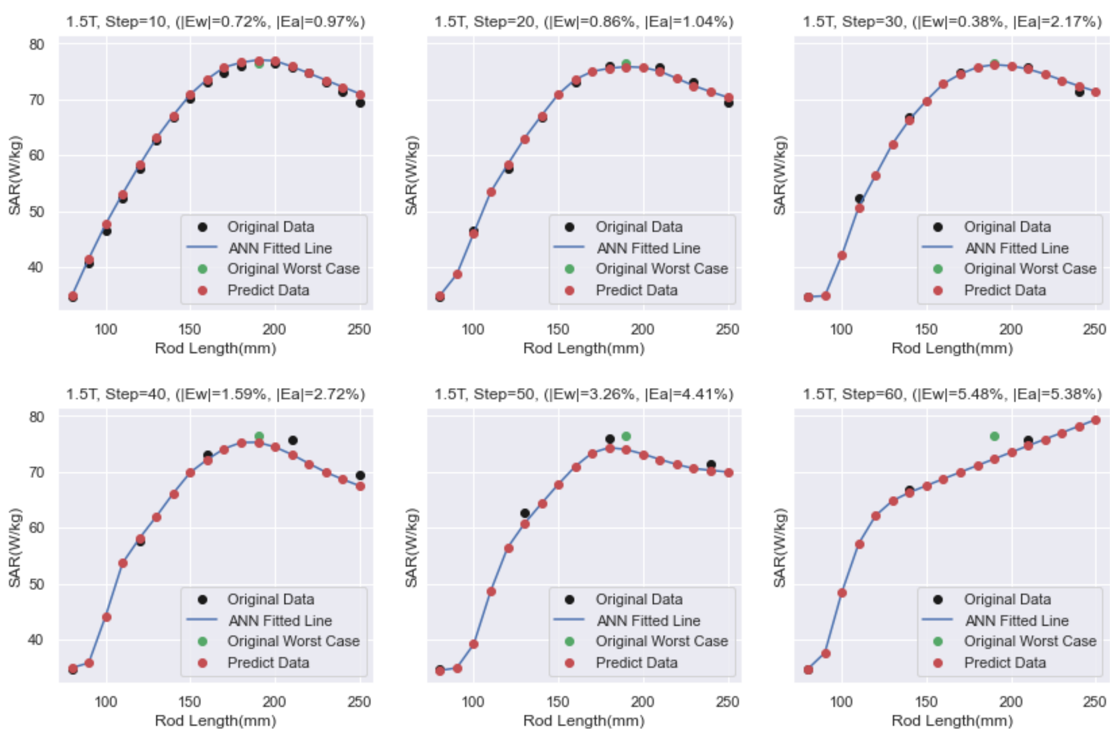


Figure 5. The ANN results (SAR_{10g}) at 1.5 T system with different step sizes of rod length.

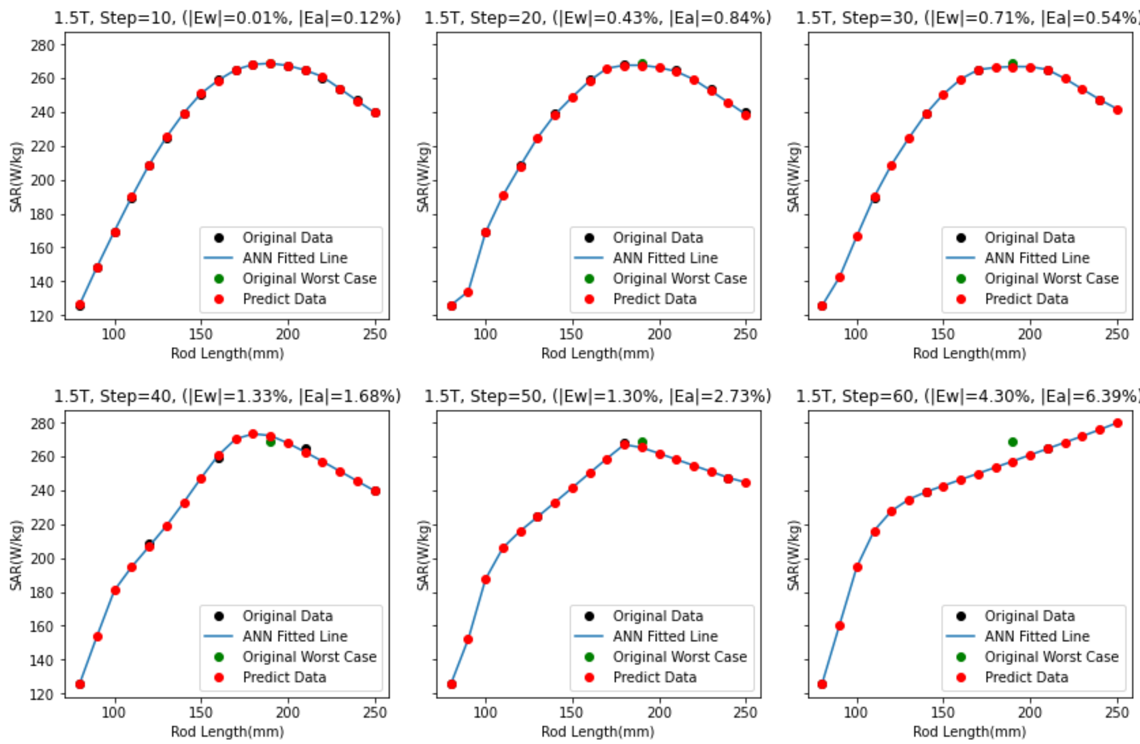


Figure 6. The ANN results (SAR_{1g}) at 1.5 T system with different step sizes of rod length.

The ANN results at 3 T system with a different step size of rod length were shown in Figure 7 and Figure 8. The prediction errors were small when the step size of rod length was less than 20 mm. It is more obvious for the worst-case SAR_{10g} prediction error which was larger than 10.54% when the network was trained using a step size of 30 mm. In the same circumstance, the worst-case SAR_{1g} prediction error was larger than 6.63%. In a 3 T system, the neural network needs more training samples with a smaller sampling step size compared to a 1.5 T system. The step size used to select the training set of the neural network should be less than 20 mm in a 3 T system.

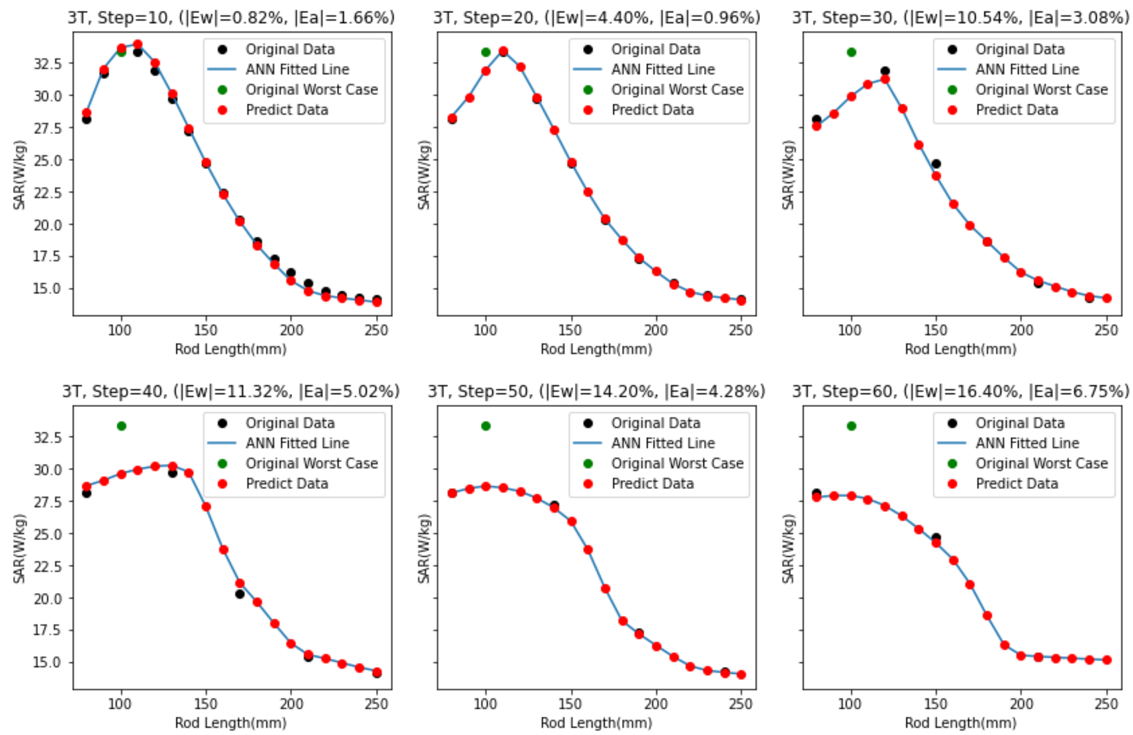


Figure 7. The ANN results (SAR_{10g}) at 3 T system with different step sizes of rod length.

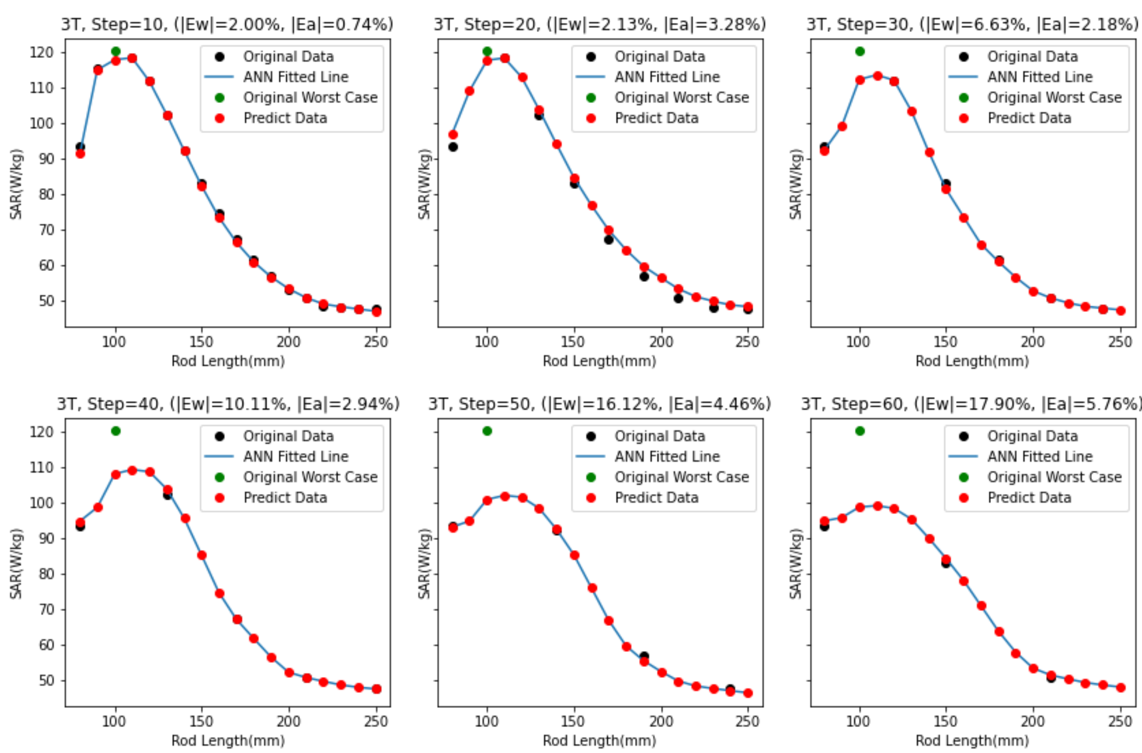


Figure 8. The ANN results (SAR_{1g}) at 3 T system with different step sizes of rod length.

Based on the study, to ensure the worst-case heating construct can be identified with ANN, it is suggested that one should use dimensional variations between training elements should be less than 1/10 of the wavelength. Overall, it should be around 24 mm. The wavelength in the lossy medium having both conductivity loss and polarization loss is calculated as

$$\begin{aligned} &\lambda = \frac{2\pi}{k'}, \\ &k' = \mathrm{Re}(k) = \mathrm{Re}(\omega\sqrt{\mu\varepsilon_c}), \\ &\varepsilon_c = \varepsilon_0\varepsilon_r - j\frac{\sigma}{\omega}, \end{aligned} \tag{12}$$

where λ is the wavelength, k' is the real part of the wavenumber in the medium, μ is the permeability of the medium, ϵ_c is the complex permittivity of the medium, and σ is the electrical conductivity of the medium. The medium used in the study is defined in Clause 8.2 of the ASTM F2182-11a [x], the plastic box was filled with gelled-saline, which was set to be ϵ_r= 80.38, σ = 0.47 S/m both at 1.5 T and 3 T. Thus, the wavelength at 1.5 T ($\lambda_{1.5T}$) is 432.20 mm and the wavelength at 3 T (λ_{3T}) is 243.93 mm.

6.3 Gain Analysis for the Neural Network

The gain of the neural network can be obtained if the criterion for the training set has been identified. In this study, the gain of the network was defined as the percentage of data saved in the global sample space by using the identified training data set criterion. Generally, the parameterized neural network can be used if the implantable medical device can be described by several parameters. The dimension of each parameter was bounded in a range $[d_{min}, d_{max}]$. The number of configurations (N_c) needed for numerical simulations for one parameter can be calculated as

$$N_c = \frac{(d_{\max} - d_{\min})}{(\lambda / 10)} \tag{13}$$

where d_{max} is the max dimension for the parameter of the device and d_{min} is the smallest dimension for the parameters of the device in clinical applications. λ denotes the wavelength at different operating frequencies under the MRI environment. The global sample space can be reduced to a smaller sample space if the device dimension was in a continuous sample space. For the device dimension that was in a discrete sample space, it also works if the number of values (v) for the parameter satisfied the condition for $v > 2$ and $v > N_c$. Otherwise, it's not necessary to use the identified training data set criterion as the global sample space has been very small. Assume that the implants were described by M parameters, and there were n ($n \leq M$) parameters that can be reduced to smaller sample space. Then the gain of the network (G_{NN}) can be defined as

$$G_{NN} = \left[1 - \frac{\prod_{i=1}^{n} \frac{(d_{\max(i)} - d_{\min(i)})}{\lambda / 10}}{\prod_{i=1}^{n} (d_{\max(i)} - d_{\min(i)})} \right] = \left[1 - \frac{u}{(\lambda / 10)^n} \right] \tag{14}$$

where i is the i-th parameters, u is the 1 unit of the dimension. In this study, $u = 1\ mm$, and the wavelength (λ) should be chosen based on the MRI operating frequency, for example, $\frac{\lambda_{1.5T}}{10} \approx 24\ mm$ at 1.5 T system and $\frac{\lambda_{3T}}{10} \approx 44\ mm$.

6.4 Application of Compression Implants

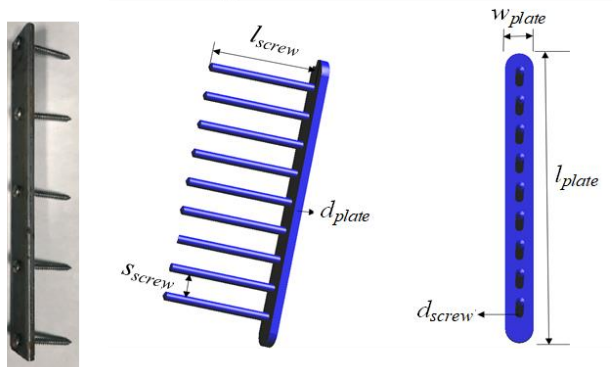


Figure 9. Parameters of the general compression plate.

The general compression plate system can be used as one study subject to validate the lambda/10 rule. This type of implant is commonly used to replace missing bone or to support fractured bones. The implants can be described by 6 parameters using plate length (l_{plate}), plate width(w_{plate}), plate depth(d_{plate}), the screw length(l_{screw}), screw diameter(d_{screw}), and screw spacing ($s_{spacing}$) as shown in Figure 9. The detailed device dimension was shown in Table 2. As we can see, the plate length can be in the range from 50 mm to 300 mm in the continuous space. The plate depth can be 3 mm or 5 mm, the plate width can be 10 mm or 20 mm, the screw spacing can be 10 mm or 20 mm, the screw length can be in the range from 20 to 100 mm in the continuous space.

Table 2. Dimension of General compression plate.

Parameters	Device Dimension(mm)
l_{plate}	[50-300]
d_{plate}	[3,5]
w_{plate}	[10,20]
s_{screw}	[10,20]
l_{screw}	[20-100]
d_{screw}	[3,5]

6.4.1 Numerical Simulations

Numerical simulations were performed to obtain the ground truth data. In this study, for the geometrical features, the minimum step size of lambda/10, and lambda/8 were adopted to get all the simulation cases as shown in Table 3. The total number of unique configurations at 3T are 1472 and 624 unique configurations at 1.5 T.

Table 3. Parameter dimensions used for simulations of compress plate.

Parameters	3 T		1.5 T	
	Dimension(mm)	No. of Values	Dimension(mm)	No. of Values
l_{plate}	[50,74,80,98,110,122, 140,146,170,194,200,218, 230,242,260,266,290,300]	18	[50,94,105,138,160,182,215,226, 270, 300]	10
d_{plate}	[3, 5]	2	[3, 5]	2
w_{plate}	[10, 20]	2	[10, 20]	2
s_{screw}	[10,20]	2	[10,20]	2
l_{screw}	[20,44,50,68,80,92,100]	7	[20,64,75,100]	4
d_{screw}	[3,5]	2	[3,5]	2

Numerical simulations were conducted both at 1.5 T and at 3 T using a full-wave electromagnetic solver based on the FDTD method to get the RF-induced heating in the ASTM phantom. Two high passes non-physical RF transmits body coils were adopted to model the MRI RF body coil operated at 64 MHz and 128 MHz respectively. The RF coil was loaded with a model of the ASTM phantom. Eight current sources were placed on the rungs of the coil to generate a uniform magnetic field inside the coil. Absorbing boundary conditions were used on all sides of the simulation boundaries.

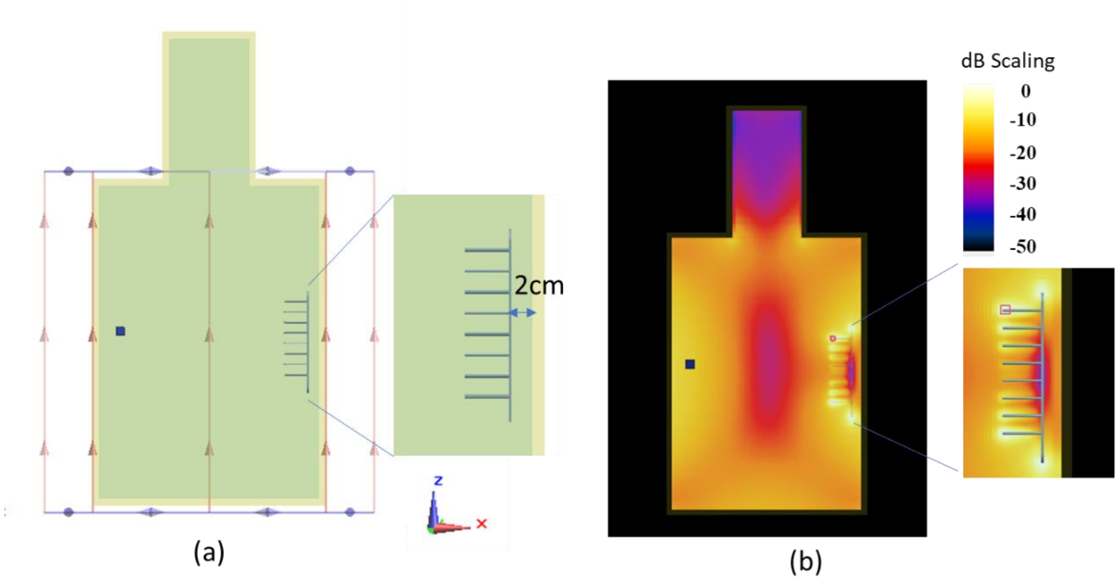


Figure 10. (a) The device was placed at the vertical center on the right side 2cm away from the phantom wall and at the center along the bore direction. (b) Example of RF-induced heating under 3 T system.

The ASTM phantom was a plastic container with a relative dielectric constant $\epsilon_r = 3.7$ and an electrical conductivity $\sigma = 0$ S/m. As defined in Clause 8.2 of the ASTM F2182-11a, the plastic box was filled with gelled-saline, which was set to be $\epsilon_r = 80.38$ and $\sigma = 0.47$ S/m at both 64 MHz and 128 MHz. The plate device was placed at the vertical center on the right side 2 cm away from the phantom wall and at the center along the bore direction (the location which provides maximum and uniform electric field-induced heating inside the phantom) as shown in Figure 10. An example of the RF-induced heating of the compression plate device under the 3 T system showed that the hot spot occurred at the end of the screw. In the numerical simulation, all metallic materials were modeled as perfect electric conductor (PEC).

The non-uniform mesh was used in the simulations to approach the balance between accuracy and complexity because the size of the coil, phantom, and devices was different. The larger mesh step size can reduce the total simulation time, but the coarse mesh cannot represent the device structure. The smaller mesh steps unbearable computational burdens and the divergent results. It was determined with several convergence analyses that the mesh size of 1 mm was applied to the plate devices. The mesh size of the gelled saline was 5 mm, and the plastic box was 10 mm. The grating ratio of the mesh size was set to 1.15. To ensure convergence, the simulation time was set for 25 periods. The convergence of the numerical simulation is quantified with a Convergence Level (CL) which is based on the difference between the last two estimations in the frequency domain [41]. Lower CL leads to more accurate estimation but requires longer simulation time. All the numerical simulations reached a CL of -35 dB at 1.5 T and -50 dB at 3 T as shown in Figure 11. All the results were normalized to a whole-body SAR of 2 W/kg.

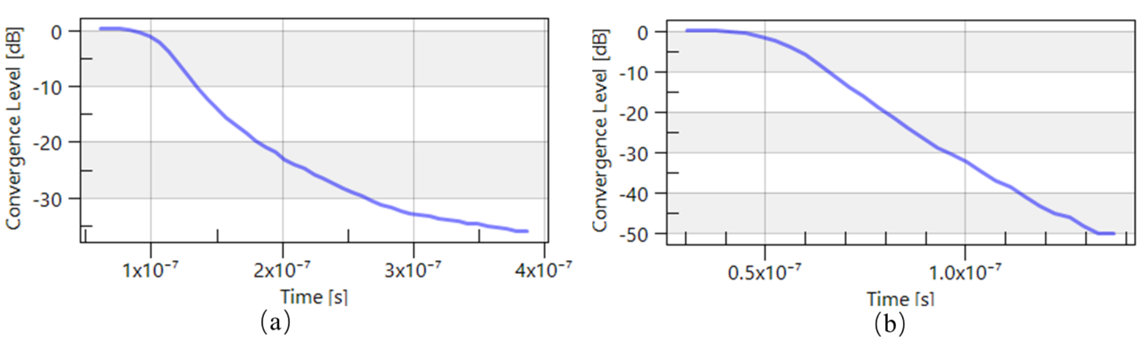


Figure 11. Numerical Simulations Convergence Analysis at (a) 1.5 T and (b) 3T with 25 periods for Compression Plate.

6.4.2 Training Set

The study parameters with different step sizes at the 1.5 T and 3 T system were shown in Table 4. These training sets with different step sizes will be used to train the NN. Once it has been trained, it can be validated using the test data which was the whole data set from simulation to predict the RF-induced heating.

Table 4. The training set for general compression plate.

Parameters	3 T			1.5 T		
	$\lambda_{3T}/10$ (mm)	$\lambda_{3T}/8$ (mm)	$\lambda_{3T}/5$ (mm)	$\lambda_{1.5T}/10$ (mm)	$\lambda_{1.5T}/8$ (mm)	$\lambda_{3T}/5$ (mm)
Plate Length	[50,74,98,122,146,170,194,218,242,266,290,300]	[50,80,110,140,170,200,230,260,290,300]	50,98,146,194,242,290,300]	[50,94,138,182,226,270, 300]	[50,105,160,215,270,300]	50,138,226,300]
Plate Depth	[3, 5]	[3, 5]	[3, 5]	[3, 5]	[3, 5]	[3, 5]
Plate Width	[10, 20]	[10, 20]	[10, 20]	[10, 20]	[10, 20]	[10, 20]
Screw Spacing	[10,20]	[10, 20]	[10, 20]	[10,20]	[10, 20]	[10, 20]
Screw Length	[20,44,68,92,100]	[20,50,80,100]	[20,68,100]	[20,64,100]	[20,75 ,100]	[20,100]
Screw Diameter	[3,5]	[3,5]	[3,5]	[3,5]	[3,5]	[3,5]

6.4.3 Results

The simulation results of the compression plate system were shown in Figure 12. It has a mean SAR_{10g} of 16.09 W/kg at 3 T system and a mean SAR_{10g} of 30.50 W/kg at 1.5 T system. The RF-induced heating in terms of SAR_{1g} also was evaluated. The results have shown that it has a mean SAR_{1g} of 44.49 W/kg at 3 T system and a mean SAR_{1g} of 83.73 W/kg at 1.5 T system. The statistics of all the simulation results for the compression implants were shown in Table 5.

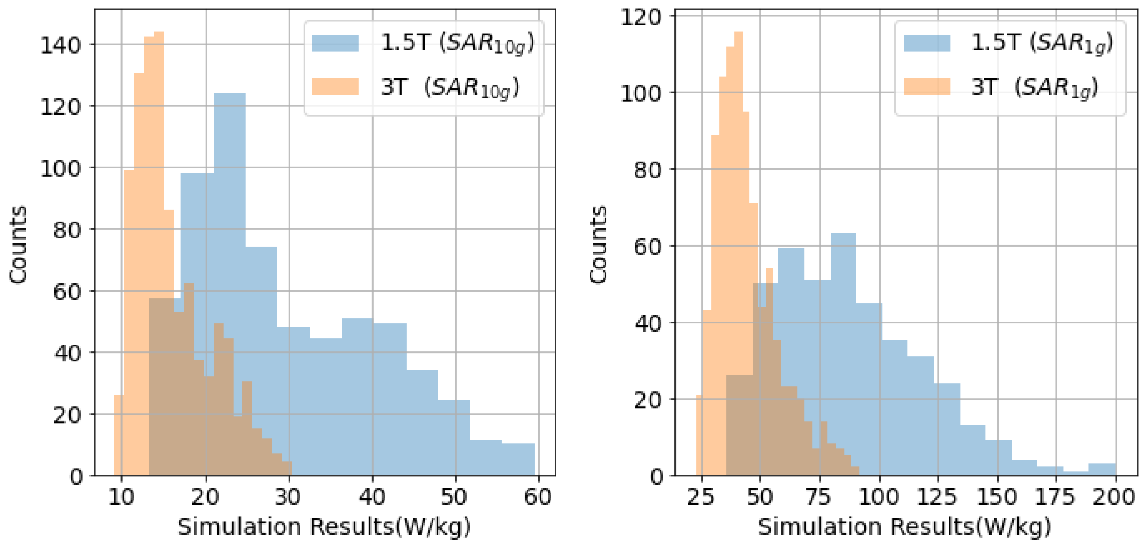
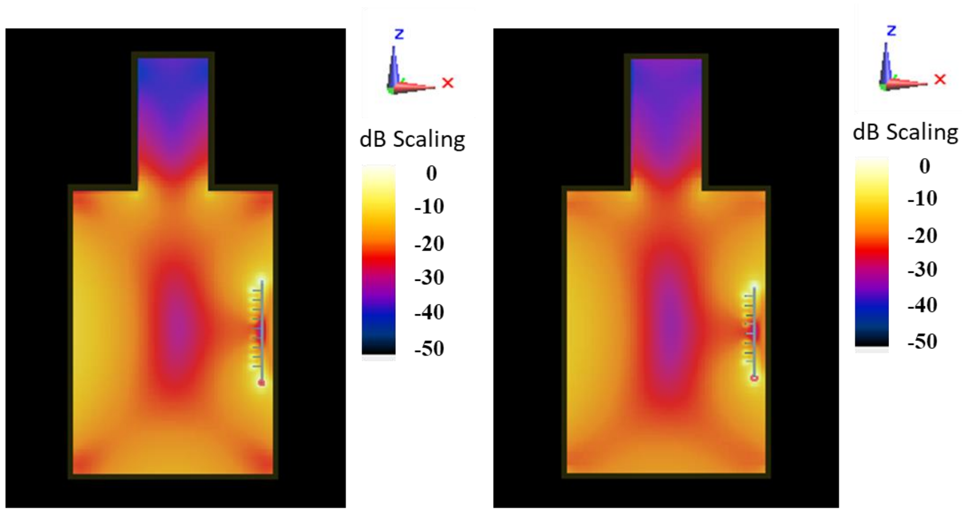


Figure 12. Simulation results for the compression plate system.

The worst-case SAR_{10g} is 30.45 W/kg and the worst-case SAR_{1g} is 91.85 W/kg for the compression plate system at 3 T system. The corresponding worst-case configuration at the 3 T system has a plate length of 98 mm, a plate width of 10 mm, a plate depth of 3 mm, a screw length of 20 mm, a screw diameter of 3 mm, and a screw spacing of 20 mm. The worst-case SAR_{10g} is 66.93 W/kg and the worst-case SAR_{1g} is 200.21 W/kg for the compression plate system at 1.5 T system. In this case, the corresponding worst-case configuration at 1.5 system has a plate length of 194 mm, a plate width of 10 mm, a plate depth of 3 mm, a screw length of 20 mm, a screw diameter of 3 mm, and a screw spacing of 20 mm. The example of worst-case RF-induced heating for the compression plate system was shown in Figure 13. The hot spots usually occurred at the end of the plate.

Table 5. Statistics of the simulation results for the compression implants.

Statistics	SAR_{10g}		SAR_{1g}	
	3 T	**1.5 T**	**3 T**	**1.5 T**
Min	9.19 (W/kg)	13.27 (W/kg)	22.79 (W/kg)	35.86 (W/kg)
Max	30.45 (W/kg)	66.93 (W/kg)	91.85 (W/kg)	200.21 (W/kg)
Mean	16.09 (W/kg)	30.50 (W/kg)	44.49 (W/kg)	83.73(W/kg)
Variance	20.51 (W^2/kg^2)	129.03 (W^2/kg^2)	173.09(W^2/kg^2)	907.11 (W^2/kg^2)



(a) (b)

Figure 13. Example of worst-case RF-induced heating of compression plate system at (a) 1.5 T and (b) 3 T.

The ANN results (SAR_{10g}) at 3 T system with a different step size of compression plate system were shown in Figure 14. The network was first trained by the minimum step size of lambda/10. The results indicated that the correlation coefficient of the ANN was larger than 0.90 and the mean absolute percentage error (MAPE) was close to 5.10%. The worst-case prediction error was less than 0.60 %. The ANN has learned the non-linear relationship between the parameterized features and the RF-induced heating by using the minimum step size of lambda/10. However, the correlation coefficient of the ANN was less than 0.80 and the worst-case error was larger than 12.56% when the ANN was trained by the minimum step size of lambda/8. The network will not learn the non-linear relationship if trained by using the step size of lambda/5 as the correlation coefficient was less than 0.51 and the MAPE was as high as 21.00%.

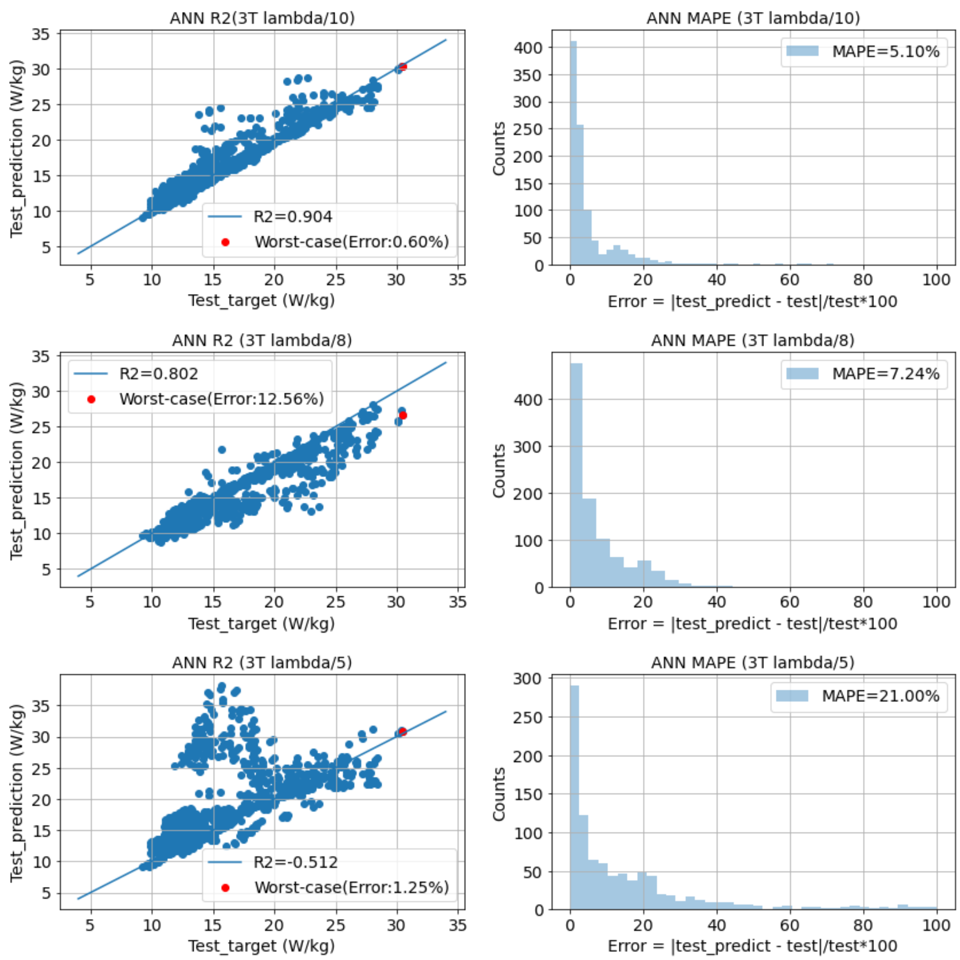


Figure 14. The NN testing results (SAR_{10g}) were trained by different step sizes for compression plate system at 3 T system.

Although the RF-induced heating in terms of SAR_{1g} was much higher than SAR_{10g}, the ANN still can be used to predict the worst-case and the overall heating using a minimum step size of lambda/10 as shown in Figure 15. The worst-case error was less than 1.66%

and the MAPE was less than 4.76% at the 3 T system when the network was trained by the minimum step size of lambda/10. The performance of the network will be low as the worst-case error will be as high as 14.31% when using a step size of lambda/8. Thus, it's recommended to use a minimum step size lambda/10 for the compression plate fast prediction using ANN at the 3T system.

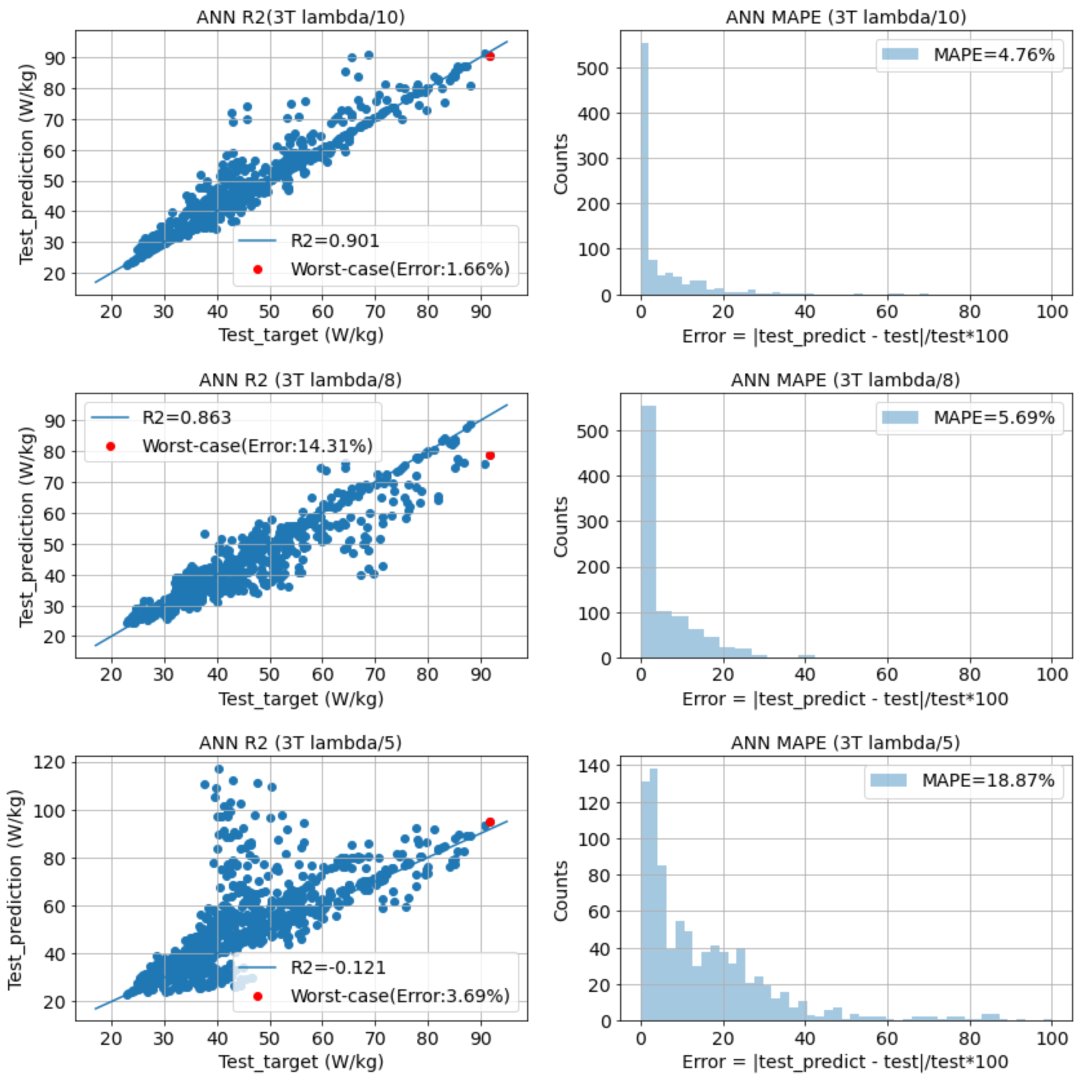


Figure 15. The NN testing results (SAR_{1g}) trained by different step sizes for compression plate system at 3 T system.

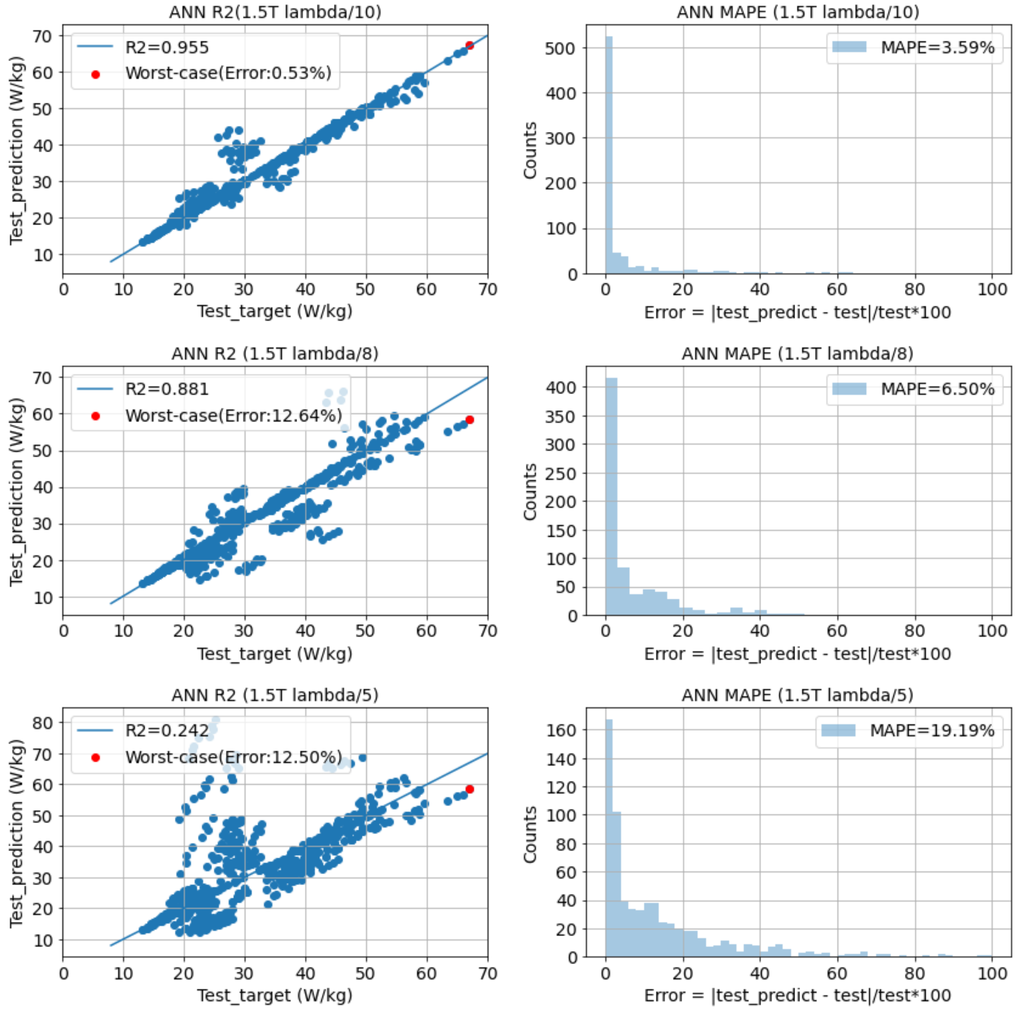


Figure 16. The NN testing results (SAR_{10g}) were trained by different step sizes for compression plate system at 1.5 T system.

The ANN results (SAR_{10g}) at 1.5 T system with different step sizes of compression plate system were shown in Figure 16. The network can predict the RF-induced heating with a small error rate when it was trained by the step size of lambda/10. The correlation coefficient of the network was larger than 0.95 and the MAPE was less than 3.59%. The network can predict the worst-case SAR_{10g} with an error that was less than 0.53%. The worst-case error was much larger compared to a minimum step size of 10/lambda when trained by the minimum step size of lambda/8. The performance of the network will not be guaranteed if trained by the step size of lambda/5. The correlation coefficient of the network was very small. The prediction error became very large which means the network was not converged. The ANN results (SAR_{1g}) at 1.5 T system with different step sizes of compression plate system were shown in Figure 17. The results have shown that the ANN can be used to predict the SAR_{1g} at 1.5 T system if the network was trained by a minimum step size of lambda/10. The worst-case prediction error was less than 4.81% and the MAPE was less than 3.92%. However, the worst-case prediction error will be larger than 15.08% if the network was trained by a minimum step size larger than lambda/8. Therefore, it's also recommended to use a minimum step size of lambda/10 for compression plate fast prediction at the 1.5 T system.

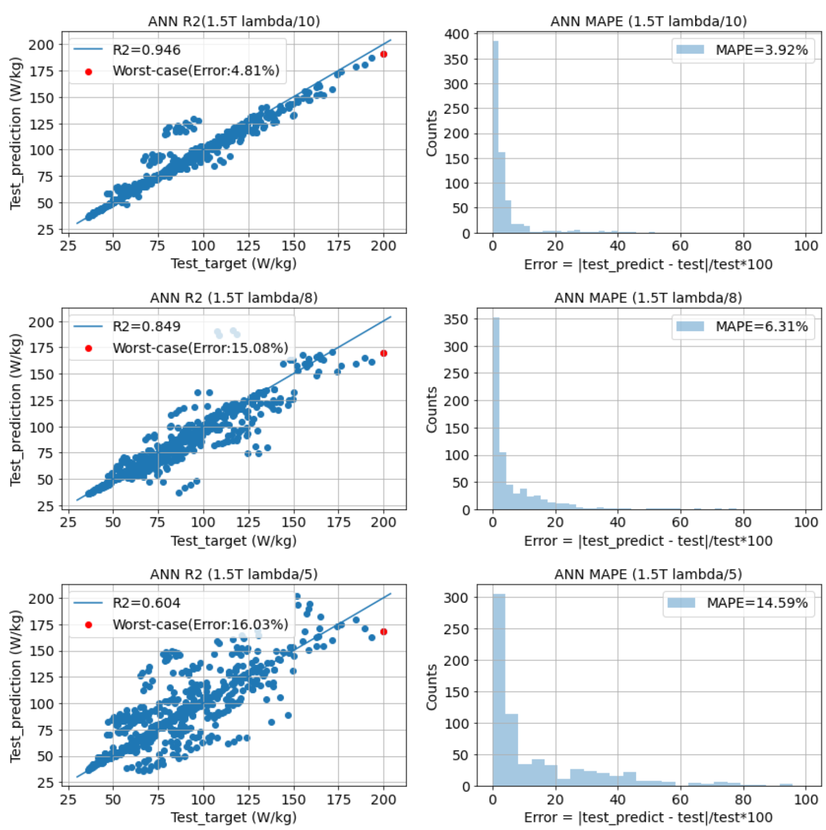


Figure 17. The NN testing results (SAR_{1g}) trained by different step sizes for compression plate system at 1.5 T system.

For the device parameters in the continuous sample space, the parameters were swept in fine step size. The other parameters remained unchanged when a selected parameter was swept. The training data set used a converged step size of lambda/10 and the errors between the prediction results and the simulation results both at the 1.5 T and 3 T system were shown in Figure 18. The largest predicted error for the plate length was less than 11% and the largest predicted error for the screw length was less than 14% at the 3 T system. The largest predicted error for the plate length was less than 9% at the 1.5 T system, however, the largest error for the screw length was close to 13%. This indicated that the RF-induced heating studied by parameters will not follow a simple linear relationship for the compression plate system.

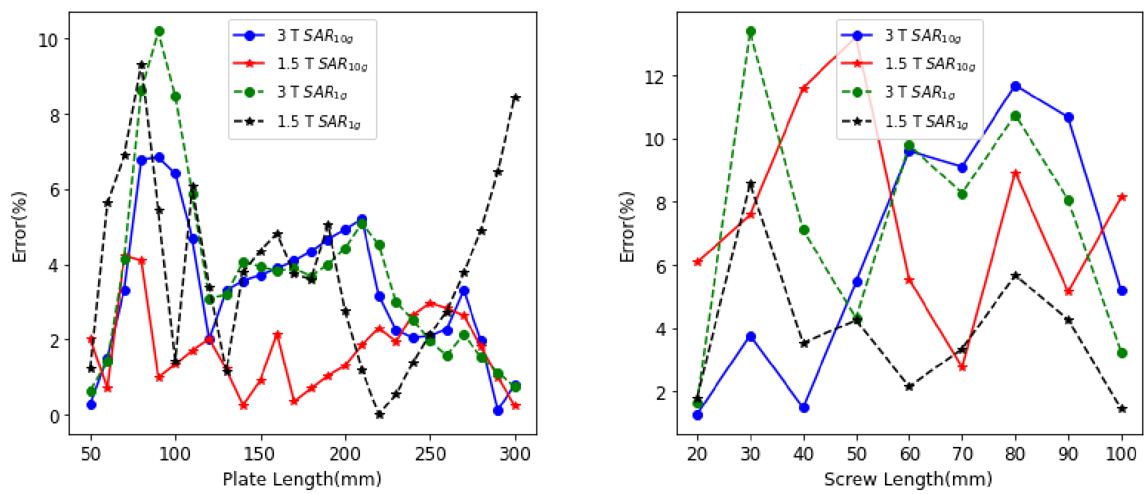


Figure 18. NN prediction error for compression plate system with a fine step of 10 mm.

For the compression plate system, there are two parameters used for the ANN that can be reduced to smaller sample space. One of the parameters is the plate length. Another parameter is the screw length. It's recommended to use a converge step size of lambda/10 to get more accurate prediction results. Therefore, the gain of the network (G_{NN}) for the study of compression plate system at 1.5 T system is $G_{NN(1.5T)} = \left[1 - \frac{1}{(24)^2}\right] \approx 99.82\%$. The gain of the network (G_{NN}) for the study of compression plate system at 3 T system is $G_{NN(3T)} = \left[1 - \frac{1}{(44)^2}\right] \approx 99.94\%$.

6.5 Application of Cervical Plate System

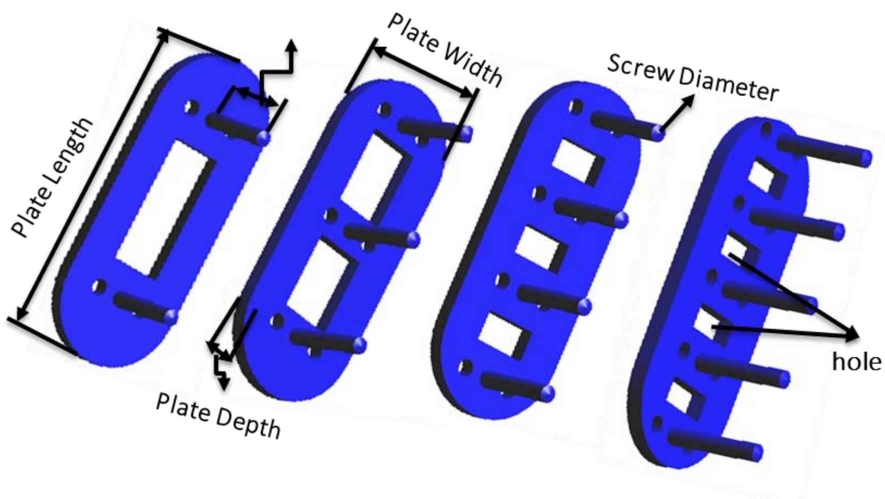


Figure 19. Cervical Plate System Parameters.

The cervical plate system is commonly used for the treatment of cervical degenerative diseases, trauma, and tumor. The device can be parameterized by plate length, screw length, screw spacing, and screw diameter as shown in Figure 19. The dimension of the cervical plate system was shown in Table 6, the plate length was in the range from 20 mm to 130 mm in the continuous space. The plate width can be 10 mm or 20 mm, screw length can be 3 mm or 5 mm, screw diameter can be 5 mm or 10 mm, the screw spacing can be 20 mm to 135 mm in the continuous space.

Table 6. The device dimension of the cervical plate system.

Parameters	Device Dimension(mm)
Plate Length	[30-140]
Plate Width	[18,25]
Plate Depth	[3,5]
Screw Length	[10,20]
Screw Diameter	[3, 5]
Number of Holes	[1,2,3,4]

6.5.1 Numerical Simulations

Numerical simulations were performed to obtain the ground truth data. In this study, for the geometrical features, the minimum step size of lambda/10, and lambda/8 were adopted to get all the simulation cases as shown in Table 7. The total number of unique configurations at 3T is 528 and 320 unique configurations at 1.5T.

Table 7. Parameter dimensions used for simulations of cervical plate system.

Parameters	3 T		1.5T	
	Dimension (mm)	No. of Values	Dimension (mm)	No. of Values
Plate Length	[30, 54, 60,78, 90, 102, 120, 126,140]	9	[30, 74, 85, 118,140]	5
Plate Width	[18,25]	2	[18,25]	2
Plate Depth	[3,5]	2	[3,5]	2
Screw Length	[10,20]	2	[10,20]	2
Screw Diameter	[3, 5]	2	[3, 5]	2
Number of Holes	[1,2,3,4]	4	[1,2,3,4]	4

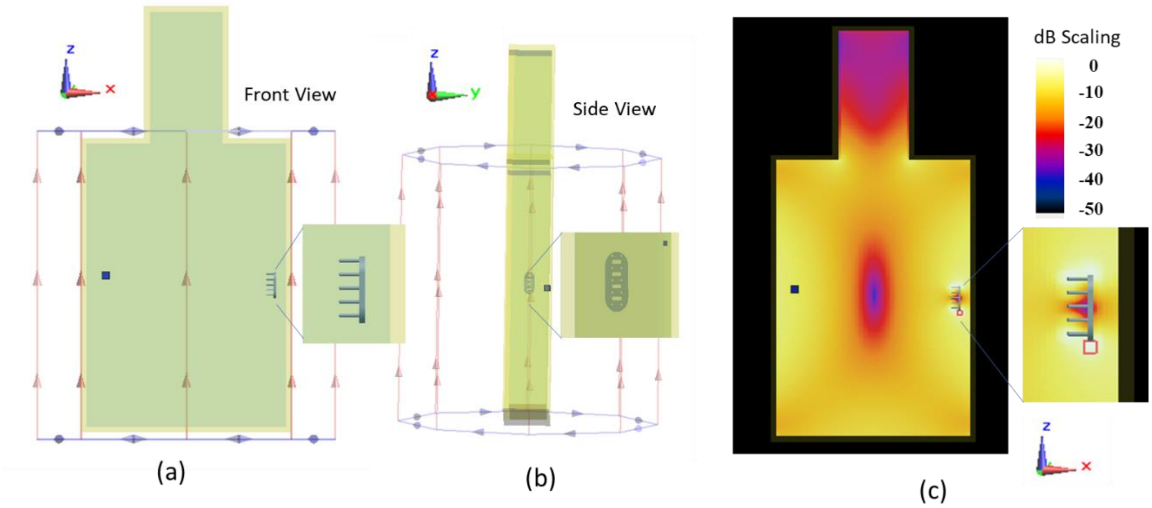


Figure 20. Numerical Simulation of the Cervical Plate Conducted at 3T: (a) Front view, (b) Side View (c) Example of the RF-induced Heating.

Numerical simulations were also performed both at 1.5 T and 3 T system based on the FDTD method to get the RF-induced heating in the ASTM phantom. The high pass non-physical RF transmits body coil and the ASTM phantom were the same as the study of compression plate system. The cervical plate device was placed at the vertical center on the right side 2 cm away from the phantom wall and the center along the bore direction as shown in Figure 20. An example of the RF-induced heating of the complex shape plate device under the 3 T system showed that the hot spot occurred at the end of the plate. In the numerical simulation, all metallic materials were modeled as perfect electric conductor (PEC).

The non-uniform mesh was used in the simulations. A mesh size of 1 mm was applied to the plate devices. The mesh size of the gelled-saline was 5 mm and the plastic box was 10 mm. The grating ratio of the mesh size was set to 1.15. To ensure convergence, the simulation time was set for 25 periods. All the numerical simulations reached a CL of -50 dB at 3 T and -30 dB at 1.5 T. All the results were normalized to a whole-body SAR of 2 W/kg.

6.5.2 Training Set

The study parameters with different step sizes at the 3 T system was shown in Table 8. For the plate length in the continuous space, the training data points were chosen based on the step size. Once the network has been trained, it can be used to predict the device dimension shown in Table 6.

Table 8. The training set was used for the cervical plate system at the 3 T system.

Parameters	3 T			1.5 T		
	$\lambda_{3T}/10$ (mm)	$\lambda_{3T}/8$ (mm)	$\lambda_{3T}/5$ (mm)	$\lambda_{1.5T}/10$ (mm)	$\lambda_{1.5T}/8$ (mm)	$\lambda_{1.5T}/5$ (mm)
Plate Length	[30, 54, 78, 102, 126, 140]	[30,60,90,120,140]	[30, 78, 102, 140]	[30, 74, 118, 140]	[30,85,140]	[30,118, 140]
Plate Width	[18,25]	[18,25]	[18,25]	[18,25]	[18,25]	[18,25]
Plate Depth	[3,5]	[3,5]	[3,5]	[3,5]	[3,5]	[3,5]
Screw Length	[10,20]	[10,20]	[10,20]	[10,20]	[10,20]	[10,20]
Screw Diameter	[3, 5]	[3, 5]	[3, 5]	[3, 5]	[3, 5]	[3, 5]
Number of Holes	[1,2,3,4]	[1,2,3,4]	[1,2,3,4]	[1,2,3,4]	[1,2,3,4]	[1,2,3,4]

6.5.3 Results

The simulation results of the cervical plate system were shown in Figure 21. The simulation results indicated that the mean SAR_{10g} at the 3 T system is 17.29 W/kg and the mean SAR_{10g} at the 1.5 T system is 33.21 W/kg. The mean value of RF-induced heating in terms of SAR_{1g} at the 3 T system is 42.43 W/kg and the mean SAR_{1g} at the 1.5 T system is 80.27 W/kg.

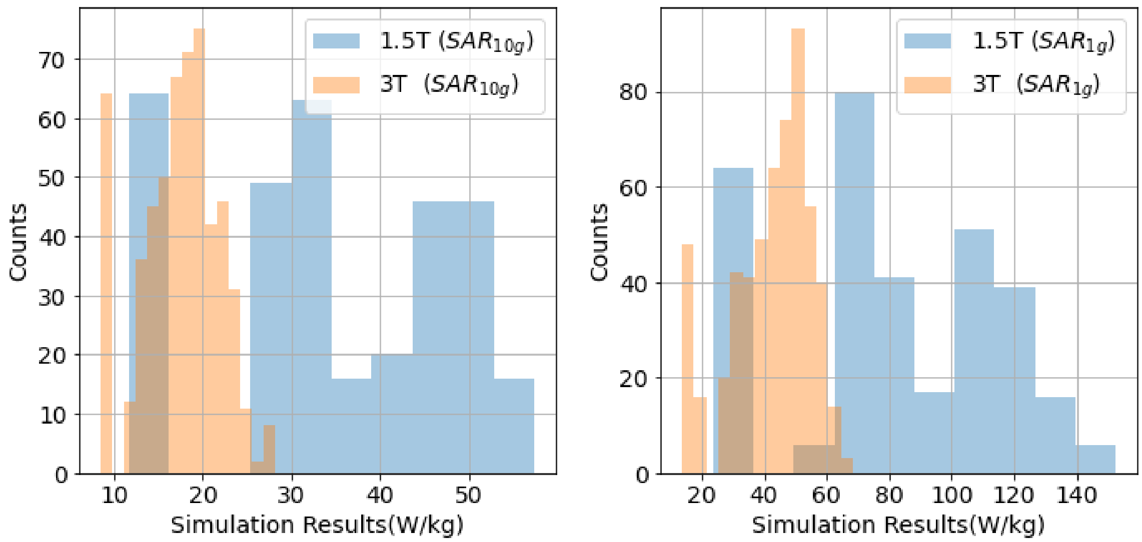


Figure 21. Simulation results for the cervical plate system.

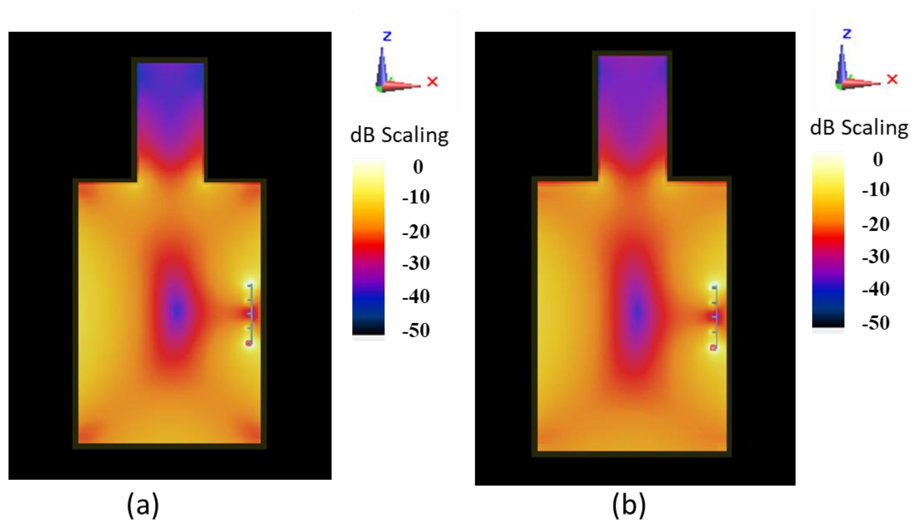


Figure 22. Illustration of worst-case RF-induced heating of cervical plate system at (a) 1.5 T system and (b) 3 T system.

The statistics of the simulation results for the cervical plate system were shown in Table 9. The worst-case SAR_{10g} is 28.15 W/kg and the worst-case SAR_{1g} is 68.51 W/kg for the cervical plate system at 3 T. The corresponding worst-case configuration has a plate length of 140 mm, a plate width of 18 mm, a plate depth of 3 mm, a screw length of 10 mm, a screw diameter of 5 mm, and 4 holes. The worst-case SAR_{10g} is 57.42 W/kg and the worst-case SAR_{1g} is 152.42 W/kg for the compression plate system at 1.5 T system. The corresponding worst-case configuration has a plate length of 102 mm, a plate width of 18 mm, a plate depth of 3 mm, a screw length of 10 mm, a screw diameter of 5 mm, and 1 hole. The illustration of worst-case RF-induced heating for the cervical plate system was shown in Figure 22. The hot spots usually occurred at the screw tips.

Table 9. Statistics of the simulation results for the cervical plate system.

Statistics	**SAR_{10g}**		**SAR_{1g}**	
	3 T	**1.5 T**	**3 T**	**1.5 T**
Min	8.47 (W/kg)	11.65 (W/kg)	13.66 (W/kg)	23.71 (W/kg)
Max	28.15 (W/kg)	57.42 (W/kg)	68.51 (W/kg)	152.42 (W/kg)
Mean	17.29 (W/kg)	33.21 (W/kg)	42.43 (W/kg)	80.27(W/kg)
Variance	20.15 (W^2/kg^2)	185.51 (W^2/kg^2)	163.42(W^2/kg^2)	1184.92 (W^2/kg^2)

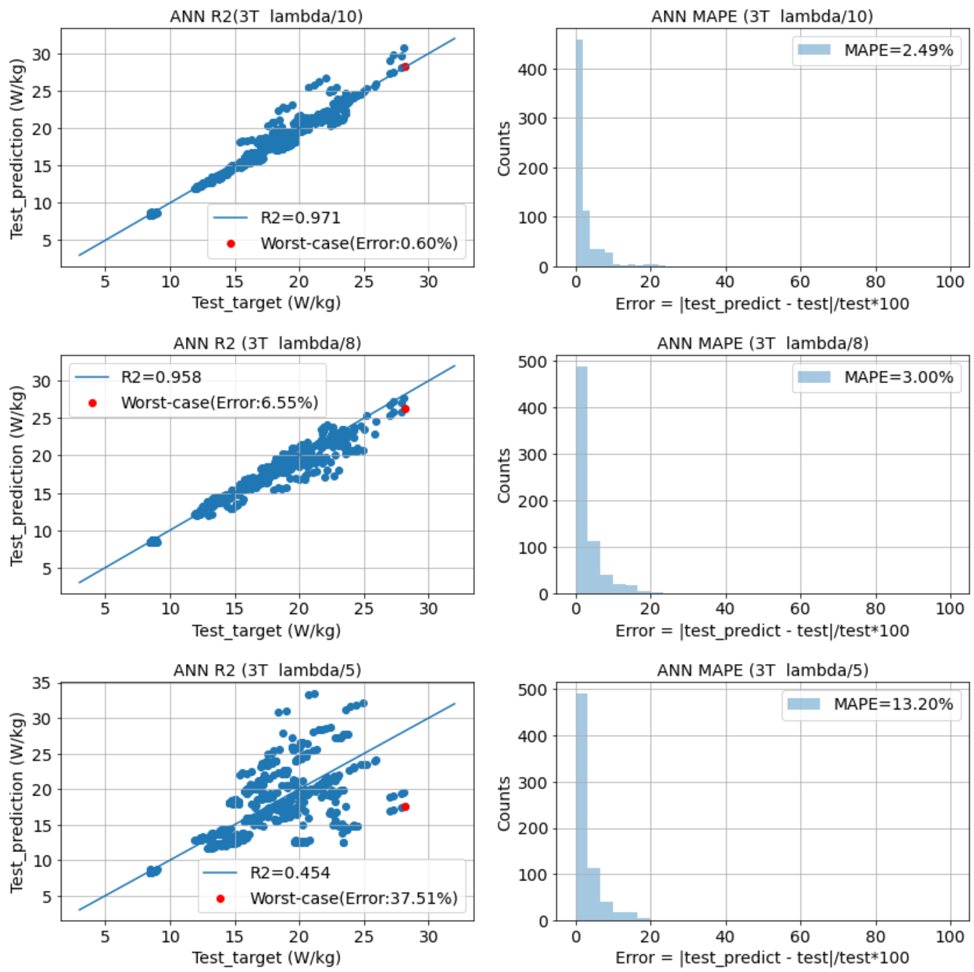


Figure 23. The NN testing results (SAR_{10g}) were trained by different step sizes for the cervical plate system at 3 T system.

The ANN results (SAR_{10g}) at 3 T system with a different step size of cervical plate system were shown in Figure 23. The ANN can predict the worst-case configuration and the corresponding RF-induced heating with small error rates by using a minimum step size of lambda/10. The correlation coefficient between the testing and predicted results was larger than 0.97. The MAPE was small and close to 2.49%. The worst-case prediction error was less than 0.60%. The performance of the network still can achieve acceptable worst-case prediction results when it was trained by a minimum step size of lambda/8. The worst-case prediction error was less than 6.55% and the MAPE was less than 3.00%. In this trained network, the correlation coefficient was larger than 0.95 which indicates the network still can capture the non-linear relationship between the input and output. However, the worst-case prediction error was larger than 37.51% when the ANN was trained by the minimum step size of lambda/5. Furthermore, the network could underestimate the worst-case heating if trained by the step size of lambda/5. The ANN results (SAR_{1g}) at 3 T system with a different step size of cervical plate system were shown in Figure 24. The network can get accurate worst-case prediction results and overall prediction results when it was trained a minimum step size of lambda/10. However, the worst-case error will be very large as it was close to 17.65% when the network was trained by a minimum step size of lambda/8.

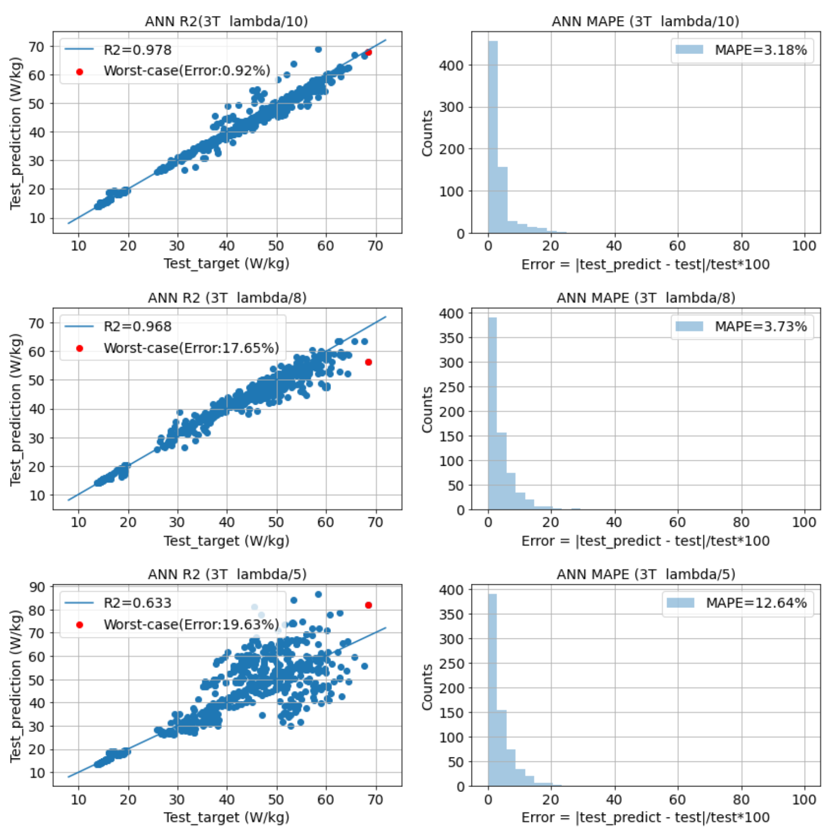


Figure 24. The NN testing results (SAR_{1g}) trained by different step sizes for the cervical plate system at 3 T system.

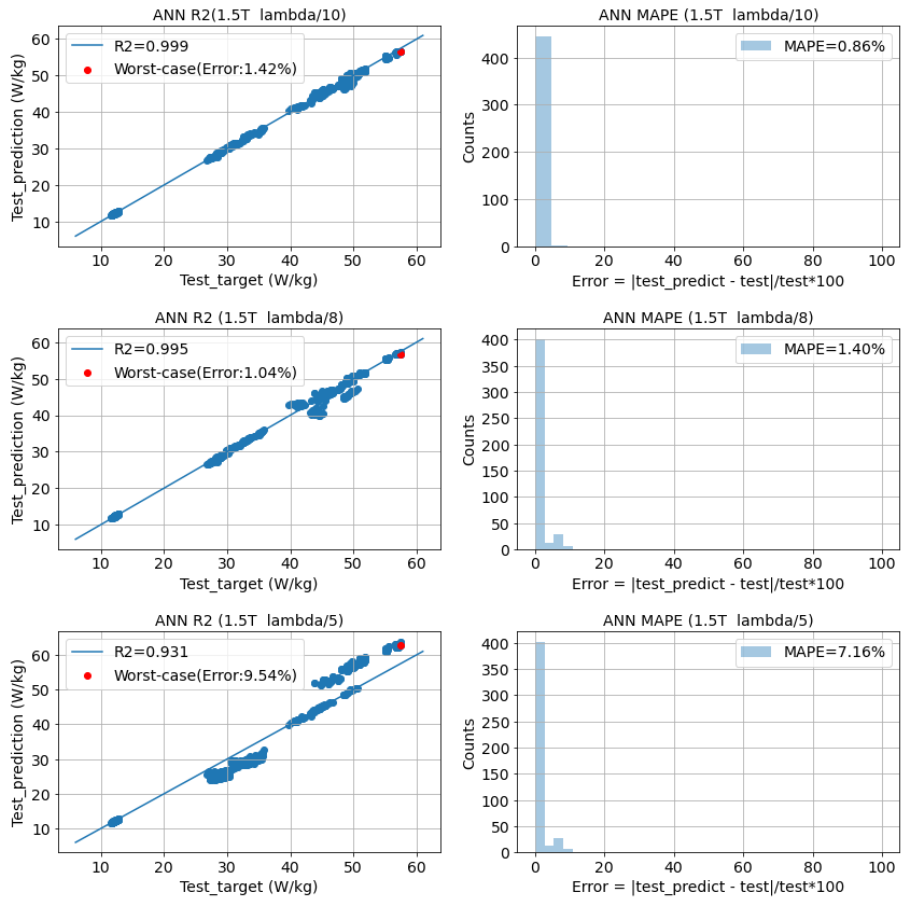


Figure 25. The NN testing results (SAR_{10g}) were trained by different step sizes for the cervical plate system at 1.5 T system.

The ANN results (SAR_{10g}) at 1.5 T system with different step sizes for cervical plate system were shown in Figure 25. As shown in the figure, the network converges very well when trained by a minimum step size that was less than lambda/8. The correlation coefficient of the network was larger than 0.99 and the prediction error was also very small. The MAPE was less than 1.40% and the worst-case prediction error was less than 1.04% if trained by a step size of lambda/8. However, the worst-case prediction error was larger than 9.54% and the MAPE will be larger than 7.16% when the network was trained by a step size of lambda/5. It can be indicated that the step size should be smaller than lambda/8 to get more accurate prediction results for the cervical plate system. The ANN results

(SAR_{1g}) at 1.5 T system with different step sizes for cervical plate system were shown in Figure 26. The network can get accurate worst-case prediction results and overall prediction results in terms of SAR_{1g} when it was trained a minimum step size smaller than lambda/8. In this study, the network still converges well with a minimum step size of lambda/5 as the worst-case prediction error was less than 2.9% and the MAPE was less than 6.17%. To get the conservative prediction results, it's recommended to use a minimum step size of lambda/8 for the RF-induced heating fast evaluation of the cervical plate system.

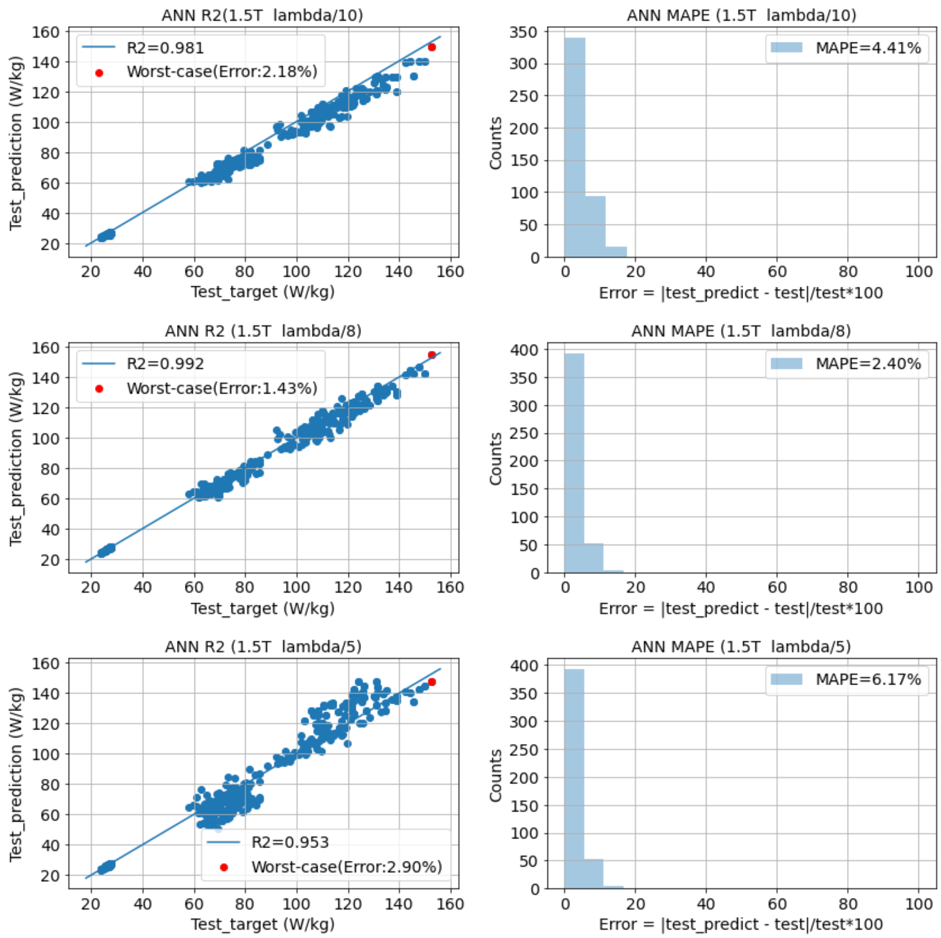


Figure 26. The NN testing results (SAR_{1g}) trained by different step sizes for the cervical plate system at 1.5 T system.

To study the performance of the network to predict the whole sample space using the smallest number of training data set. The trained network was used to predict the device parameters in the continuous sample space and these parameters were swept in fine step size. The other parameters remained unchanged when a selected parameter was swept. The training data set for the cervical plate devices used a converged step size of lambda/8 and the errors between the prediction results and the simulation results were shown in Figure 27. The plate length defined in a continuous sample space was studied as shown in Table 6. The largest predicted error was less than 15% with the different number of holes at the 3 T system. The largest predicted error was less than 7% with the different number of holes at the 1.5 T system.

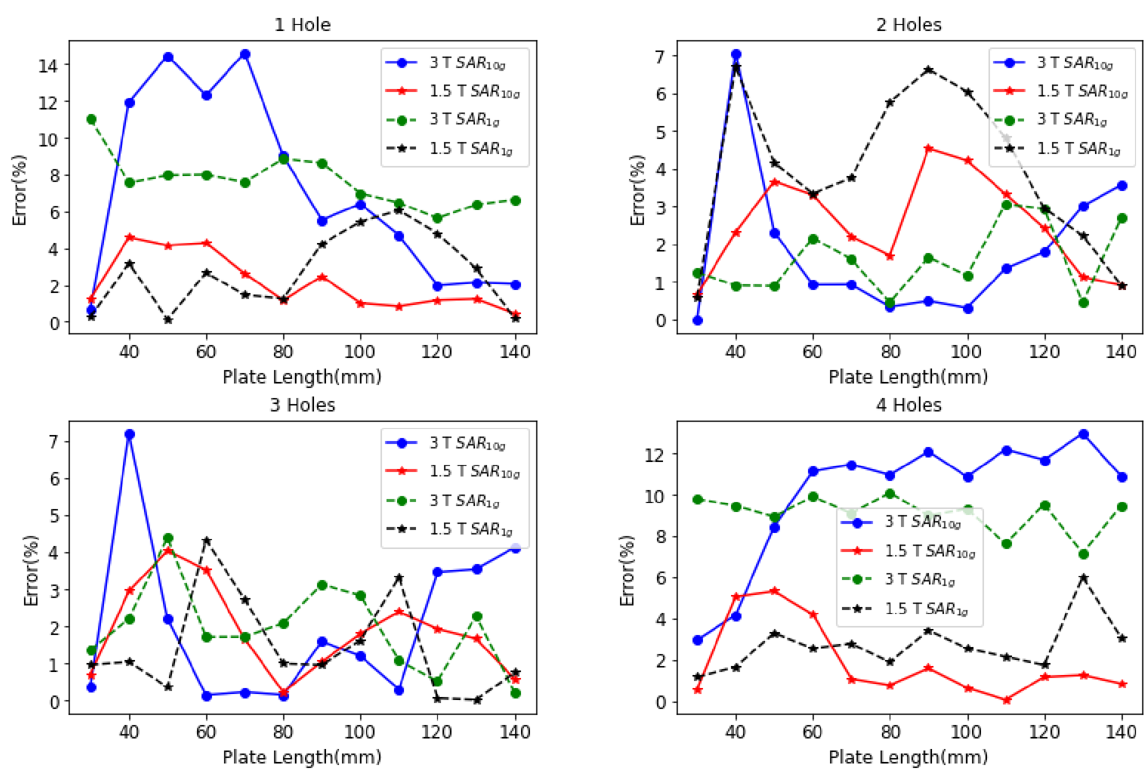


Figure 27. NN prediction errors for the cervical plate system with a fine step of 10 mm.

For the cervical plate system, the plate length is only one parameter used for the ANN can be reduced to smaller sample space. It's also recommended to use a converge step size of lambda/10 to get more accurate prediction results. Therefore, the gain of the network (G_{NN}) for the study of cervical plate system at 1.5 T system is $G_{NN(1.5T)} = \left[1 - \frac{1}{24}\right] \approx$ 95.83%. The gain of the network (G_{NN}) for the study of cervical plate system at 3 T system is $G_{NN(3T)} = \left[1 - \frac{1}{44}\right] \approx 97.72\%$.

6.6 Application of Thoracolumbar Device

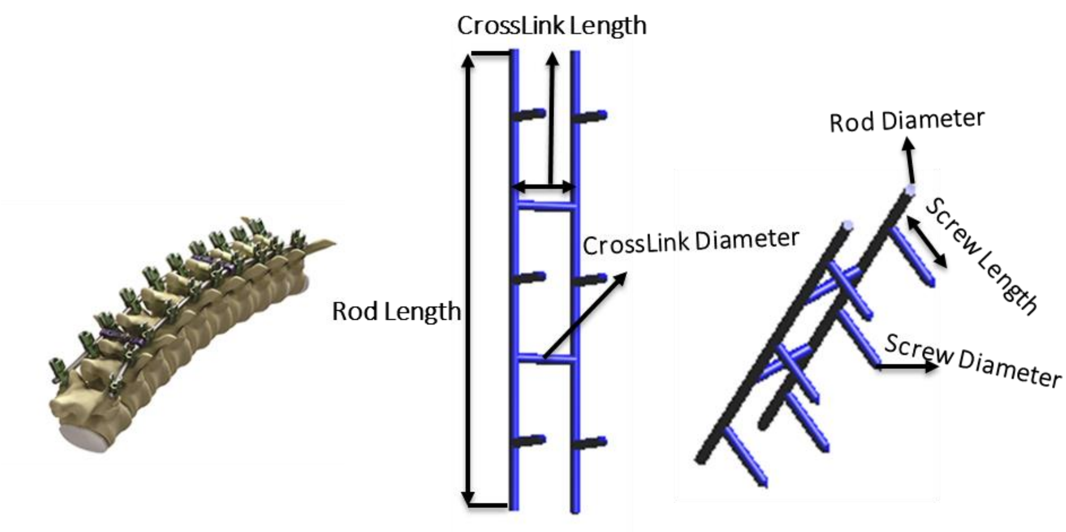


Figure 28. Thoracolumbar Device.

The thoracolumbar device is commonly used for deformities treatment. It consisted of rods, interconnected components (crosslink), and screws. The general spinal cord fixator was shown in Figure 28. The detailed device dimension was shown in Table 10.

Table 10. The detail device dimension for the thoracolumbar device.

Parameters	Device Dimension(mm)
Rod Length	[40-520]
Rod Diameter	[2.5,5]
CrossLink Length	[20,40]
CrossLink Diameter	[5]
Screw Length	[20-116]
Screw Diameter	[3,5]

6.6.1 Numerical Simulation

Numerical simulations were performed to obtain the ground truth data. In this study, for the geometrical features, the minimum step size of lambda/10, and lambda/8 were adopted to get all the simulation cases as shown in Table 11. A total number of 720 unique configurations at 3 T and 280 unique configurations at 1.5 T were studied.

Table 11.Parameter dimensions used for simulations of thoracolumbar device.

Parameters	3 T		1.5 T	
	Dimension (mm)	No. of Values	Dimension (mm)	No. of Values
Rod Length	[40,64,70,88,100,112,130,136,160,184,190,208,220,232,250,256,280,304, 310,328,340,352,370,376,400,424,430,448,460,472,490,496,520]	33	[40,84,95,128,150,172,205,216,260,304,315, 348,370,392,425,436, 480,520]	18
Rod Diameter	[2.5,5]	2	[2.5,5]	2
CrossLink Diameter	[5]	1	[5]	1
CrossLink Length	[20,40]	2	[20,40]	2
Screw Length	[20,44,50,68,80,92,110,116]	8	[20,64,75,108,116]	5
Screw Diameter	[3, 5]	2	[3, 5]	2

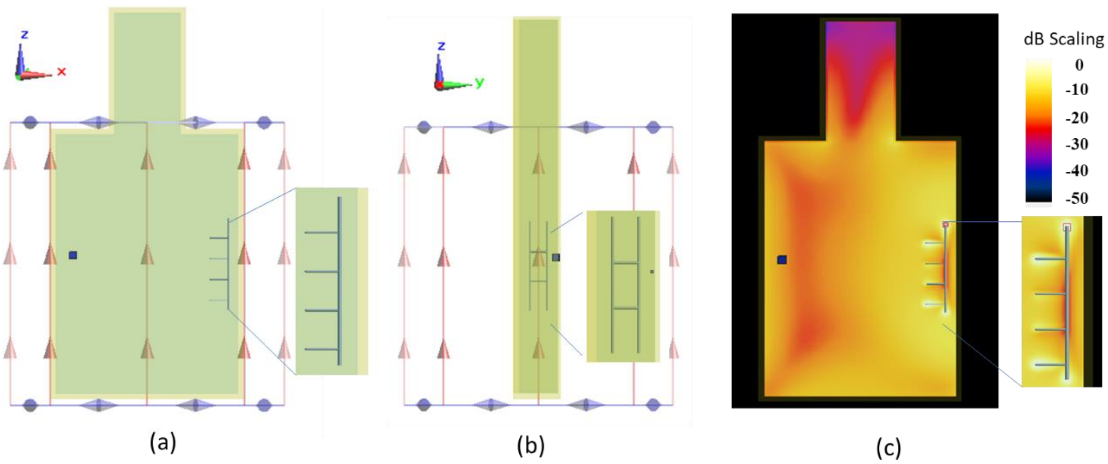


Figure 29. Numerical simulation of the thoracolumbar device conducted at 3T: (a) Front view, (b) Side view (c) Example of the RF-induced heating.

Numerical simulations were also performed both at 1.5 T and 3 T system based on the FDTD method in the ASTM phantom. The high pass non-physical RF transmits body coil and the ASTM phantom were the same as the studies of compression plate system and cervical plate system. The thoracolumbar device was placed at the vertical center on the right side 2 cm away from the phantom wall at the center along the bore direction as shown in Figure 29. An example of the RF-induced heating of the complex shape plate device under the 3 T system showed that the hot spot occurred at the end of the rod. In the numerical simulation, all metallic materials were modeled as perfect electric conductor (PEC).

The non-uniform mesh was used in the simulations. A mesh size of 1 mm was applied to the thoracolumbar devices. The mesh size of the gelled-saline and plastic box was the same as the setting of the compression plate system. The grating ratio of the mesh size was set to 1.15. To ensure convergence, the simulation time was set for 25 periods. All the numerical simulations reached a CL of -50 dB at 3 T, and a CL of 30 dB at 1.5 T. All the results were normalized to a whole-body SAR of 2 W/kg.

6.6.2 Training Set

The training set used for the thoracolumbar device was shown in Table 12. The range of the rod length is covered in 40mm to 500 mm and the crosslink length ranges from 20 mm to 100 mm. Both rod length and crosslink length used the step size of $\lambda_{3T}/10$.

Table 12. Different step sizes study for the thoracolumbar device.

Parameters	3 T			1.5 T		
	$\lambda_{3T}/10$ (mm)	$\lambda_{3T}/8$ (mm)	$\lambda_{3T}/5$(mm)	$\lambda_{1.5T}/10$ (mm)	$\lambda_{1.5T}/8$ (mm)	$\lambda_{1.5T}/10$ (mm)
Rod Length	[40,64,88,112,136,160,184,208,232,256,280,304,328,352,376,400,424,448,472,496,520]	[40,70,100,130,160,190,220,250,280,310,340,370,400,430,460,490,520]	[40,88,136,184,232,280,328,376,424,472, 520]	[40,84,128,172,216,260,304,348,392,436,480,520]	[40,95,150,205,260,315,370,425,480,520]	[40,128,216,304,392,480,520]
Rod Diameter	[2.5,5]	[2.5,5]	[2.5,5]	[2.5,5]	[2.5,5]	[2.5,5]
CrossLink Diameter	[5]	[5]	[5]	[5]	[5]	[5]
CrossLink Length	[20, 44,68,92,116]	[20, 50,80,110,116]	[20,68,116]	[20,64,108,116]	[20,75,116]	[20,108,116]
Screw Length	[20,40]	[20,40]	[20,40]	[20,40]	[20,40]	[20,40]
Screw Diameter	[3,5]	[3,5]	[3,5]	[3,5]	[3,5]	[3,5]

6.6.3 Results

The simulation results of the thoracolumbar devices were shown in Figure 30. The simulation results indicated that the mean SAR_{10g} at 3 T system is 16.58 W/kg and the mean SAR_{10g} at 1.5 T system is 22.26 W/kg. For the simulation results in terms of SAR_{1g}, the mean SAR_{1g} at the 3 T system is 57.65 W/kg and the mean SAR_{1g} at the 1.5 T system is 83.46 W/kg.

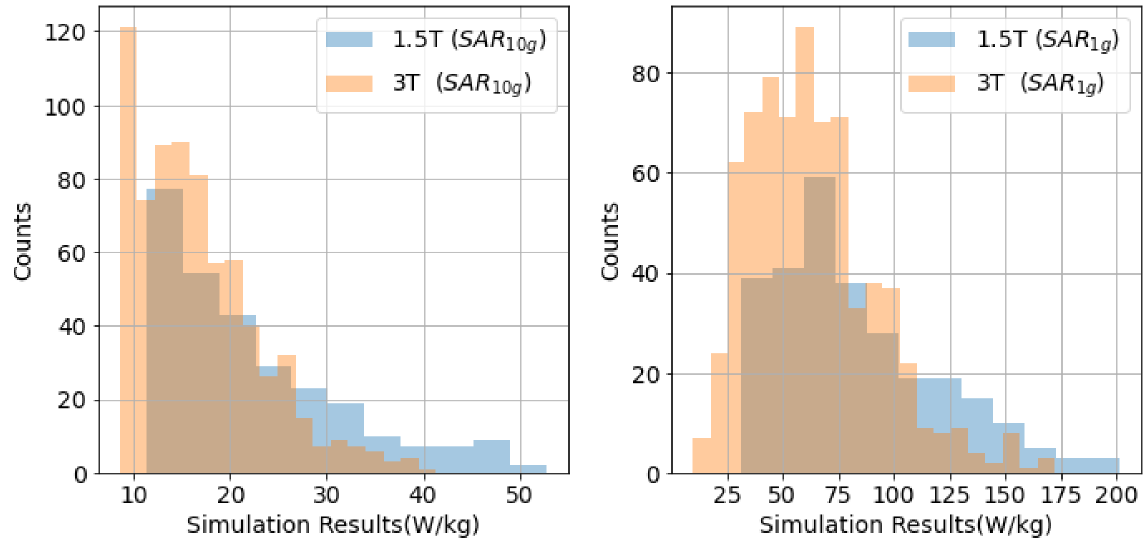


Figure 30. Simulation results for the thoracolumbar devices.

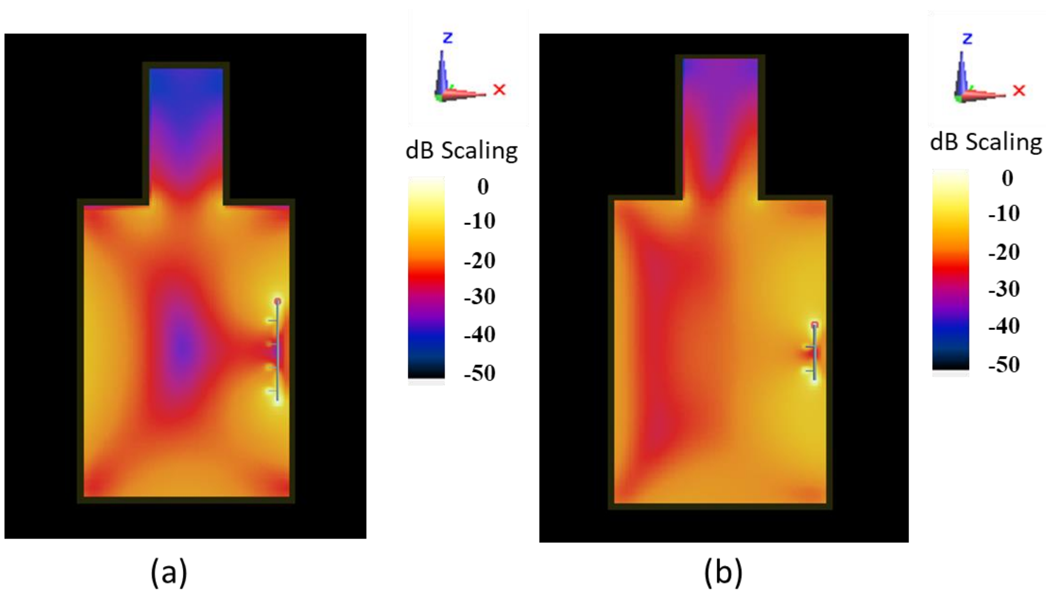


Figure 31. Illustration of the worst-case RF-induced heating of thoracolumbar device at (a) 1.5 T system and (b) 3 T system.

The worst-case RF-induced heating for the thoracolumbar devices was shown in Figure 31. The hot spots usually occurred at the top of the rod. The statistics of the simulation results for the thoracolumbar device were shown in Table 13. The worst-case SAR_{10g} is 41.33 W/kg and the worst-case SAR_{1g} is 172.21 W/kg for the thoracolumbar device at 3 T. The corresponding worst-case configuration has a rod length of 112 mm, a screw length of 20 mm, a crosslink length of 40 mm, a screw radius of 1.5 mm, a rod diameter of 2.5 mm and a crosslink radius of 2.5 mm. The worst-case SAR_{10g} is 52.79 W/kg and the worst-case SAR_{1g} is 201.64 W/kg for the thoracolumbar device at 1.5 T system. The worst-case configuration has a rod length of 205 mm, a screw length of 20 mm, a crosslink length of 40 mm, a screw radius of 1.5 mm, a rod diameter of 2.5 mm, and a crosslink radius of 2.5 mm.

Table 13. Statistics of the simulation results for the thoracolumbar devices.

Statistics	**SAR_{10g}**		**SAR_{1g}**	
	3 T	**1.5 T**	**3 T**	**1.5 T**
Min	8.49 (W/kg)	11.26 (W/kg)	9.77 (W/kg)	31.38 (W/kg)
Max	41.33 (W/kg)	52.79 (W/kg)	172.21 (W/kg)	201.64 (W/kg)
Mean	16.58 (W/kg)	22.26 (W/kg)	57.65 (W/kg)	83.46 (W/kg)
Variance	40.88 (W^2/kg^2)	89.34 (W^2/kg^2)	814.26 (W^2/kg^2)	1323.04 (W^2/kg^2)

The ANN results (SAR_{10g}) at the 3 T system with a different step size of the thoracolumbar device were shown in Figure 32. The ANN can accurately predict the worst-case SAR by using a minimum step size of lambda/10 at the 3 T system. As demonstrated in the figure, the correlation coefficient of the ANN was larger than 0.93 by evaluating the ground-truth simulation results and the predicted results by using a minimum step size of lambda/10. The worst-case prediction error of the ANN was less than 2.02%. The prediction accuracy of the ANN tends to be decreased dramatically when it was trained by a minimum step size of lambda/8. In this case, the correlation coefficient of the ANN was less than 0.86 and the worst-case prediction error was larger than 13.03%. For all the testing cases, the MAPE was larger than 6.63 %.

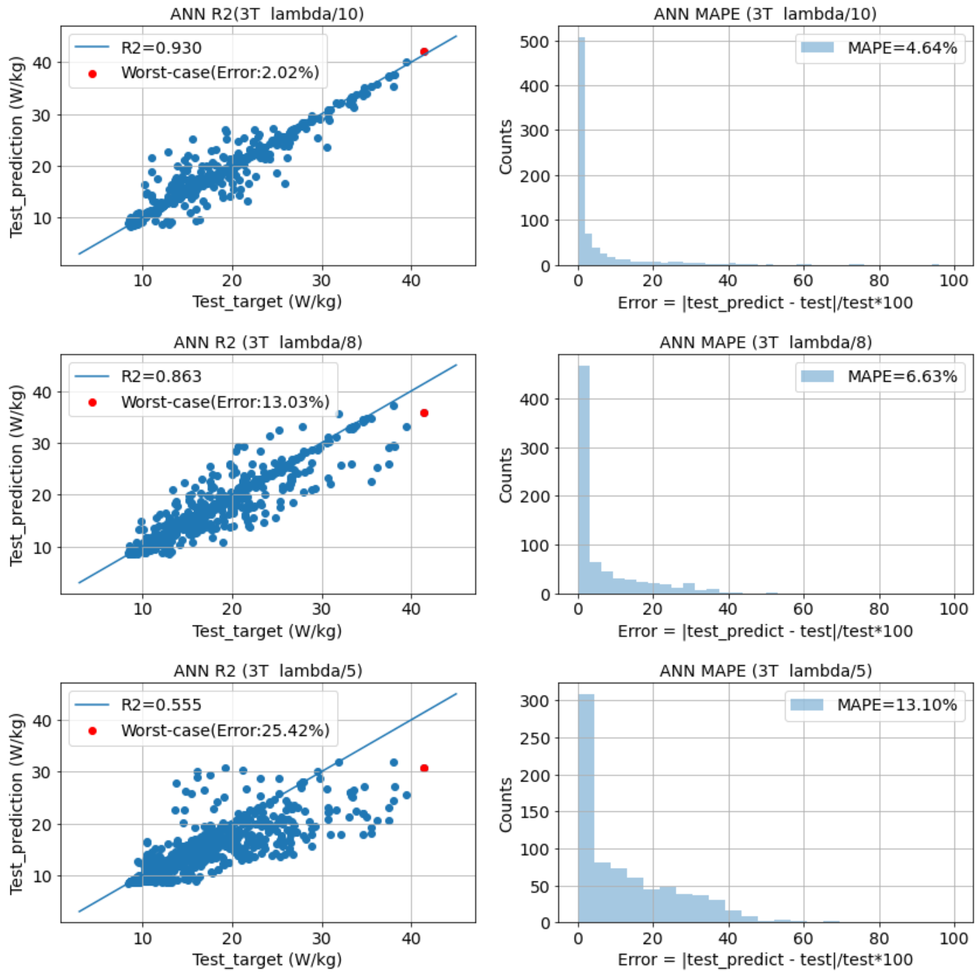


Figure 32. The NN testing results (SAR_{10g}) were trained by different step sizes for the thoracolumbar device at the 3 T system.

The ANN results (SAR_{1g}) at the 3 T system with a different step size of the thoracolumbar device were shown in Figure 33. The correlation coefficient of the network was larger than 0.90 when it was trained by a minimum step size of lambda/10. The predicted worst-case error will be less than 3.11% and the MAPE will be less than 6.86%. However, the worst-case prediction error will be larger than 23.42% when the network was trained by a minimum step size of lambda/8. Thus, the criterion of the lambda/10 still can be applied to the thoracolumbar devices at the 3 T system.

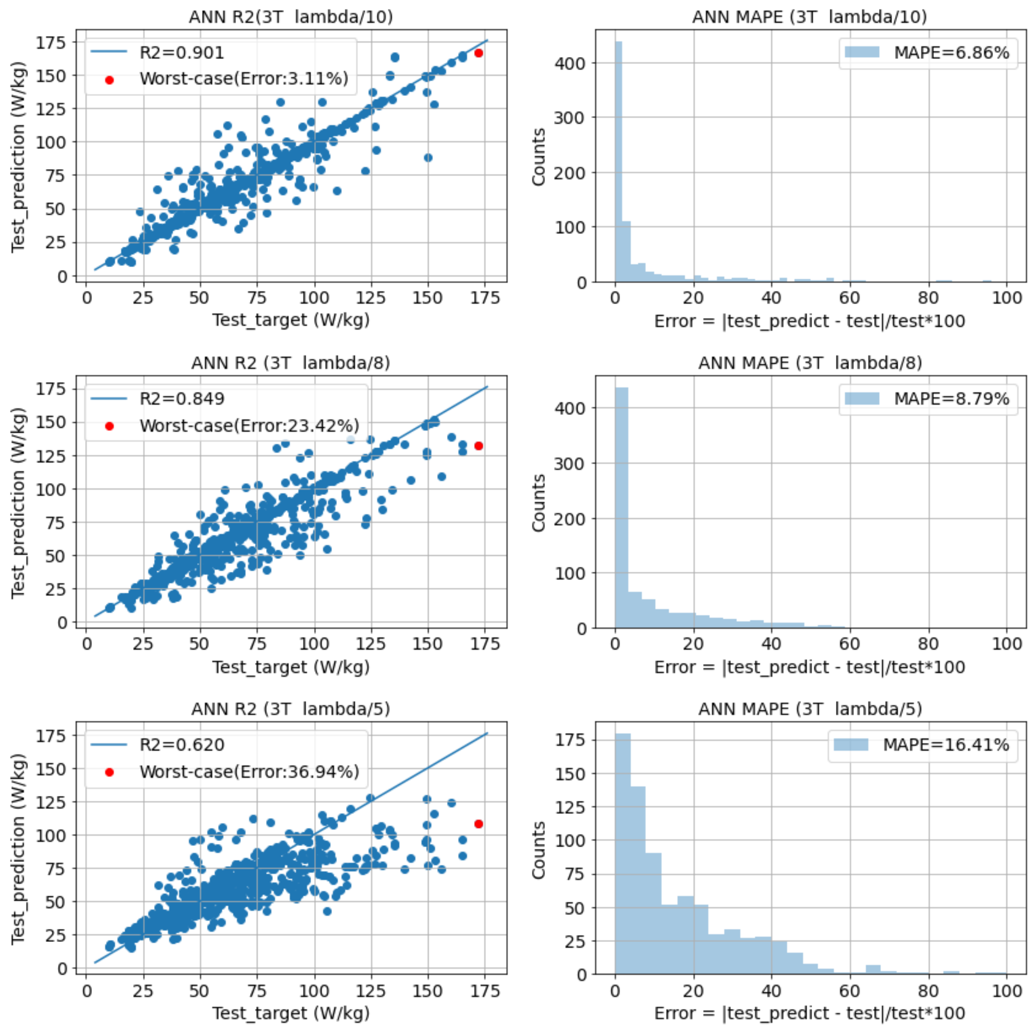


Figure 33. The NN testing results (SAR_{1g}) trained by different step sizes for the thoracolumbar device at the 3 T system.

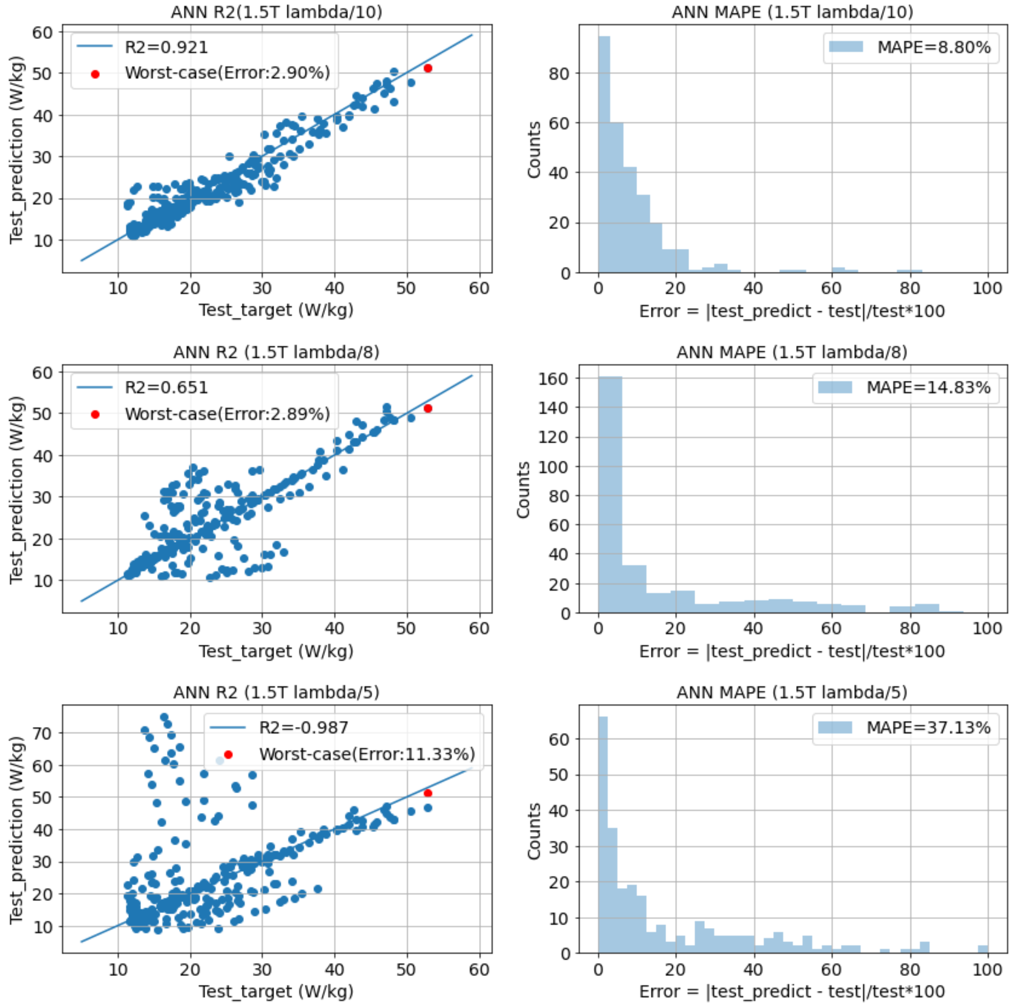


Figure 34. The NN testing results (SAR_{10g}) were trained by different step sizes for the thoracolumbar device at a 1.5 T system.

The ANN results (SAR_{10g}) at the 1.5 T system with different step sizes of the thoracolumbar device were shown in Figure 34. The worst-case prediction error was low when the network was trained by a minimum step size of lambda/10 and the MAPE of the ANN was less than 8.80% for all the testing data. However, the overall prediction error will be increased if the network was trained by using a minimum step size of lambda/8. As shown in the figure, the MAPE will be increased to 14.83%. As the step size was larger, the network performance will decrease dramatically.

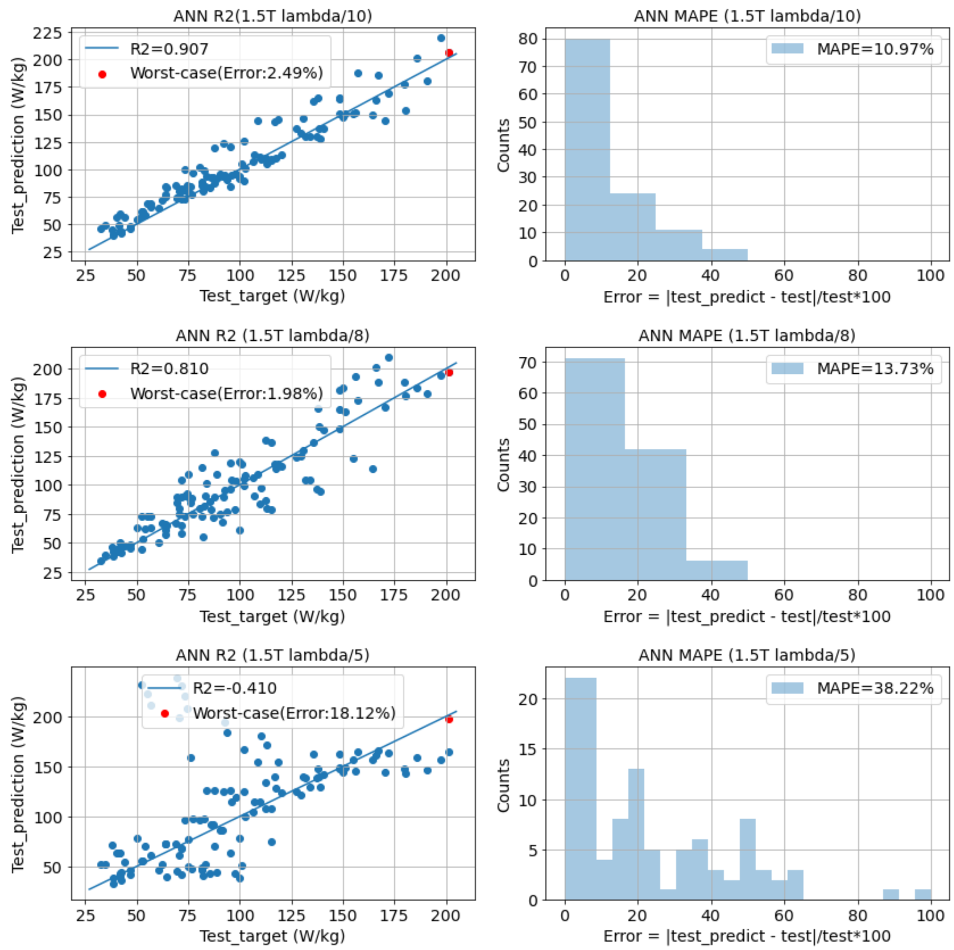


Figure 35. The NN testing results (SAR_{1g}) trained by different step sizes for the thoracolumbar device at 1.5 T system.

The ANN results (SAR_{1g}) at the 1.5 T system with different step sizes of the thoracolumbar device were shown in Figure 35. The network can learn the non-linear relationship between the parameters and the SAR_{1g} when the network was trained by a minimum step size of lambda/10 because the correlation coefficient of the network was larger than 0.90. The worst-case prediction error for the SAR_{1g} evaluation was less than 2.49% and the MAPE was less than 10.97%. The MAPE of the network prediction results will be larger than 13.73% when the network was trained by using a minimum step size of lambda/8. Thus, to get more accurate RF-induced heating, it is recommended to use a minimum step size of lambda/10 to train the network for the thoracolumbar devices.

To study the general performance of the network to predict the whole sample space using the smallest number of training data set. The trained network with a converged step size of lambda/10 was used to predict RF-induced heating using the device parameters in the continuous sample space. These parameters were swept in fine step size and other parameters remained unchanged when a selected parameter was swept. The errors between the prediction results and the simulation results were shown in Figure 36. In this study, the rod length, and screw length were in the continuous sample space. The structure of the thoracolumbar device was more complicated than the compression plate system and the cervical plate system. Thus, network predicted results tend to diverge and errors would be larger. The largest predicted error was less than 18% for the study of rod length and the largest predicted error was smaller than 5% for the screw length at the 3 T system. The largest predicted error was also less than 18% for the study of rod length and the largest error was less than 9% for the study of screw length at the 1.5 T system.

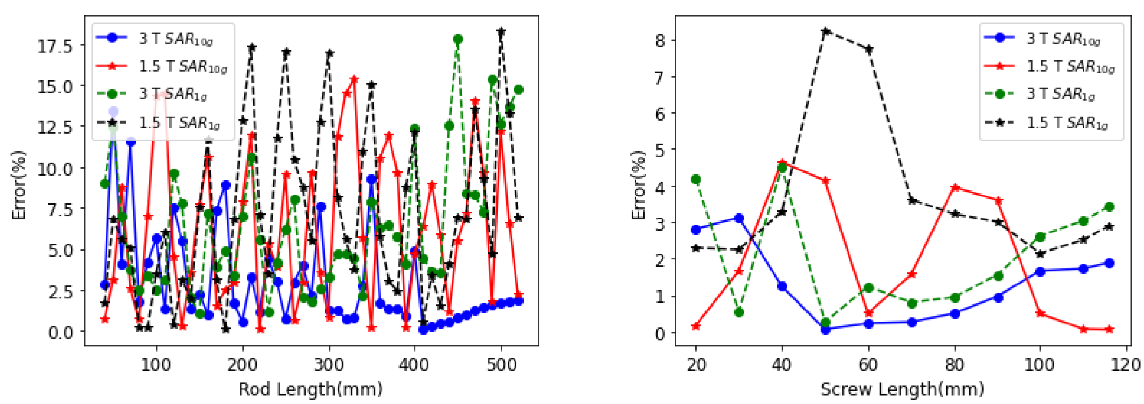


Figure 36. NN prediction errors for the thoracolumbar device with a fine step of 10 mm.

For the thoracolumbar devices, both the rod length and screw length which was used as the input of the ANN can be reduced to a smaller sample space. It's also recommended to use a converge step size of lambda/10 to get more accurate prediction results. Therefore, the gain of the network (G_{NN}) for the study of thoracolumbar devices at 1.5 T system is $G_{NN(1.5T)} = \left[1 - \frac{1}{(24)^2}\right] \approx 99.82\%$. The gain of the network (G_{NN}) for the study of thoracolumbar devices at 3 T system is $G_{NN(3T)} = \left[1 - \frac{1}{(44)^2}\right] \approx 99.94\%$.

6.7 Time Analysis for Parameterized NN

The upper bound of time complexity of the parameterized NN can be expressed as,

$$O(n \times t \times x \times l \times k^2), \tag{15}$$

where n is the number of training samples, x is the number of features, l is the number of layers, k is the max number of nodes in all layers, and the training epochs (iterations) of the network. The time cost of the entire training of the network for our 4-layer network with thousands of samples, devices described by 11 features, and trained by 2000 epochs was within 120s. This is much less than the full-wave modeling of the devices. Because each device inside the phantom will cost more than 2 hours by the numerical simulation using an NVIDIA C2075 high-performance graphics processing unit (GPU) which had a normal 1150 MHz clock rate of 448 CUDA cores. Once the network has been trained, the 3-layer network predictive models are capable of producing thousands of RF-induced heating predictions within seconds.

7 External Fixation Use in-vitro Phantom with Parameterized NN

The partially implanted devices that in in-vitro conditions are used as the study subjects. One rod with two pins was created as a study example to determine the criterion for the training data set. Then one representative partially implanted device was used to validate the criterion for the training data set. To obtain the ground-truth data, the RF-induced heating for the devices was numerically investigated based on the FDTD method. The same NN architecture was constructed the same as the architecture used by fully implanted devices but with different input dimensions. At last, the NNs were trained by the training data set determined by the criterion, while the rest were used for fast evaluation of RF-induced heating and testing performance of the NNs.

7.1 The Criterion for Training Data Set

The external fixation device is partially implanted in the body. The wavelength/10 rule may not apply since the incident field can change. Thus, we use the rod with two pins to determine the criterion. The device can be described by two parameters, the rod length, and insertion depth. The insertion depth the part of the device inserted into the gel.

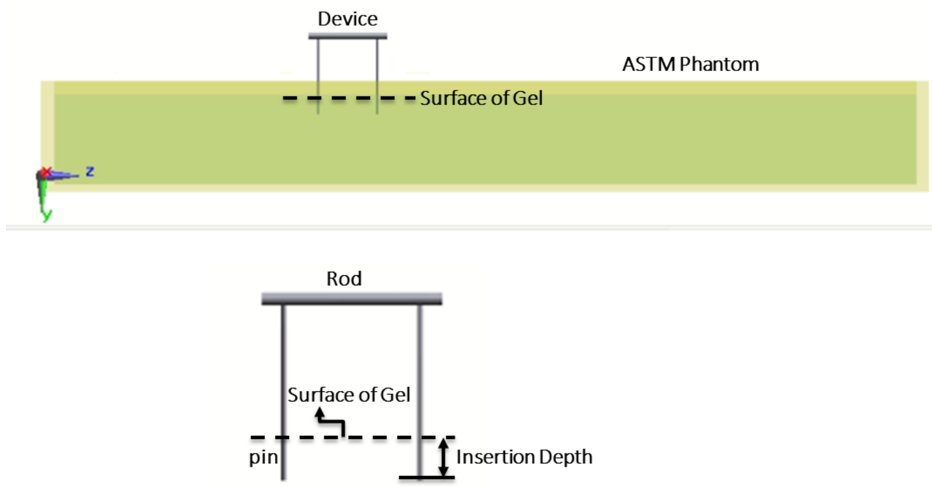


Figure 37. The study rod with two pins.

7.1.1 Numerical Simulation

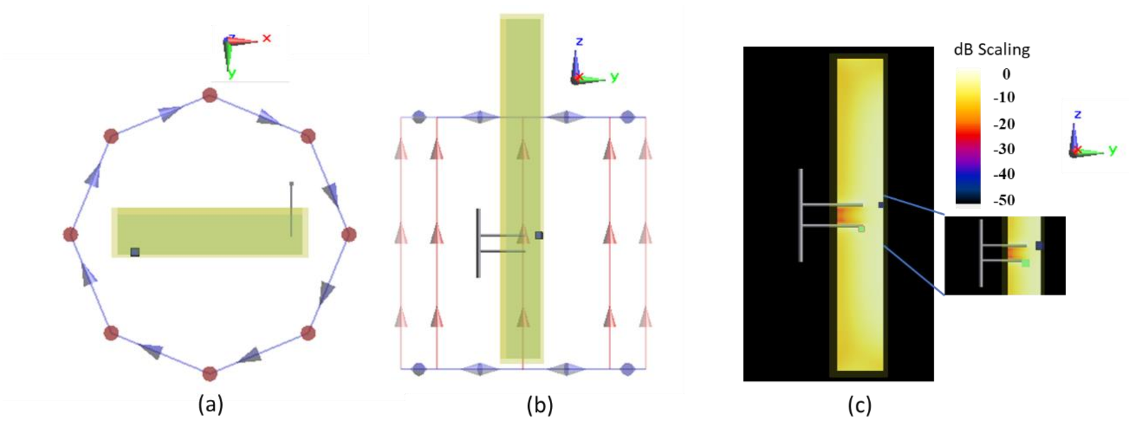


Figure 38. Numerical simulation of the simple external fixator conducted at 3T: (a) Bottom view, (b) Side view (c) Example of the RF-induced heating.

The external fixation devices are usually made of metallic materials and were set to be the perfect electric conductor (PEC) in the simulations. An example of the simple external fixator conducted at the 3 T system in the ASTM phantom is shown in Figure 38. The hot spot was close to the end of the pins. The high pass non-physical RF transmits body coil and the ASTM phantom were also the same as the studies of compression plate system. Adaptive meshing was adopted at 3 T, and other settings like mesh size and grating ratio were the same as the setting of the compression plate system. All the voxelized models were checked to avoid meshing errors. All the simulation results were checked to ensure convergence. All the results in terms of SAR_{10g} were also normalized to a whole-body SAR of 2 W/kg.

The rod length can be in the range from 80 mm to 500 mm, and insertion depth can be in the range from 20 mm to 70 mm. Both rod length and insertion depth were using a step size of 10 mm. The commonly used rod diameter of 5 mm and a screw diameter of 3 mm were studied. Thus, there are a total number of $43 \times 6 = 258$ configurations.

Table 14. The study parameters and step size for the rod with two pins.

Parameters	**Values (mm)**	**Step Size (mm)**	**No. of Values**
Rod Length	[80~500]	10	43
Insertion Depth	[20~70]	10	6
Rod Diameter	5	/	/
Pin diameter	3	/	/

7.1.2 Training Set

Part of the data was selected and used as the training data set which contains different step sizes as shown in Table 15. For each MRI operating frequency, the training set was selected as step size of 10 mm, 20 mm, 30 mm, 40 mm, 50 mm, and 60 mm. The worst-case configuration was not included in the training set.

Table 15. Training data set for the rod with two pins.

Parameters	Values (mm)	Step Size (mm)	No. of Values
Rod Length	[80~500]	[10,20,30,40,50,60]	43
Insertion Depth	[20~70]	[10,20,30,40,50,60]	6
Rod Diameter	5	/	/
Pin diameter	3	/	/

7.1.3 Results

The worst-case configuration at 1.5 T has a rod length of 130 mm and an insertion depth of 20 mm with a SAR_{10g}=14.84 W/kg and SAR1g = 36.45 W/kg. The worst-case configuration at 3 T has a rod length of 500 mm and an insertion depth of 20 mm with a SAR_{10g} = 12.76 W/kg and SAR_{1g} = 33.96 W/kg.

The ANN results (SAR_{10g}) at a 1.5 T system with different step sizes were shown in Figure 39. The neural network was first trained with a small step size of 10 mm and the worst-case configuration was excluded from training. Then, the step size will be increased by 10 mm each time to validate the predicted results of the network. The prediction errors were small when the step size was less than 50 mm. However, the prediction errors will larger than 3.62 % when the step size was larger than 50 mm. Furthermore, the correlation coefficient between the parameterized features and the output SAR_{10g} was smaller than 0.72. This indicates the network will not learn a strong non-linear relationship between input and output. The predicted results could be underestimated if it is lower than the ground-truth RF-induced heating. The overall prediction errors for SAR_{10g} will be larger than 5.51% if the network was trained by a step size of 60 mm. The ANN results (SAR_{1g}) at a 1.5 T system with different step sizes were shown in Figure 40. The overall prediction errors for SAR_{1g} will be larger than 5.68% if the network was trained by a step size of 50 mm. Therefore, the step size used to select the training set of the neural network should be less than 50 mm in a 1.5 T system.

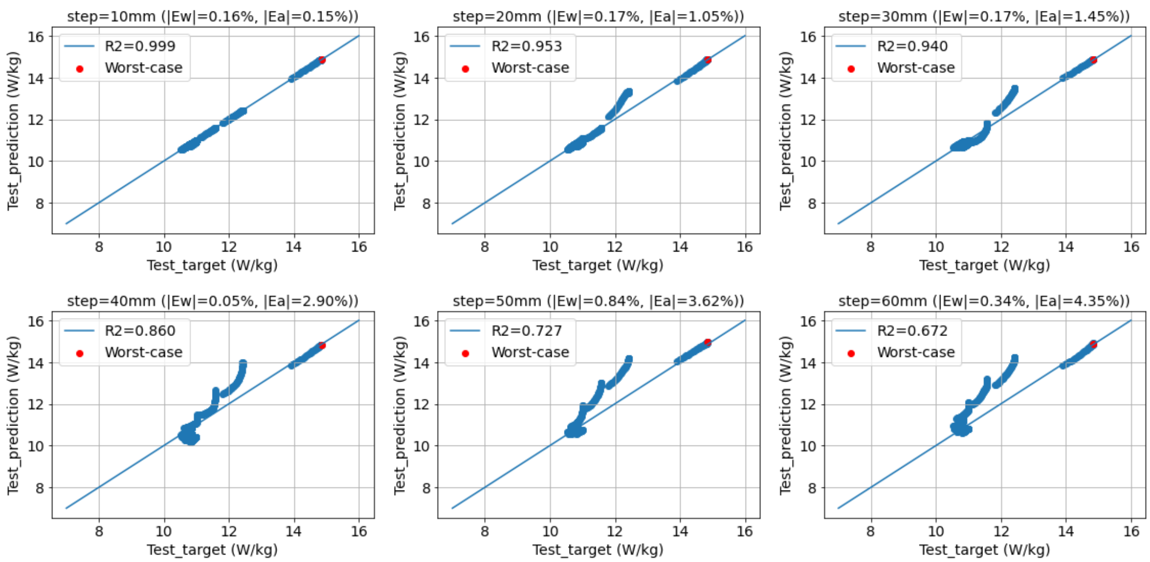


Figure 39. The ANN results (SAR_{10g}) at 1.5 T system with different step sizes for simple external fixator.

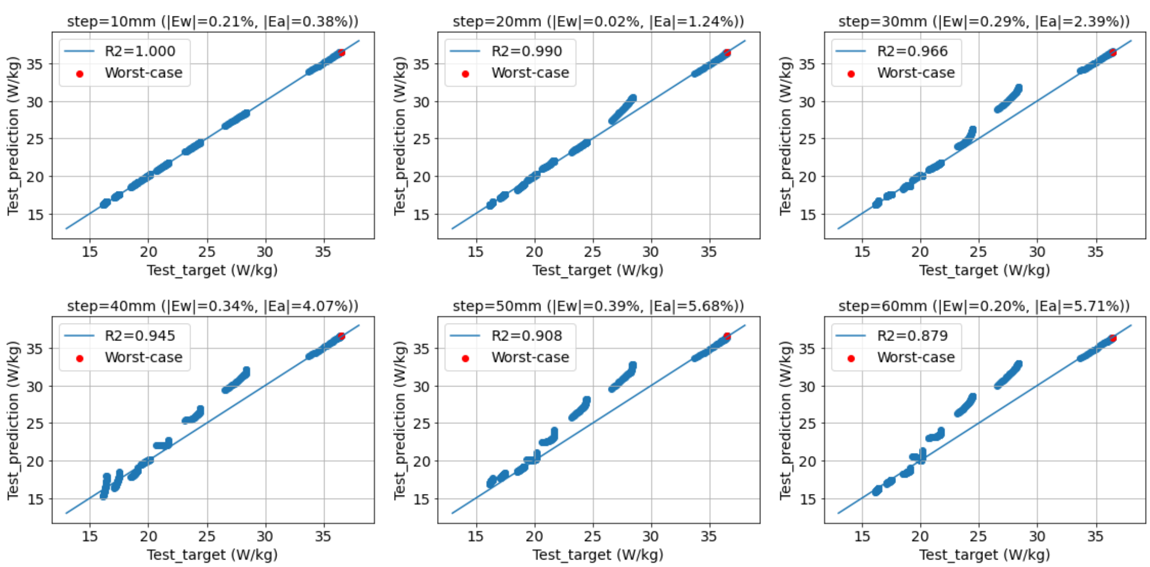


Figure 40. The ANN results (SAR_{1g}) at 1.5 T system with different step sizes for simple external fixator.

The ANN results (SAR_{10g}) at the 3 T system with a different step size of the simple external fixator were shown in Figure 41. The overall prediction errors were small when the step size was less than 20 mm. The worst-case prediction error was less than 1.50% when the step size was smaller than 20 mm. The performance of the network will decrease dramatically when the step size was larger than 30 mm as the worst-case error was larger than 27.26% and the correlation coefficient was very small. Similarly, the ANN results (SAR_{1g}) at the 3 T system with a different step size of the simple external fixator were shown in Figure 42. The overall error will be larger than 16.78% if the network was trained by a step size larger than 30 mm. Thus, the step size used to select the training set of the neural network should be less than 30 mm in a 3 T system.

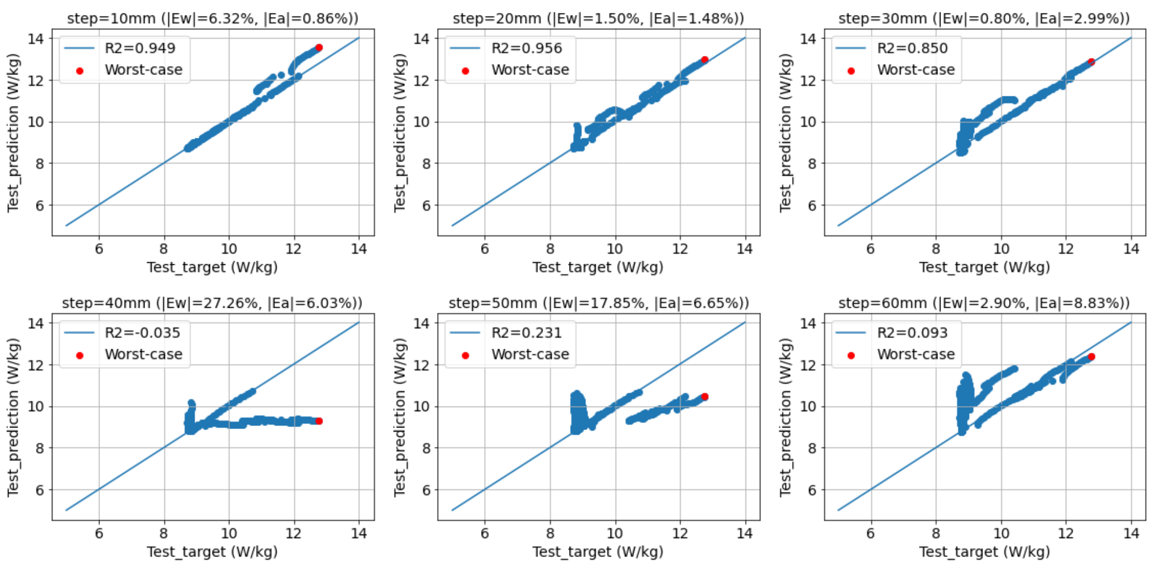


Figure 41. The ANN results (SAR_{10g}) at 3 T system with different step sizes for simple external fixator.

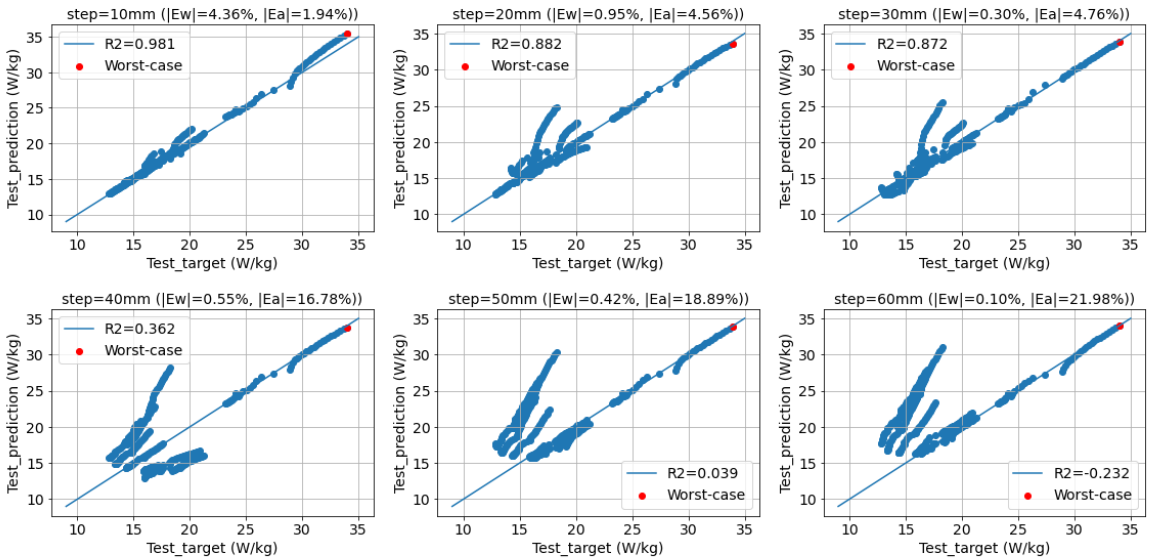


Figure 42. The ANN results (SAR_{1g}) at 3 T system with different step sizes for simple external fixator.

Based on the study, to ensure the worst-case heating construct can be identified with ANN, it is suggested that one should use dimensional variations between training elements that should be less than 1/10 of the wavelength at 1.5 T system ($\approx 50mm$). For the devices at the 3 T system, the dimensional variations between training elements should be less than 1/8 of the wavelength ($\approx 30\ mm$).

7.2 Application of Complex External Fixator

The more complex external fixator will be more stable as it contains clamps to stable multiple screws as shown in Figure 43. The typical parameters will be the rod length, screw radius, screw spacing, frame distance, and insertion depth.

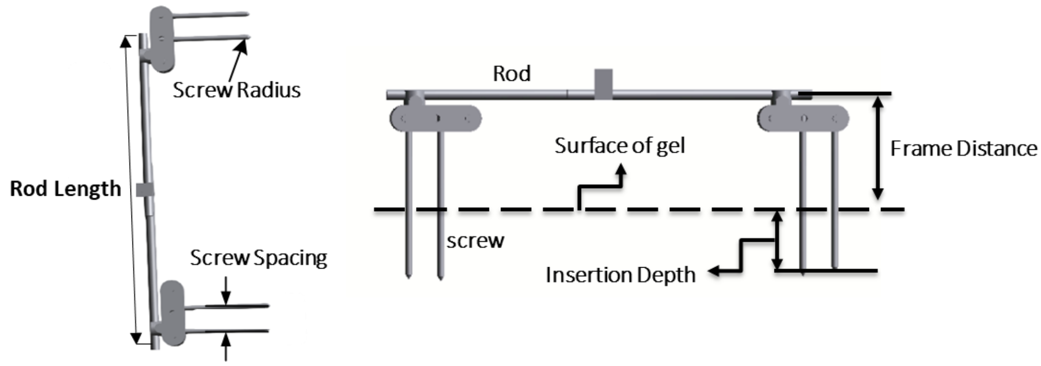


Figure 43. The Complex External Fixator.

The detail dimension for each parameter was shown in Table 16. The rod length, insertion depth, and frame distance were in continuous sample space. The rod length is in the range from 64 mm to 140 mm, the insertion depth is in the range from 20 mm to 80 mm, and the frame distance is in the range from 40 mm to 100 mm. For the screw radius and screw spacing, these two parameters can take optional values for dimension changes. The commonly used rod radius of 2.5 mm was used in this study.

Table 16. The device dimension of the complex external fixator.

Parameters	**Device Dimension(mm)**
Rod Length	[128-280]
Insertion Depth	[20-80]
Frame Distance	[40-100]
Screw Radius	[3,5]
Screw Spacing	[10,20]

7.2.1 Numerical Simulations

Numerical simulations were performed to obtain the ground truth data. In this study, for the parameterized features, the minimum step size of lambda/10, and lambda/8 were adopted to get all the simulation cases as shown in Table 17. A total number of 716 unique configurations were studied at 3 T and 292 unique configurations were studied at 1.5T.

Table 17.Parameter dimensions used for simulations of the complex external fixator.

Parameters	3 T		1.5 T	
	Dimension (mm)	No. of Values	Dimension (mm)	No. of Values
Rod Length	[128,152,158,176,188,200,218,224,248,272,278,280]	12	[128,172,183,216,238,260,280]	7
Insertion Depth	[20,44,50,68,80]	5	[20,64,75,80]	5
Frame Distance	[40,64,70,88,100]	5	[40,84,95,100]	4
Screw Radius	[3,5]	2	[3,5]	2
Screw Spacing	[10,20]	2	[10,20]	2

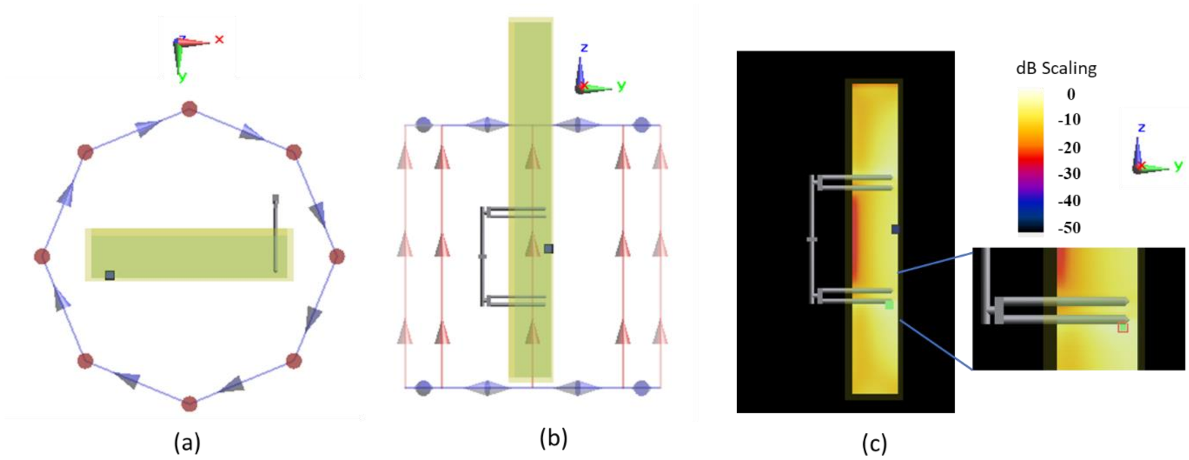


Figure 44. Numerical simulation of the complex external fixation device conducted at 3T: (a) Bottom view, (b) Side view (c) Example of the RF-induced heating.

An example of the complex external fixator conducted at 3 T system in ASTM phantom is shown in Figure 44. The hot spot was close to the end of the pins. Other settings were keeping the same as the settings of the simple external fixator. All the simulation results were checked to ensure convergence. All the numerical simulation results reached a CL of -50 dB at 3 T and -30 dB at 1.5 T. All the results were normalized to a whole-body

SAR of 2 W/kg.

7.2.2 Training set

The different step sizes at the 3 T system for the studied parameters of the complex fixator was shown in Table 18. The range of rod length is from 64 mm to 144 mm, the insertion depth is with a range from 20 mm to 80 mm, and frame distance is a range from 40 mm to 100 mm. The screw radius usually very small, thus it can be set to two optional values. The screw spacing was used to identify the distance between screws. The network will be trained by each training set and tested by all the configurations.

Table 18. The different step sizes for studied parameters of the complex external fixator.

Parameters	3 T			1.5 T		
	$\lambda_{3T}/10$ (mm)	$\lambda_{3T}/8$ (mm)	$\lambda_{3T}/5$ (mm)	$\lambda_{1.5T}/10$ (mm)	$\lambda_{1.5T}/8$ mm	$\lambda_{1.5T}/10$ (mm)
Rod Length	[128,152,176,200,224,248,272,280]	[128, 158,188,218,248,278,280]	[128,176, 224,272, 280]	[128,172,216,260, 280]	[128,183,238,280]	[128,216, 280]
Insertion Depth	[20,44,68,80]	[20,50,80]	[20,68,80]	[20,64,80]	[20,75,80]	[20,80]
Frame Distance	[40,64,88,100]	[40,70,100]	[40,88,100]	[40,84,100]	[40,95,100]	[40,100]
Screw Radius	[3,5]	[3,5]	[3,5]	[3,5]	[3,5]	[3,5]
Screw Spacing	[10,20]	[10,20]	[10,20]	[10,20]	[10,20]	[10,20]

7.2.3 Results

The simulation results of the complex external fixator were shown in Figure 45. The mean SAR_{10g} at the 3 T system is 11.95 W/kg and the mean SAR_{10g} at the 1.5 T system is 21.26 W/kg from all the simulations. The mean SAR_{1g} at 3 T is 26.52 W/kg and the mean SAR_{1g} at 1.5 T is as high as 47.35 W/kg.

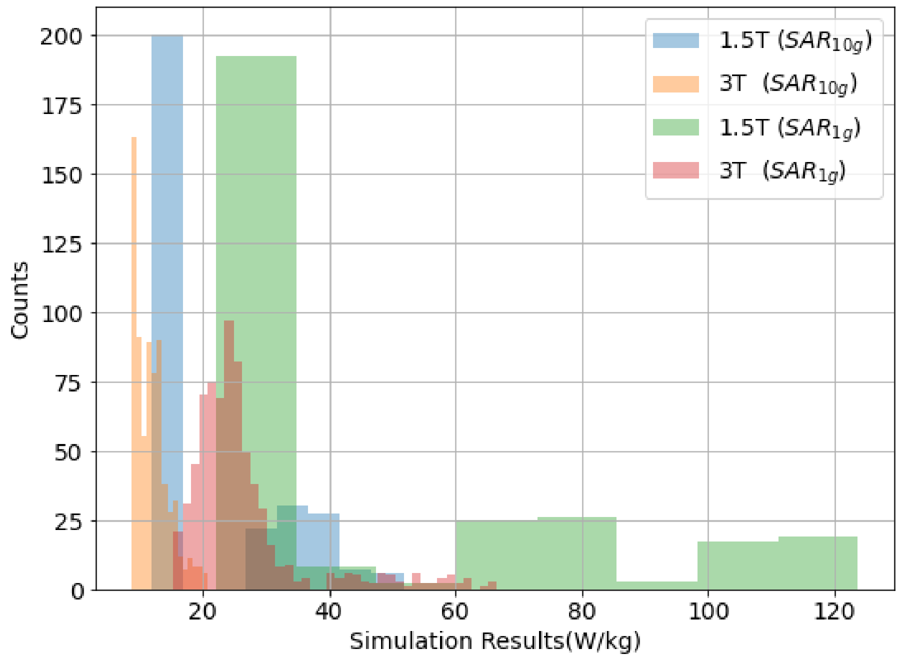


Figure 45. Simulation results for the complex external fixators.

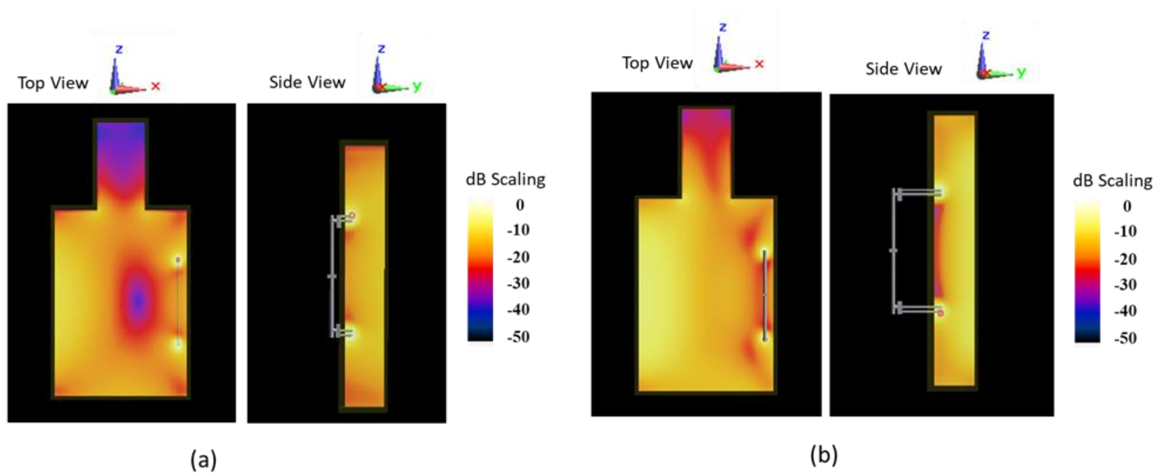


Figure 46. Illustration of worst-case RF-induced heating of complex external fixator at (a) the 1.5 T system and (b) the 3 T system.

The worst-case RF-induced heating for the complex external fixators was shown in Figure 46. Typically, the hot spots would occur at the location that was close to the tip of the screw. The statistics of the simulation results for the complex external fixator were shown in Table 19. The worst-case SAR_{10g} for a complex external fixator device at 3 T is

20.86 W/kg. The corresponding worst-case configuration has a rod length of 280 mm, an insertion depth of 20 mm, a frame distance of 100 mm, a screw radius of 3 mm, and a screw spacing of 10 mm. The worst-case SAR_{10g} of the complex external fixator device at 1.5 T system is 51.72 W/kg. The corresponding worst-case configuration has a rod length of 280 mm, an insertion depth of 20 mm, a frame distance of 40 mm, a screw radius of 3 mm, and a screw spacing of 10 mm. The worst-case SAR_{1g} for a complex external fixator device at 3 T is 66.59 W/kg. The worst-case SAR_{1g} of the complex external fixator device at 1.5 T system is 123.68 W/kg. The worst-case configurations of the RF-induced heating quantified SAR_{1g} was the same as the worst-case configurations quantified by SAR_{10g}.

Table 19. Statistics of the simulation results for the complex external fixator.

Statistics	**SAR_{10g}**		**SAR_{1g}**	
	3 T	**1.5 T**	**3 T**	**1.5 T**
Min	8.74 (W/kg)	11.79 (W/kg)	15.42 (W/kg)	22.12 (W/kg)
Max	20.86 (W/kg)	51.72 (W/kg)	66.59 (W/kg)	123.68 (W/kg)
Mean	11.95 (W/kg)	21.26 (W/kg)	26.52 (W/kg)	47.35 (W/kg)
Variance	6.77 (W^2/kg^2)	120.64 (W^2/kg^2)	90.83(W^2/kg^2)	940.89 (W^2/kg^2)

The ANN results (SAR_{10g}) at 3 T system with a different step size of complex external fixator were shown in Figure 47. The network can get low prediction errors when using a minimum step size of lambda/10 as the training data set. The correlation coefficient of the ANN was larger than 0.98 for all the testing data. The MAPE of the ANN was less than 1.89% and the worst-case prediction error was less than 0.55%. The network still can be used to predict the worst-case RF-induced heating even using a minimum step size of lambda/8 as the training set. The worst-case prediction error was less than 3.80% and the overall prediction error was less than 2.53%. The performance of the network continued decreasing if the step size was larger. The predicted RF-induced heating will deviate far away from the ground-truth values if the network was trained by a step size of lambda/5. Although the overall RF-induced heating in terms of SAR_{1g} was higher than SAR_{10g}, the neural network still performs well on predicting the overall heating and the worst-case heating if using proper training data set as shown in Figure 48. It can be indicated that the neural network can be used to predict the overall SAR_{1g} and worst-case SAR_{1g} if it was trained by a step size less than lambda/5. Otherwise, the worst-case SAR_{1g} prediction errors will be larger than 8.61%. Therefore, the worst-case heating can be accurately predicted by the network using the step size of lambda/8 as the training set at 3 T.

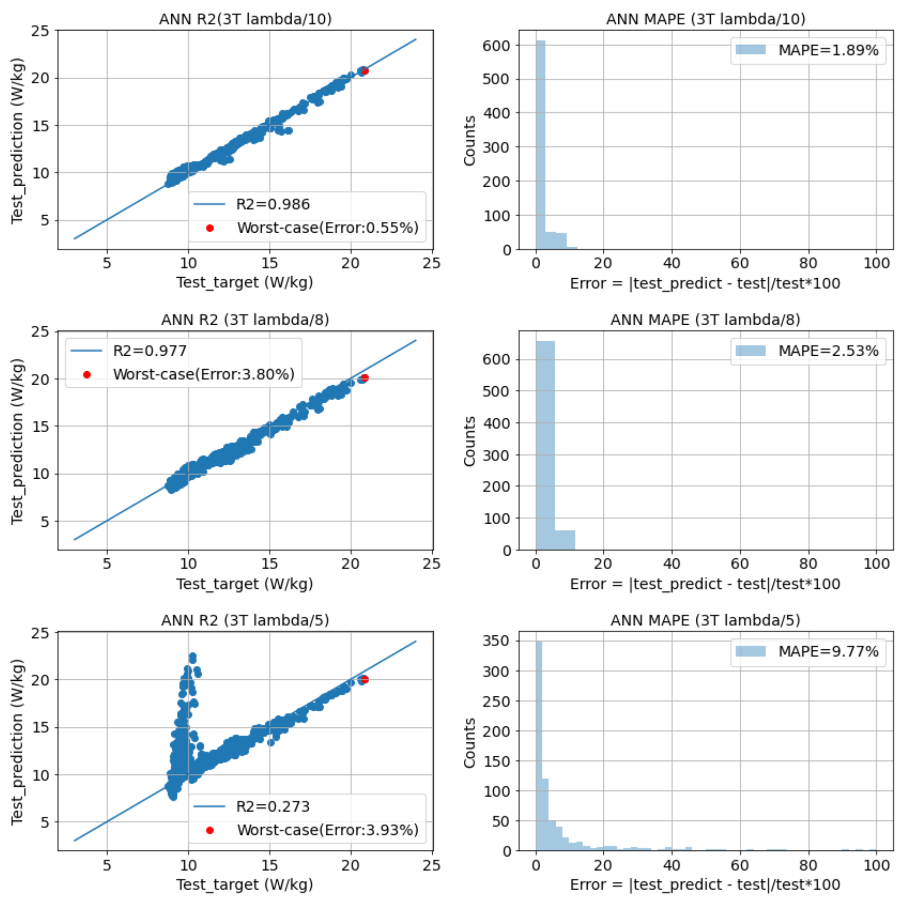


Figure 47. The NN testing results (SAR_{10g}) were trained by different step sizes for complex external fixator at 3 T system.

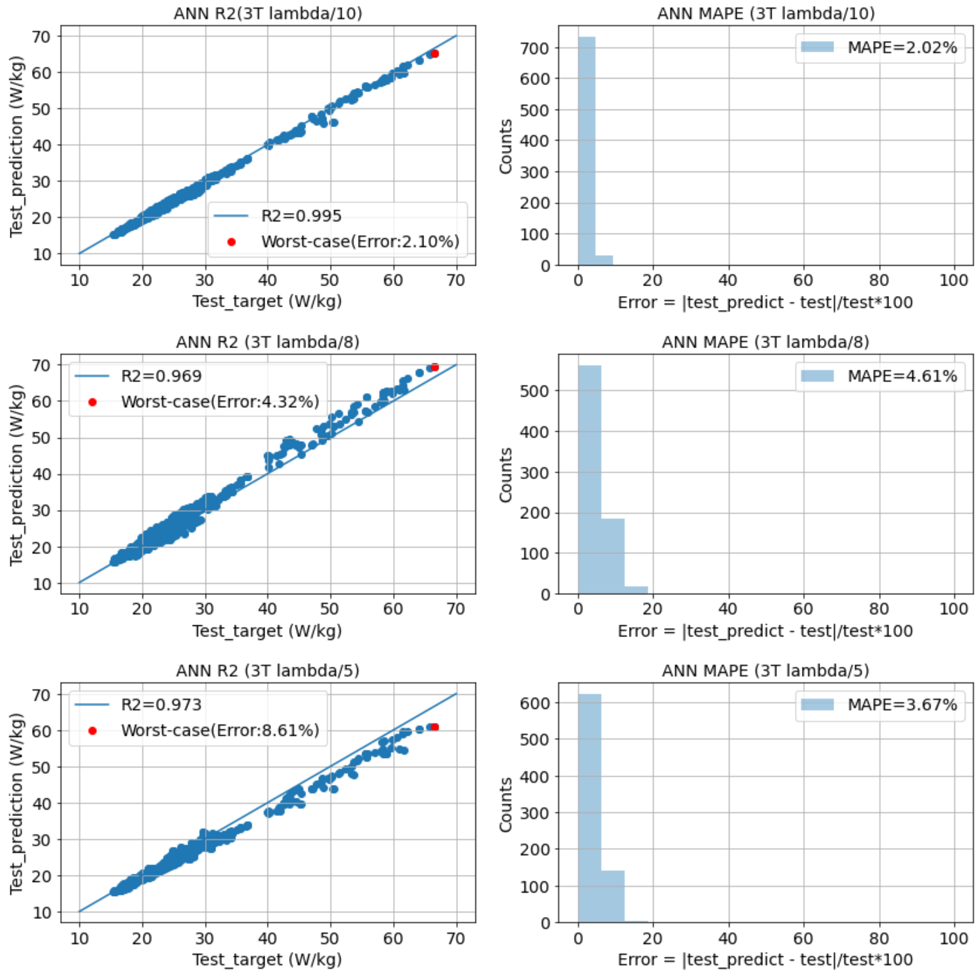


Figure 48. The NN testing results (SAR_{1g}) trained by different step sizes for complex external fixator at 3 T system.

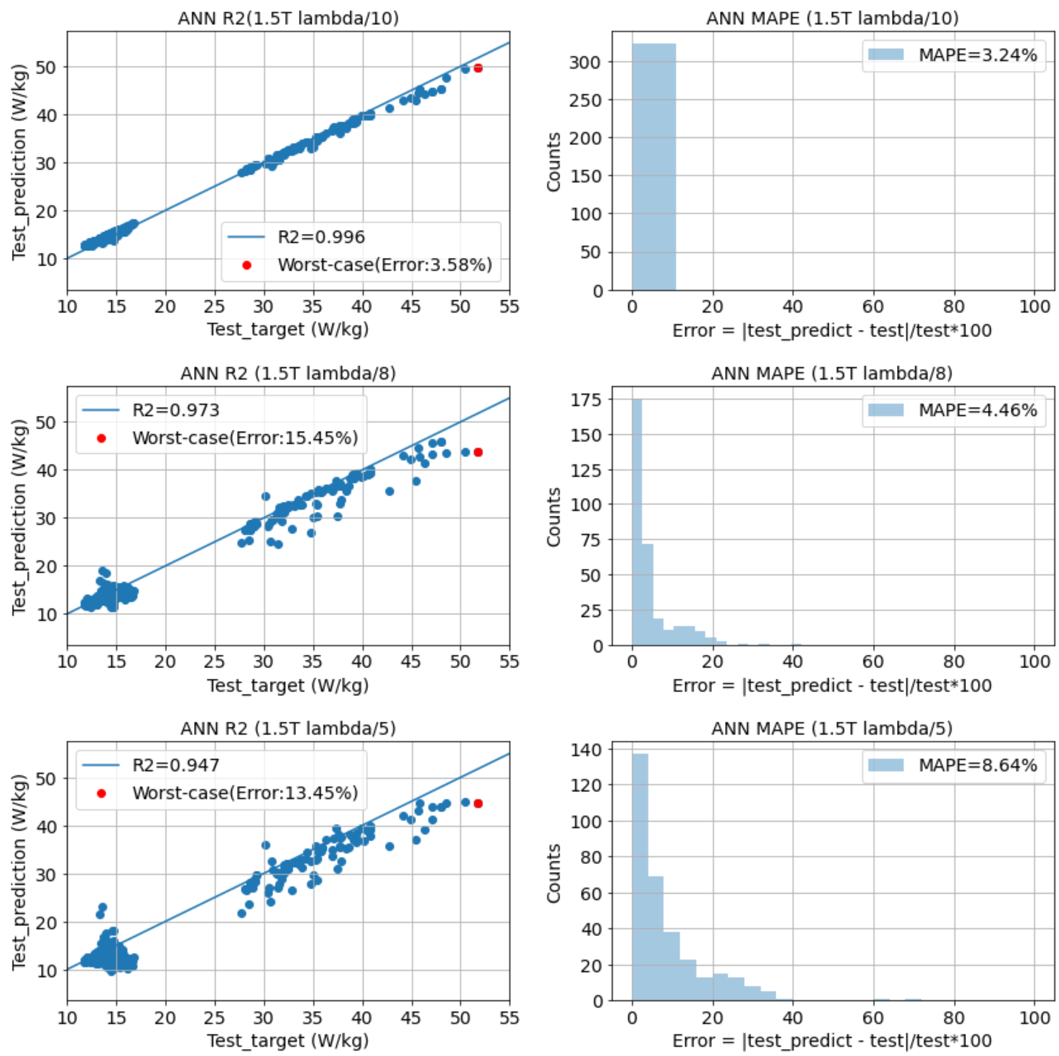


Figure 49. The NN testing results (SAR_{10g}) were trained by different step sizes for complex external fixator at 1.5 T system.

The ANN results (SAR_{10g}) at a 1.5 T system with different step sizes of complex external fixator were shown in Figure 49. The ANN results at the 1.5 T system show that the correlation coefficient for all the test cases was larger than 0.99 by using a training set with lambda/10 step size. The overall MAPE was small which was less than 3.24%. Furthermore, the worst-case prediction error was less than 3.58%. However, the performance of the network was not guaranteed when trained by the larger step size of lambda/8 at 1.5 T. The worst-case predicted SAR_{10g} will as be high as 15.45% and will be much lower than the actual worst-case from simulations. The ANN results (SAR_{1g}) at a 1.5 T system with different step sizes of complex external fixator were shown in Figure 50. The worst-case prediction error for SAR_{1g} was smaller than 1.6% and the overall MAPE was less than 1.94% if the ANN was trained by using a step size of lambda/10. The worst-case prediction error for SAR_{1g} will be larger than 12.88% if the step size was lambda/8. Thus, the criterion of the lambda/10 still can be applied to the complex external fixator both at 1.5 T and 3 T system. Otherwise, the predicted RF-induced heating might be underestimated if the network was trained by a step size larger than lambda/10.

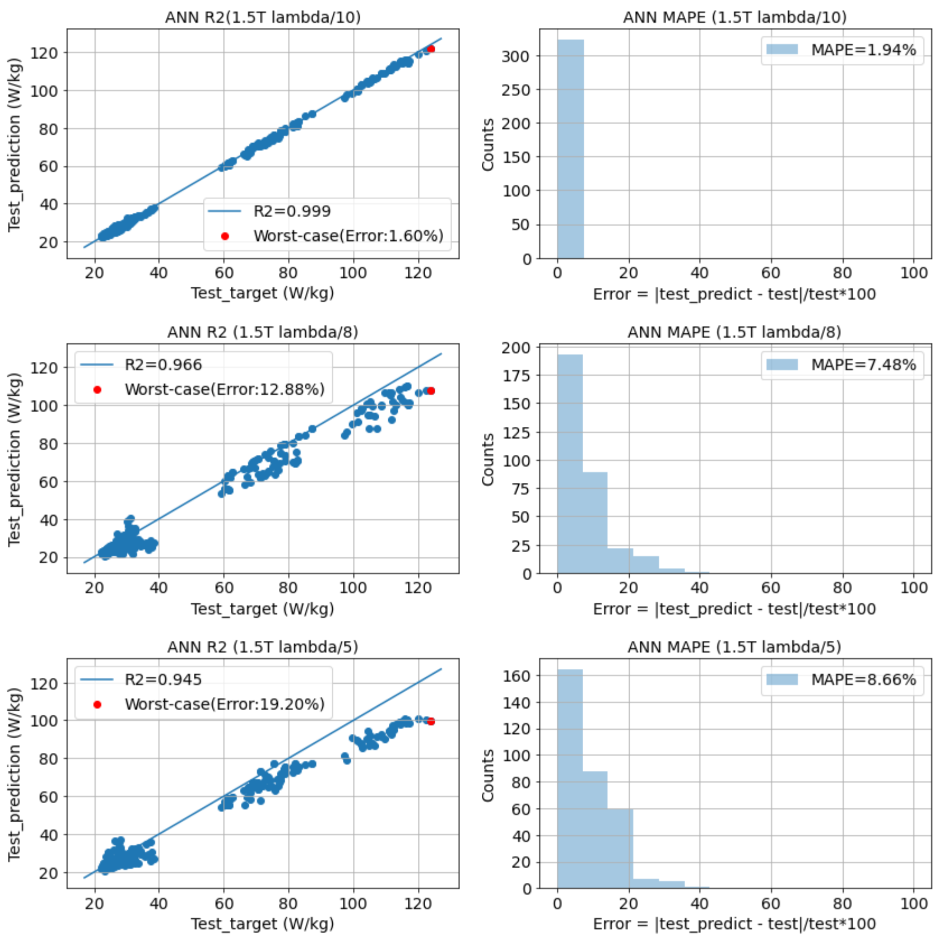


Figure 50. The NN testing results (SAR_{1g}) trained by different step sizes for complex external fixator at 1.5 T system.

To study the general performance of the network to predict the whole sample space using the smallest number of training data sets for the complex external fixator. The trained network with a converged step size was used to predict the RF-induced heating for device parameters in the continuous sample space. There are three parameters, the rod length, the insertion depth, and the frame distance in the continuous space. The other parameters keep unchanged when study one selected parameter. The errors between the prediction results and the simulation results were shown in Figure 51. The largest predicted error was less than 6% for all the study parameters at the 3 T system. The largest predicted error was less than 13% for all the studies of parameters at the 1.5 T system. The RF-induced heating of the complex external fixator with the variation of the insertion depth in y-direction induced larger errors compared to other parameters.

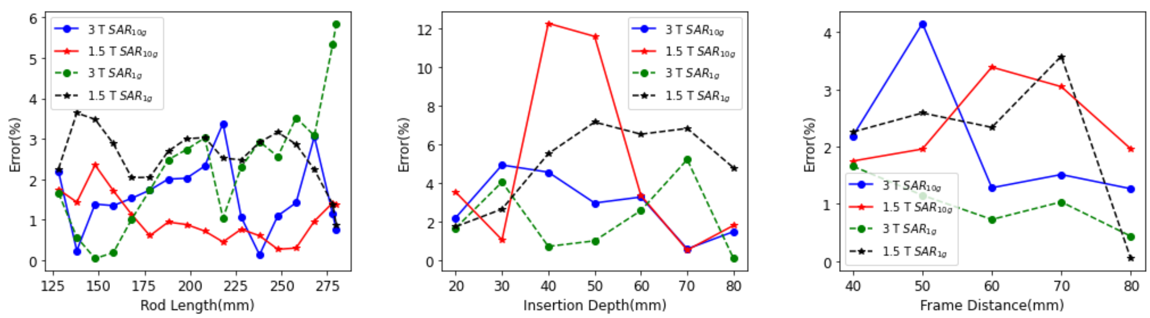


Figure 51. NN prediction errors for the complex external fixator with a fine step of 10 mm.

For the complex external fixator, the rod length, the insertion depth, and the frame distance can be reduced to the smaller sample space. It's recommended to use a converge step size of lambda/10 to get more accurate prediction results. Therefore, the gain of the network (G_{NN}) for the study of complex external fixator at 1.5 T system is $G_{NN(1.5T)} = \left[1 - \frac{1}{(24)^3}\right] \approx 99.99\%$. The gain of the network (G_{NN}) for the study of complex external fixator at 3 T system is $G_{NN(3T)} = \left[1 - \frac{1}{(44)^3}\right] \approx 99.99\%$.

8 Mesh-based Convolutional Neural Network

Many structures cannot be described using simple parameters. Based on the idea that electromagnetic simulations will use meshes in the simulations, we study the possibility of predicting the results of RF-induced heating use the device meshes with the CNN network.

8.1 Convolutional Neural Network

To accurately predict the MRI RF-induced heating and cater to clinical requirements, one specific type of NNs, convolutional neural network (CNN) was developed in this study to model the MRI RF-induced heating for the devices that cannot be parameterized using simple dimensional features. The CNN is an end-to-end architecture that can automatically capture the geometrical features without the need to get hand-craft features (length, width, thickness, etc.) as the input. Another benefit is that it can handle devices with various geometrical, spatial changes, and capture these fine features on the surface of the object. Thus, it will allow us to predict the various complex shape of medical implants without any substantial contradictions.

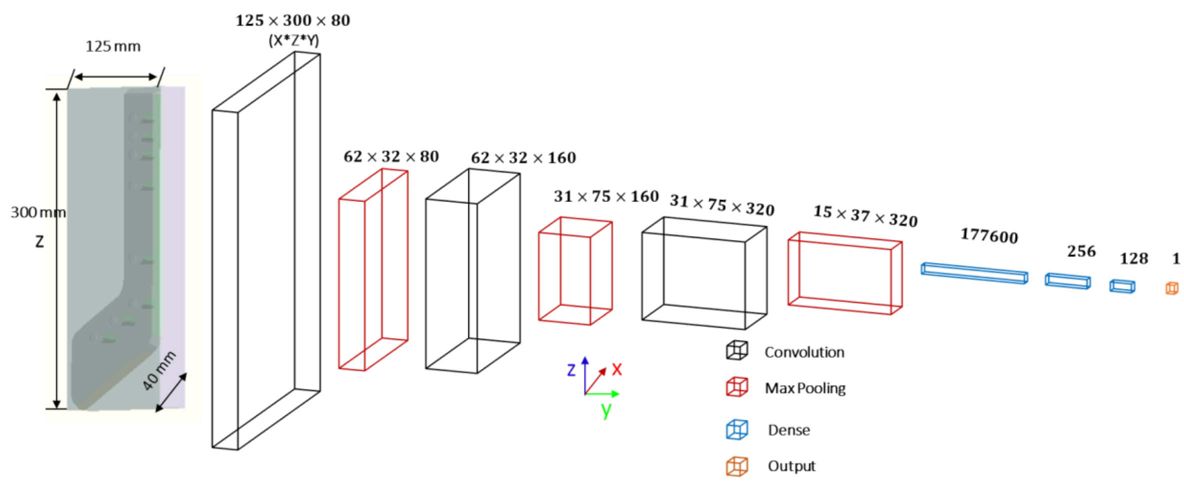


Figure 52. CNN architecture of RF-induced heating fast prediction for complex shape medical devices.

The architecture of CNN is case sensitive and depending on the data size, input, training algorithm, number of layers, etc. The architecture was similar to AlexNet [42] and implemented using the Tensorflow framework [43]. After several analyses, a convolutional neural network that had three convolutional layers with a 3×3 filter for each and three dense layers was used to predict the $psSAR_{10g}$ of the multi-configuration and complex shape plate devices as shown in Figure 52. Input will hold the raw data of the device, consists of layers of slices that have depth in the y-direction), width in the x-direction, and length in the z-direction. The convolution layer will compute the output of neurons that are connected to local regions in the input, each computing a dot product between their weights and a small region (in x-z plane/slice) they are connected to the input volume. The filter was used to find the spatial correlation in a sliced layer of image and taken as the receptive field. The detailed spatial information was captured by this receptive field after the convolutional operation. The first convolution layer was responsible for capturing low-level features such as edges, corners, shapes of the 3D device from filters. Each filter corresponding to one feature map obtained from convolution.

The max-pooling layer will perform a down-sampling operation along the spatial dimension for each slice in length and width. This also has the effect of making the resulting max-pooling feature maps more robust to changes in the position of the feature in the image. The max-pooling will take the most activated presence of a feature in the layers of images. The max-pooling layer was also used for dimension reduction to reduce computational cost which reduces the dimension of images.

With added more convolution layers, the CNN was adapted to the high-level features, which can capture the precise 3D geometrical representations of the device. This will increase the number of feature maps gradually by twice for each layer. The dense layer with relu activation function was added to learn non-linear combinations of the high-level features as represented by the output of the convolution layer. In this way, the convolution layers could automatically capture the 3D geometrical information which covers geometrical dimensions, shape, and material information.

Three dense layers formed a feed-forward neural network and backpropagation were applied to every iteration of the training process. To keep the features learned from previous convolutional layers which contain precise geometrical information, the size of the dense layers was gradually decreased to one dimension to get the $psSAR_{10g}$. Over a series of training iterations, the CNN was able to predict the $psSAR_{10g}$ with low-level features in the form of sliced images. The testing data are used to examine the validity of the results predicted by CNN.

For the mesh-based CNN, the 3D mesh can be sliced into 2D meshes in one specified direction. Thus, the 2D meshes can be used as the input of the CNN as shown in Figure 53. To cover all different dimensions of the devices, a large box needs to be defined and has the largest dimension. Then, the 3D mesh can be sliced into layers of 2D meshes according to the smallest resolution Δr and these 2D meshes contains detail geometrical features of the devices even when the geometrical features are small. According to the defined box size, a total number of 100 slices (y-direction), each slice is 100 mm x220 mm was studied. For the volume-based estimation, the resolution could be 1 mm or 2 mm since most plates would be thicker than 3mm and the screw size was more than 2 mm in diameter. The boundary edge-based meshes, another form of input by getting the difference for adjacent points for each sliced layer can also be used as the input of the CNN.

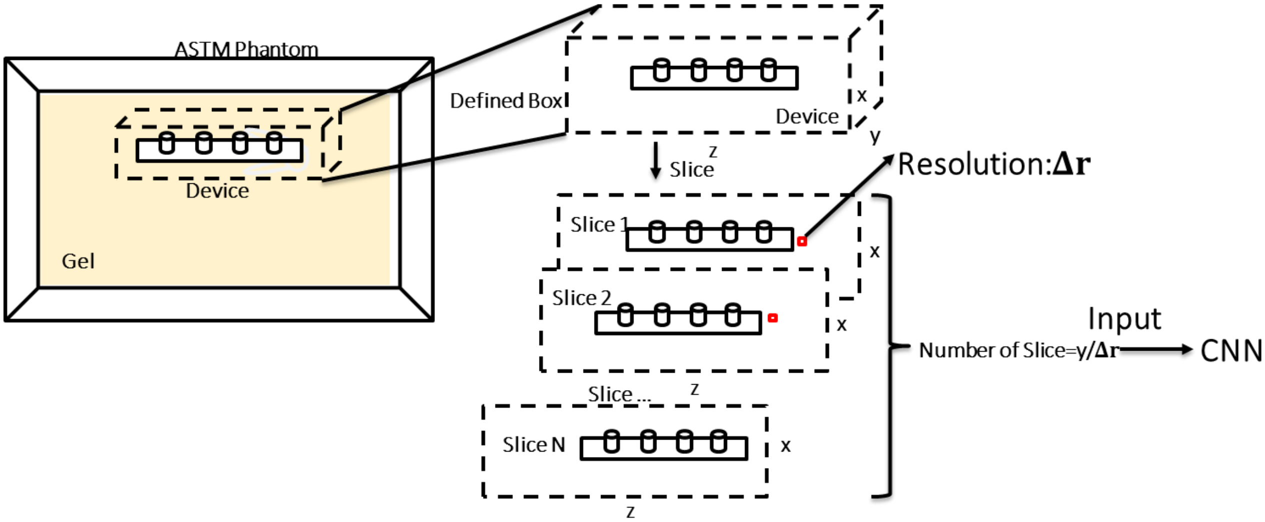


Figure 53. Get input for the CNN.

8.2 Application of General Compression Plate using Mesh-based CNN

8.2.1 Training Set

To study the mesh-based CNN, the same training set for the ANN at the 3 T system was used. These training sets with different step sizes will be used to train the CNN. Once it has been trained, it can be validated using the test data which was the whole data set from simulation to predict the RF-induced heating.

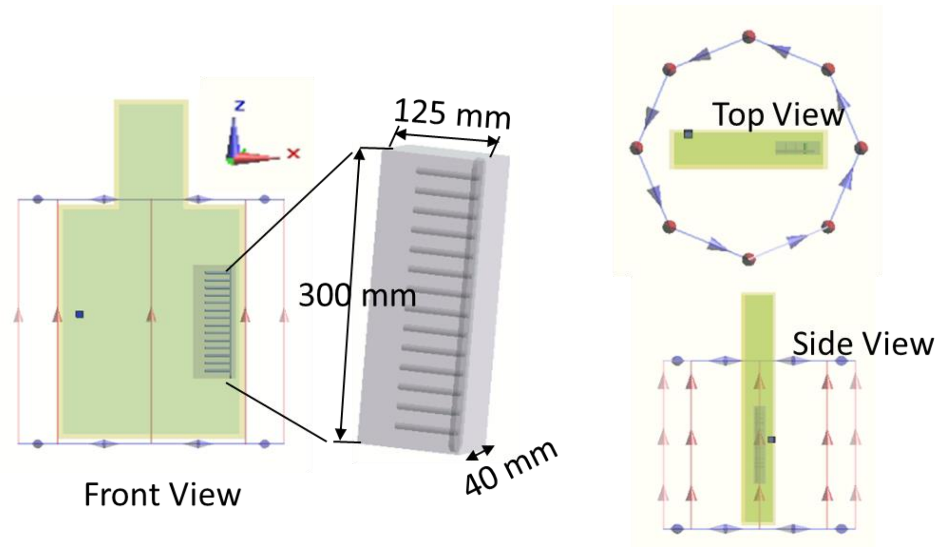


Figure 54. Compression plate used for mesh-based CNN.

The compression plate devices in 3D will be sliced into layers of images. To reserve the geometrical dimension, a box with a geometrical dimension that can cover all the configurations was first defined as shown in Figure 54. In this study, the box dimension is $125 \times 40 \times 300$ ($x \times y \times z$). The device in 3D will be sliced into layers of images in the x-z plane in a y-direction with a step size of 1 mm. Therefore, there are a total number of 40 layers for each complex shape device. Each layer covers an x dimension of 125 mm and a z dimension of 300 mm. The image plots (x-z plane) that some layers will include the devices, and other layers include the screws were shown in Figure 55. The details in a sliced layer of the compression plate device for volume-based estimation was shown in Figure 56. The materials type of PEC will be defined as 1 and other materials will be defined as 0.

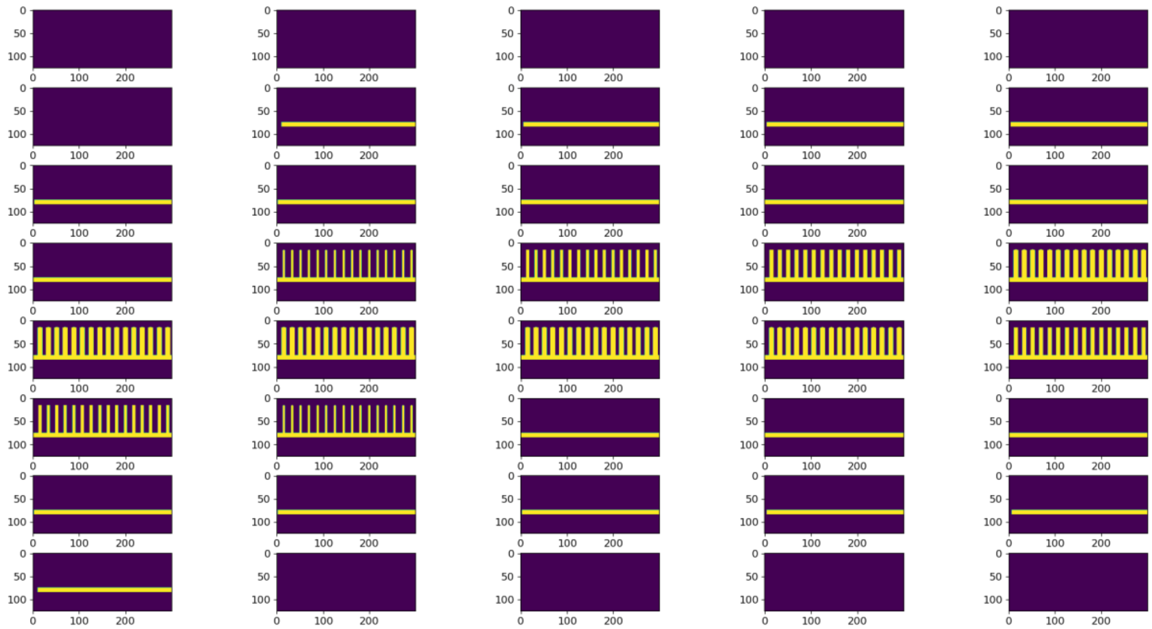


Figure 55. Forty-layers of images in y-direction for each complex shape device.

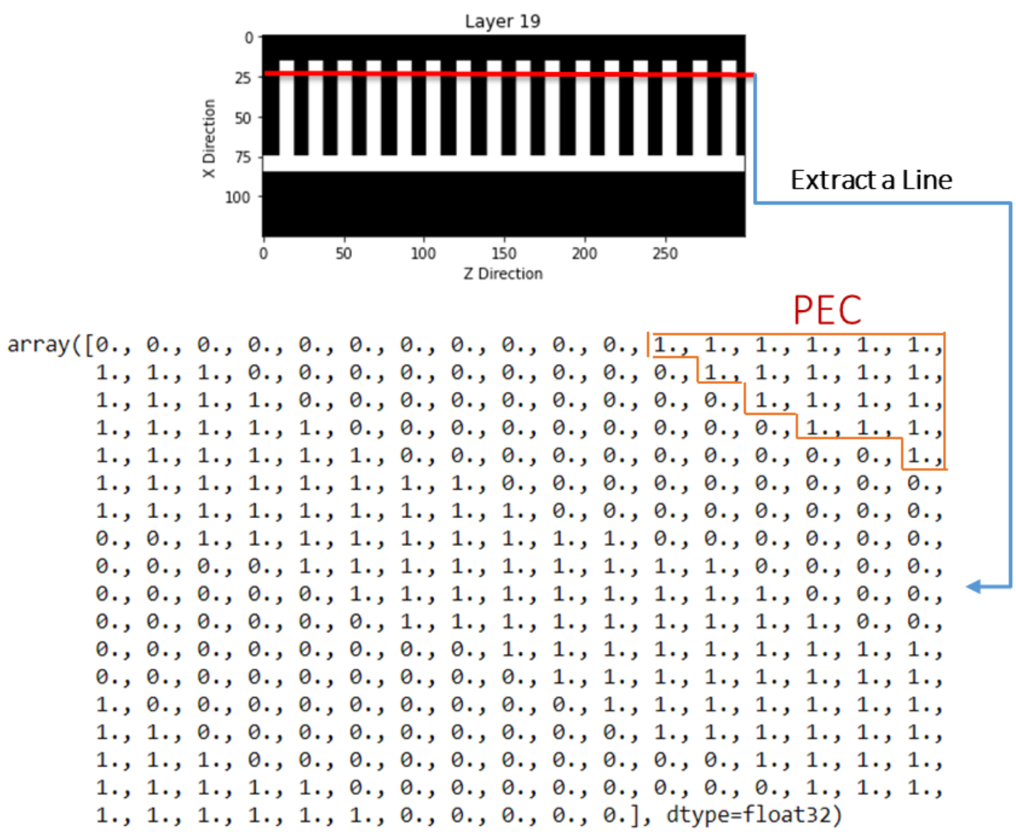


Figure 56. The details in a sliced layer of the compression plate device for volume-based estimation.

Each sliced layer of the compression plate device could be further processed by getting the difference between two adjacent points. In this way, the edge representation for the device was obtained for the input of the CNN. The details in a sliced layer of the compression plate device were shown in Figure 57. The materials type of PEC will be defined as 1 or -1, and other materials will be defined as 0.

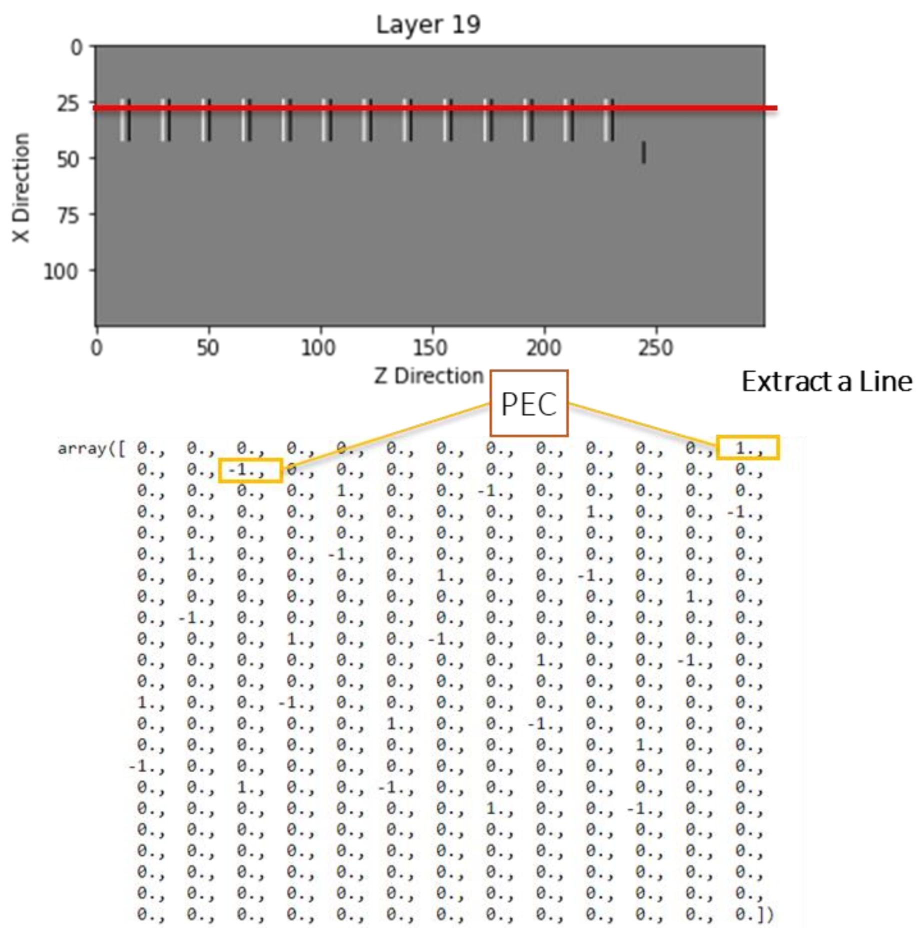


Figure 57. The details in a sliced layer of the compression plate device for edge-based estimation.

8.2.2 Results

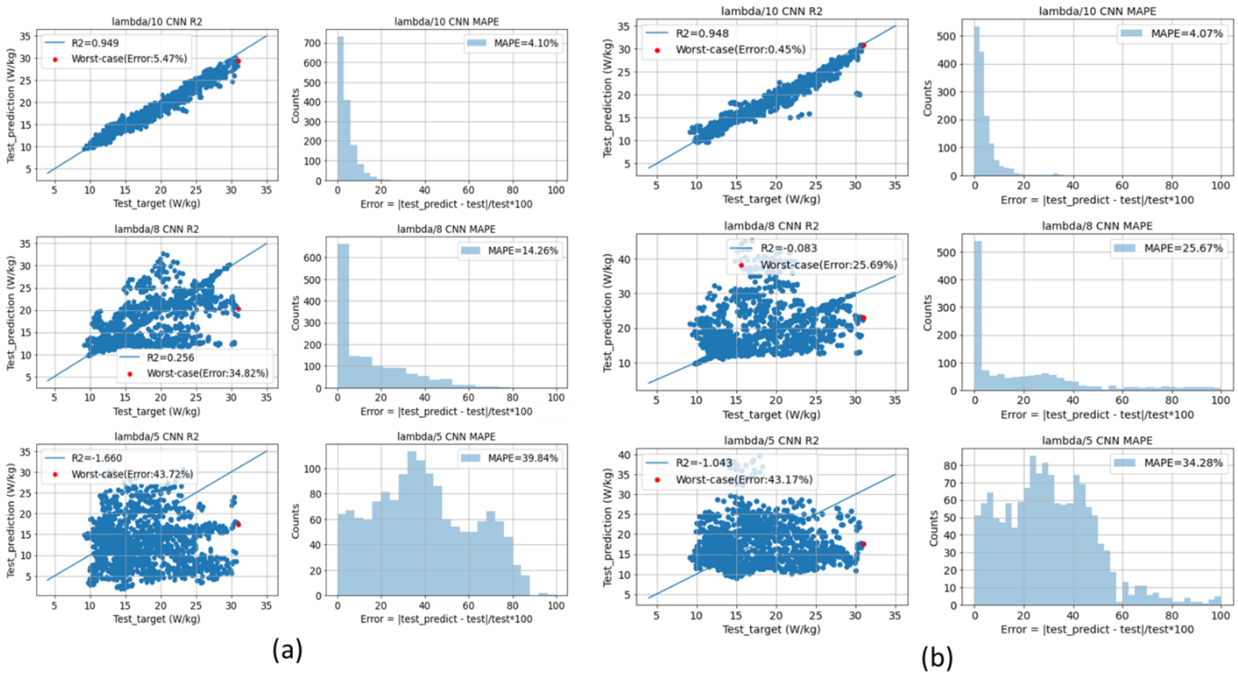


Figure 58. The mesh-based CNN testing results (SAR_{10g}) trained by different step sizes for the compression plate system at 3 T system, (a) volume-based estimation, and (b) boundary edge-based estimation.

To get the comprehensive performance comparison between the mesh-based CNN and the parameterized NN, the same training sets from the study of the parameterized NN were used for the mesh-based CNN. The mesh-based CNN testing results (SAR_{10g}) both for the volume-based estimation and boundary edge-based estimation at the 3 T system for the compression plate system were shown in Figure 58. For the volume-based estimation, the CNN trained by lambda/10 step size can predict the worst-case heating with an error rate of less than 5.47%. The boundary edge-based estimation had a worst-case prediction error of 0.45%. The MAPE of the CNN was less than 4.1% both for the volume-based estimation and boundary edge-based estimation. It can be indicated that the minimum step size to predict the worst-case and all other configurations should be less or equal to lambda/10. The worst-case prediction error was larger than 14.22% if the CNN was trained with a step size of lambda/8. The overall prediction error was larger than 32.37% when the step size was lambda/5.

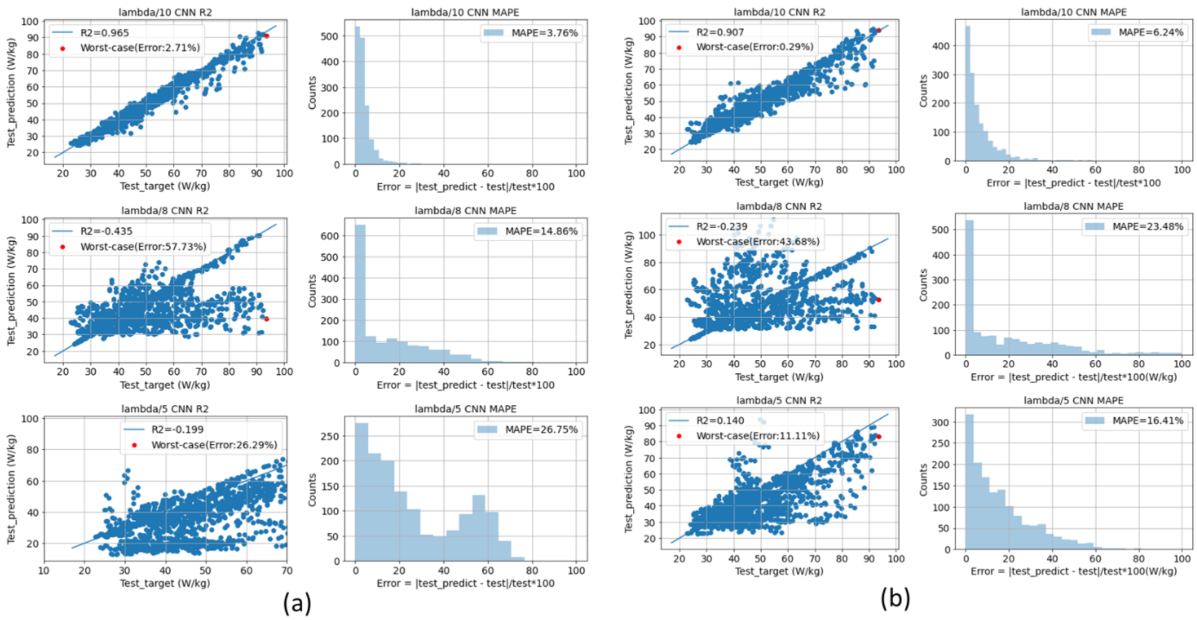


Figure 59. The mesh-based CNN testing results (SAR_{1g}) trained by different step sizes for the compression plate system at 3 T system, (a) volume-based estimation, and (b) boundary edge-based estimation.

The mesh-based CNN testing results (SAR_{1g}) both for the volume-based estimation and boundary edge-based estimation at the 3 T system for the compression plate system were shown in Figure 59. It can be indicated that the lambda/10 criterion can be applied to predict the worst-case SAR_{1g} and overall SAR_{1g} at the 3T system using the CNN. The worst-case prediction error for the volume-based estimation was less than 2.71% and the MAPE was less than 3.76%. The worst-case prediction error for the edge-based estimation was less than 0.29% and the MAPE was less than 6.24%. The CNN may not be converged if it was trained by a minimum step size larger than lambda/10.

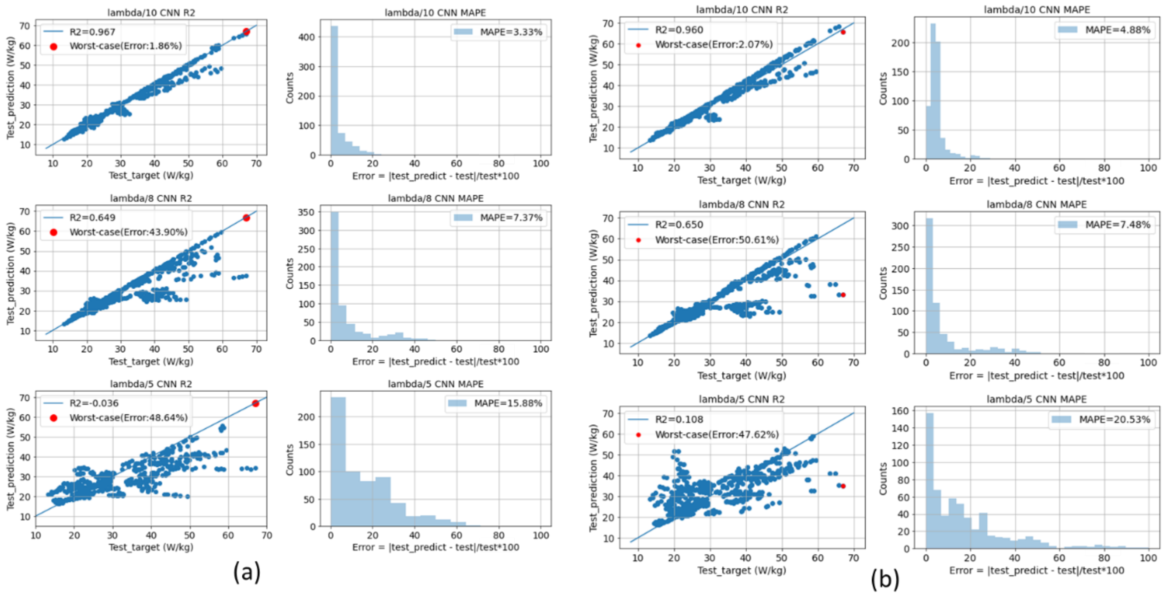


Figure 60. The mesh-based CNN testing results (SAR_{10g}) trained by different step sizes for the compression plate system at a 1.5 T system, (a) volume-based estimation, and (b) boundary edge-based estimation.

The mesh-based CNN testing results (SAR_{10g}) at the 1.5 T system for the compression plate system were shown in Figure 60. The CNN trained by lambda/10 step size can predict the worst case with an error rate of 1.86% for volume-based estimation. The worst-case prediction error for boundary edge-based estimation was smaller than 2.07%. The overall prediction error of the CNN was smaller than 4.88% both for the volume-based estimation and boundary edge-based estimation. In this case, the network has a high correlation coefficient (larger than 0.96) between the input and output. The correlation coefficient of the network was lower than 0.65 and the overall prediction error was larger than 7.37% when trained by a step size of lambda/8. The network might not be converged if it was trained by a minimum step size of lambda/5.

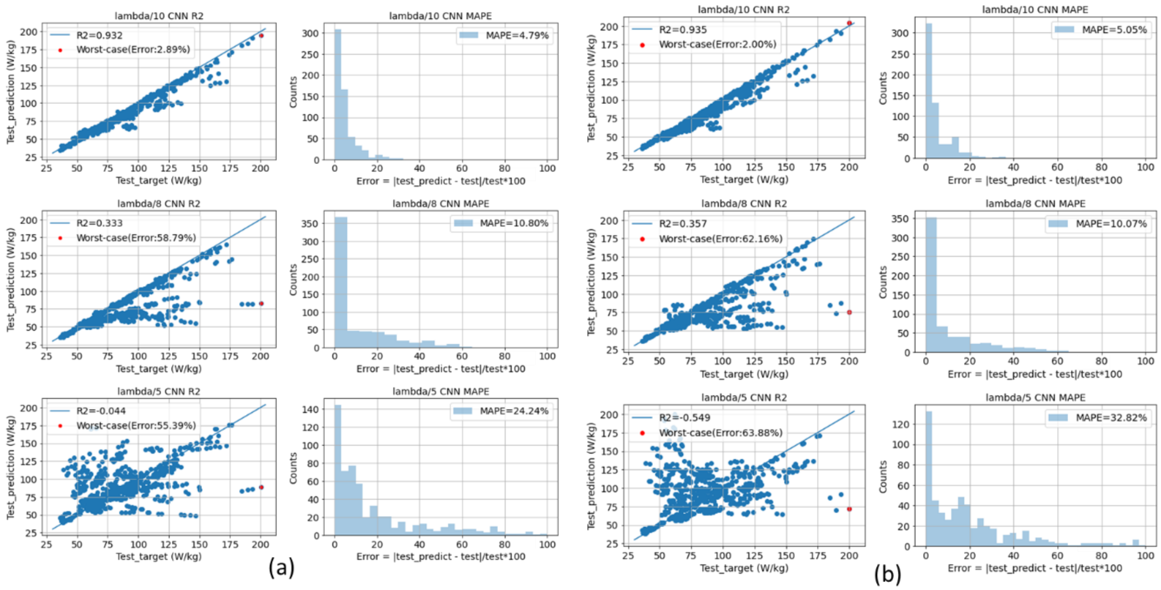


Figure 61. The mesh-based CNN testing results (SAR_{1g}) trained by different step sizes for the compression plate system at a 1.5 T system, (a) volume-based estimation, and (b) boundary edge-based estimation.

The mesh-based CNN testing results (SAR_{1g}) at the 1.5 T system for the compression plate system were shown in Figure 61. The network converged well by using the minimum step size of lambda/10. The correlation coefficient was larger than 0.93 both for the volume-based estimation and boundary edge-based estimation. The worst-case prediction error was less than 2.89% and the MAPE was less than 4.79% for the volume-based estimation. The worst-case prediction error was less than 2.00% and the MAPE was less than 5.05% for the boundary edge-based estimation.

The performance comparison between the ANN and CNN for the prediction of SAR_{10g} was shown in Figure 62. The results indicated that the performance of the ANN was better than the CNN. The prediction error of the ANN tends to be lower than the CNN with the training data set using a step size of lambda/10. In this case, the worst-case prediction error and the MAPE were lower than 6%. The prediction error with the training data set using a step size of lambda/8 tends to be much larger than a step size of lambda/10. The performance of the network was not guaranteed when it was trained by using a step size of lambda/5 as the overall MAPE was very high.

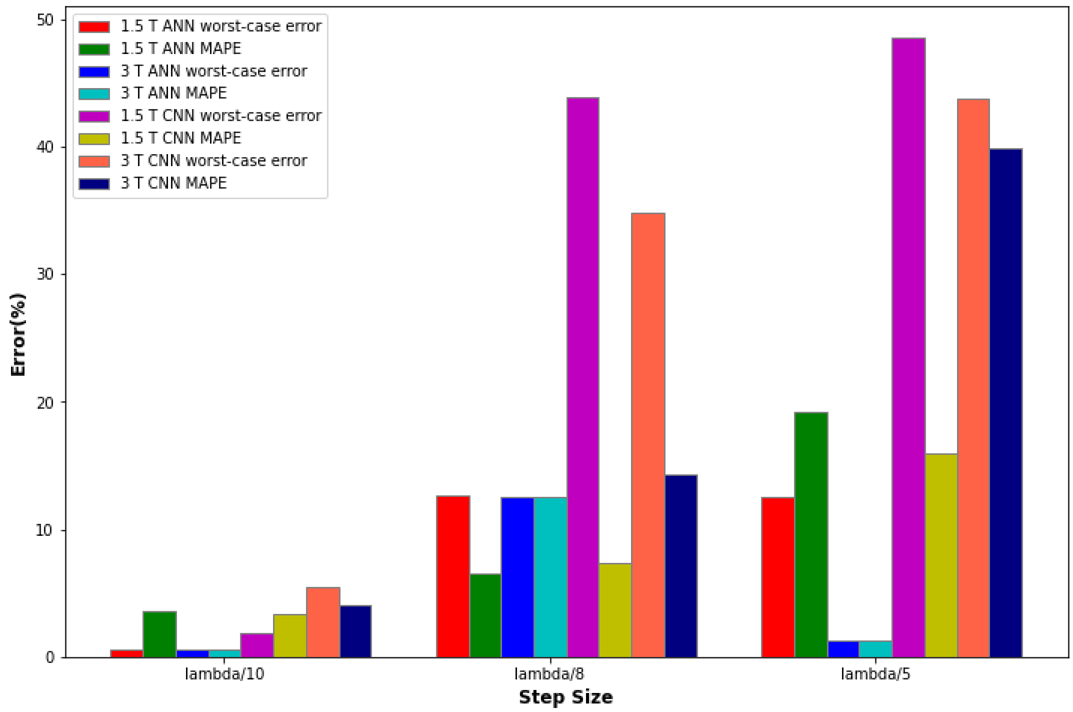


Figure 62. The performance comparison between ANN and CNN for the prediction of SAR_{10g}.

The performance comparison between the ANN and CNN for the prediction of SAR_{1g} was shown in Figure 63. The worst-case prediction error and the MAPE were also lower than 6% by using the minimum step size lambda/10. However, the worst-case error and the overall prediction error will be very large if the networks were trained by a minimum step size larger than lambda/8. Thus, its' recommended to use the minimum step size of lambda/10 for the RF-induced heating fast prediction.

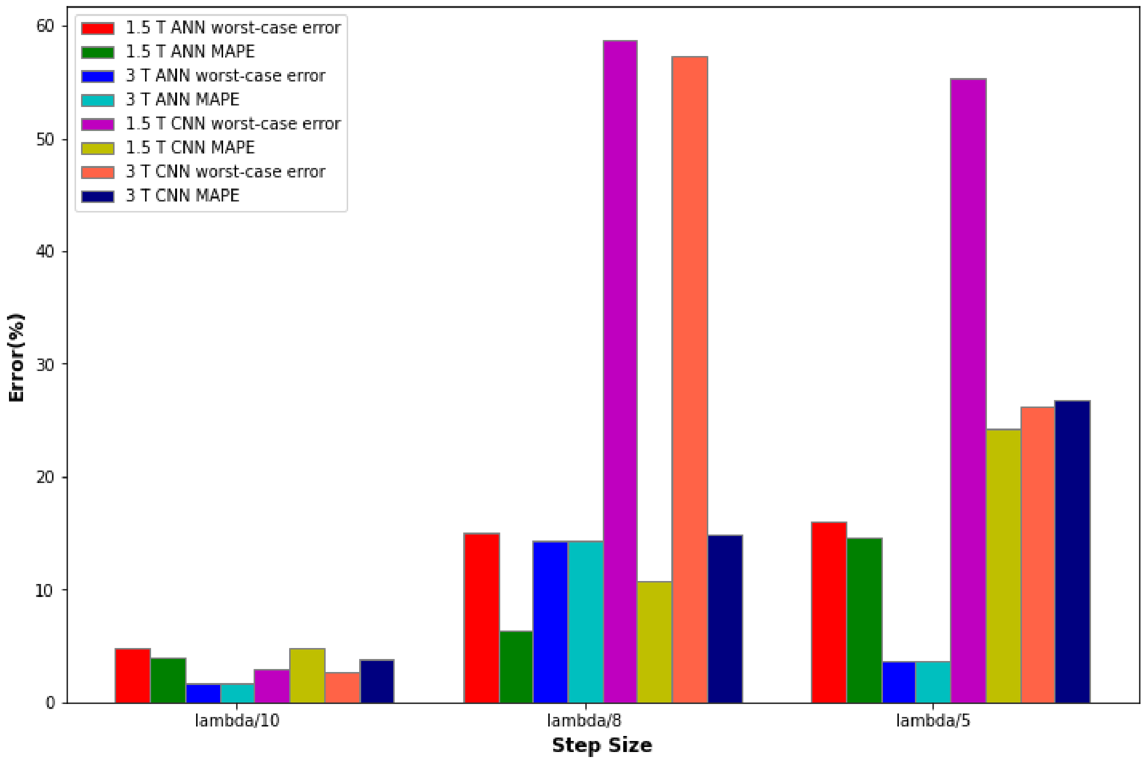


Figure 63. The performance comparison between ANN and CNN for the prediction of SAR_{1g}.

8.2.3 Resolution (Mesh Size) Study

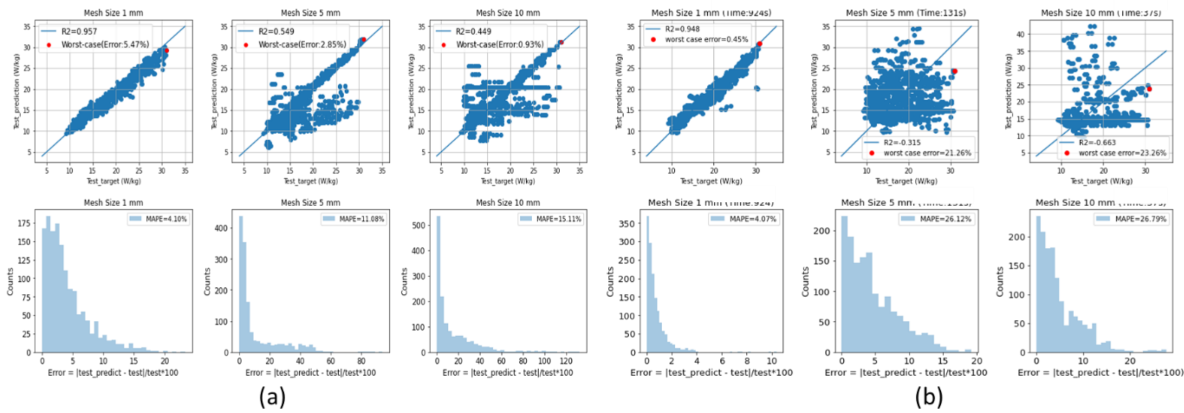


Figure 64. The mesh size study results (SAR_{10g}) for the compression plate system using CNN at 3 T system, (a) volume-based estimation, and (b) boundary edge-based estimation.

The minimum resolution used in CNN that converges and predicts the worst-case was studied. The results (SAR_{10g}) of the mesh size study at 3 T were shown in Figure 64. The correlation coefficient of the CNN will decrease and the overall prediction MAPE will increase with larger mesh size. For the volume-based estimation, the correlation coefficient of the network was lower than 0.54 and the MAPE was larger than 11.09% when trained by a mesh size of 5 mm. The CNN might not be converged if the mesh size was larger than 5 mm. The correlation coefficient was less than 0.45 and the MAPE was larger than 15.11% if trained by a mesh size of 10 mm. The CNN performance used for boundary edge-based estimation shown similar performance. The results (SAR_{1g}) of the mesh size study at 3 T were shown in Figure 65. The worst-case prediction error and the overall prediction error were small by using a mesh size of 1 mm. Otherwise, the network will not converge as the correlation coefficient was smaller than 0.59 and the overall prediction error was larger than 11%. Thus, it is recommended to use a mesh size that was smaller than 5mm to ensure better prediction performance.

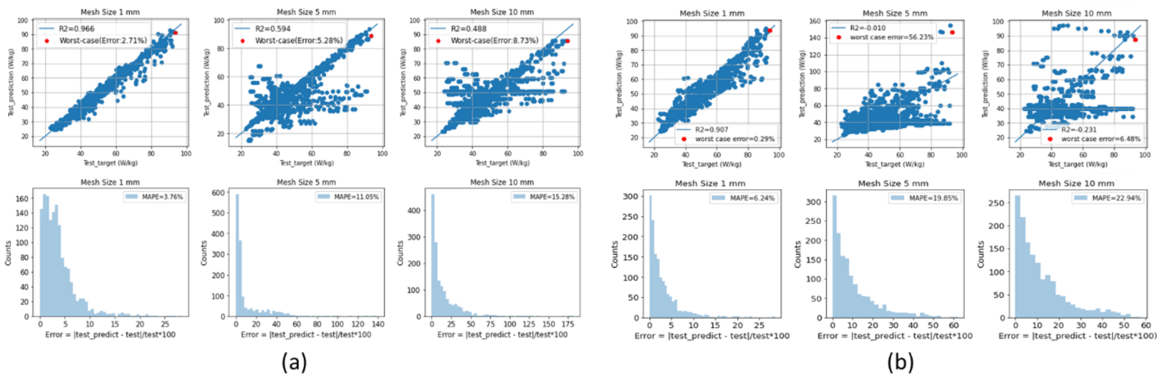


Figure 65. The mesh size study results (SAR_{1g}) for the compression plate system using CNN at 3 T system, (a) volume-based estimation, and (b) boundary edge-based estimation.

The mesh size study results (SAR_{10g}) at the 1.5 T system as shown in Figure 66. For the volume-based estimation, the worst-case prediction error using a mesh size of 5 mm was comparable to a mesh size of 1mm. However, the network would not converge as the correlation coefficient of the network was lower than 0.51 and the worst-case error was larger than 13% when the network used a mesh size of 10 mm. For boundary edge-based estimation, the network might not converge if the CNN trained by using a mesh size of 5 mm as the worst-case prediction error was larger than 14.50%.

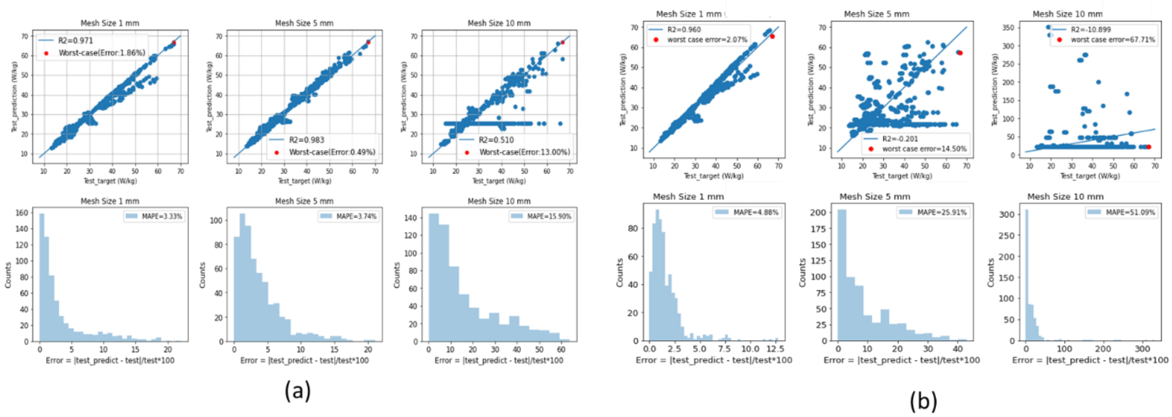


Figure 66. The mesh size study results (SAR_{10g}) for the compression plate system using CNN at 1.5 T system, (a) volume-based estimation, and (b) boundary edge-based estimation.

The mesh size study results (SAR_{1g}) at the 1.5 T system as shown in Figure 67. The performance of the CNN was similar to the mesh size study results of SAR_{10g}. For the volume-based estimation, the worst-case prediction error of a mesh size of 5 mm also was comparable to a mesh size of 1mm. The network would not converge as the correlation coefficient of the network was lower than 0.5 and the worst-case error was larger than 64% when the network used a mesh size of 10 mm. For boundary edge-based estimation, the network might not converge if the CNN trained by using a mesh size of 5 mm as the worst-case prediction error was larger than 13.50%.

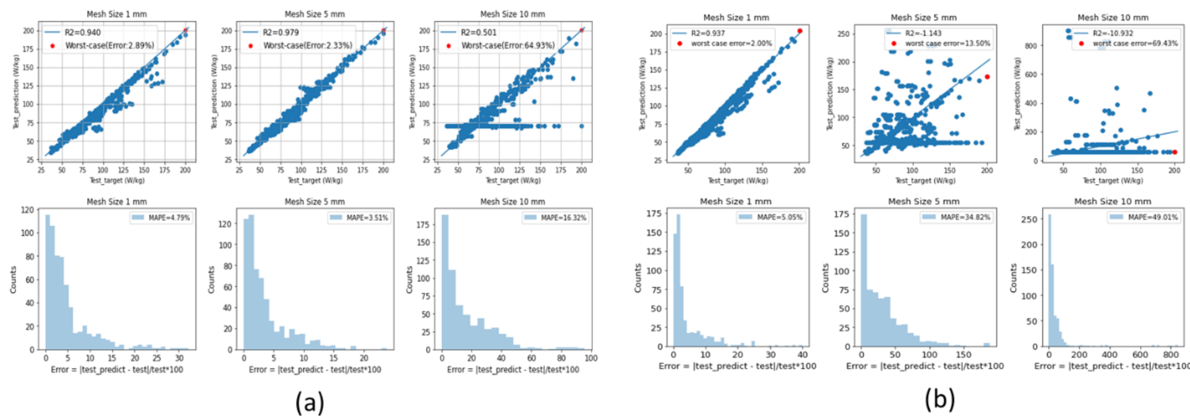


Figure 67. The mesh size study results (SAR_{1g}) for the compression plate system using CNN at 1.5 T system, (a) volume-based estimation, and (b) boundary edge-based estimation.

8.3 Application of Complex Plate System

The complex shape plate system was commonly used among various clinical scenarios for setting and stabilizing fractured bones, such as arm and leg bones. To cater to different fractured bones, the construction of the plate system will be different. Thus, each construction will be varied in size and shape of the plate. For this study, 384 constructions were constructed based on clinical applications.

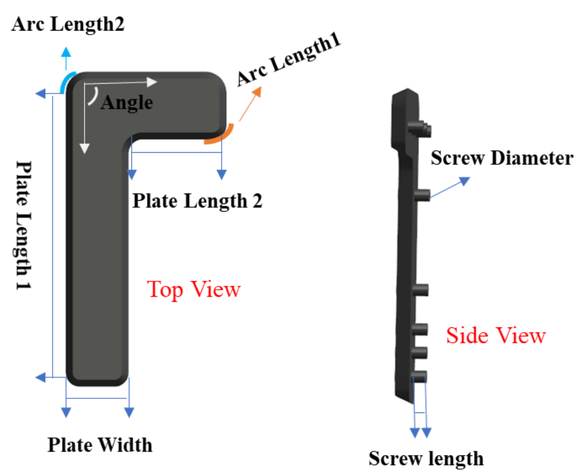


Figure 68.The complex shape plate system construction details.

The commonly used complex shape plate system, which consists of a plate and screws, was created as an example subject as showed in Figure 68. It has been known the device has angle difference, edge difference, length variations, etc. To compare with the parameterized NN, we try parameterized the device. The small variations on the edge or corner were hard to describe, but the arc length was measured in this study. The configuration of the plate has a longer length variation from 200 mm to 220 mm. The configuration of the plate has a longer length variation from 80 to 300 mm, width variation from 25 mm to 35 mm, shorter length variation from 50 mm to 60 mm, arc length variation at the bottom right corner of the shorter plate from 10 mm to 20 mm, and arc length variation at the top left corner of the longer plate from 15 mm to 25 mm. The configuration of the screw is varied with length in the range from 20 mm to 30 mm and diameter in the range from 5 mm to 8 mm. The angle between the shorter plate longer plate can be in the range from 60 degrees to 120 degrees.

Table 20. Device dimension of the complex plate system used of mesh-based CNN study.

Parameters	**Device Dimension(mm)**
Plate Length 1	[200-220] mm
Plate Width	[25,35] mm
Plate Length 2	[50,60] mm
Arc Length 1	[10,20] mm
Arc Length 2	[15,25] mm
Screw Length	[20,30] mm
Screw Diameter	[5,8] mm
Angle	$[60°, 90°, 120°]$

8.3.1 Numerical Simulation

For the complex shape plate system, numerical simulations were conducted at 3

Tesla(T) using full-wave electromagnetic solver based on FDTD method to get the RF-induced heating, in terms of peak 10 gram (g) averaged specific absorption rate ($psSAR_{10g}$) in the ASTM phantom. In this *in-vitro* way, the data collection phase for NNs will save more time for this study since the NN has been validated for in-vitro RF-induced heating fast evaluation. A high pass non-physical RF transmits body coil was adopted to model the MRI RF body coil operated at 128 MHz for a 3 T system. This non-physical coil model has been widely used and validated as in [44]. The RF coil was loaded with a model of the ASTM phantom. Eight current sources were placed on the rungs of the coil to generate a uniform magnetic field inside the coil. Absorbing boundary conditions were used on all sides of the simulation boundaries.

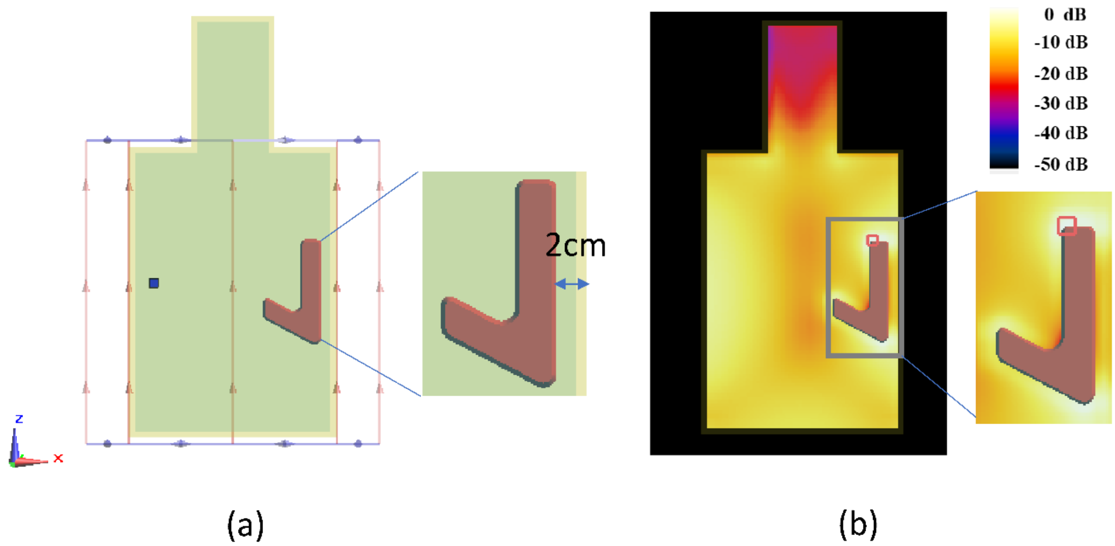


Figure 69. (a) The device was placed at the vertical center on the right side 2cm away from the phantom wall and at the center along the bore direction. (b) Example of RF-induced heating under 3 T system.

The ASTM phantom was a plastic container with a relative dielectric constant $\epsilon_r =$ 3.7 and an electrical conductivity $\sigma = 0$ S/m, filled with gelled-saline, which had the $\epsilon_r =$

3.7 and $\sigma = 0.47$ S/m. The plate device was placed at the vertical center on the right side 2 cm away from the phantom wall and at the center along the bore direction (the location which provides maximum and uniform electric field-induced heating inside the phantom) as shown in Figure 69. An example of the RF-induced heating of the complex shape plate device under the 3 T system showed that the hot spot occurred at the end of the plate. In the numerical simulation, all metallic materials were modeled as perfect electric conductor (PEC).

The non-uniform mesh was used in the simulations to approach the balance between accuracy and complexity because the size of the coil, phantom, and devices was different. The larger mesh step size can reduce the total simulation time, but the coarse mesh cannot represent the device structure. The smaller mesh steps unbearable computational burdens and the divergent results. It was determined with several convergence analyses that the mesh size of 0.5/1 mm was applied to the plate devices. The mesh size of the gelled-saline was 5 mm and the plastic box was 10 mm. The grating ratio of the mesh size was set to 1.15. To ensure convergence, the simulation time was set for 25 periods. All the results were normalized to a whole-body SAR of 2 W/kg.

8.3.2 Training Set

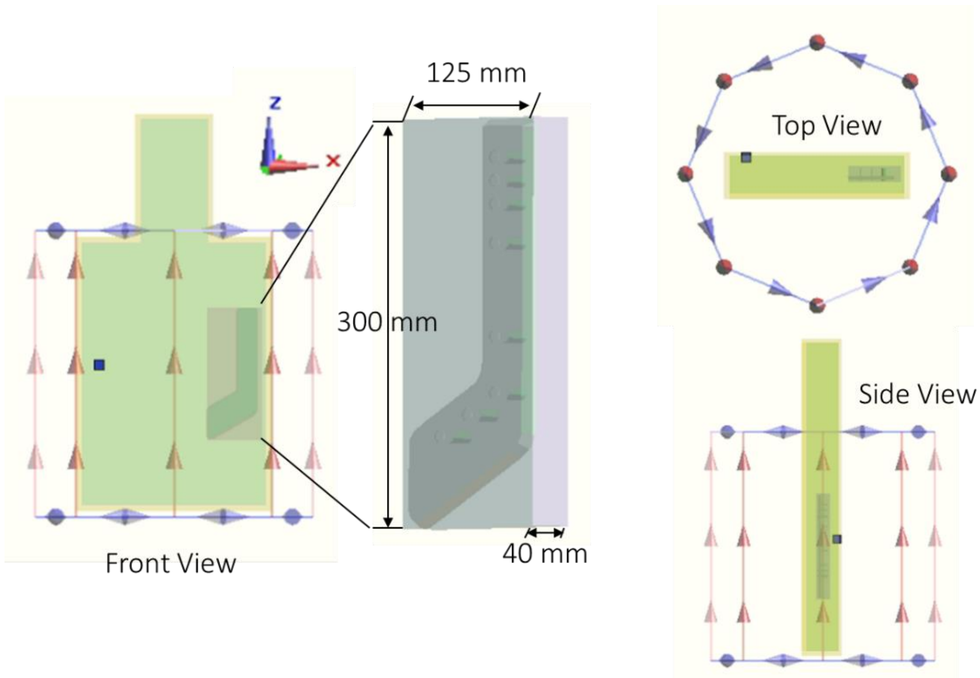


Figure 70. Complex shape device used for CNN.

To get input for the NN, the complex shape devices need to preprocess and transfer to a proper input format. The complex shape devices cannot be described by simple geometrical features. Images are the simplest input format when dealing with a 3D object. Thus, the complex shape devices in 3D will be sliced into layers of images. To reserve the geometrical dimension, a box with a geometrical dimension that can cover all the configurations was first defined as shown in Figure 70. In this study, the box dimension is $125 \times 300 \times 40$ $(x \times y \times z)$.

The device in 3D will be sliced into layers of images in the x-z plane in a y-direction with a step size of 1 mm. Therefore, there are a total number of 40 layers for each complex shape device as shown in Figure 71. Each layer covers an x dimension of 125 mm and a z dimension of 300 mm. The image plots (x-z plane) were shown that some layers will include the devices, and other layers include the screws.

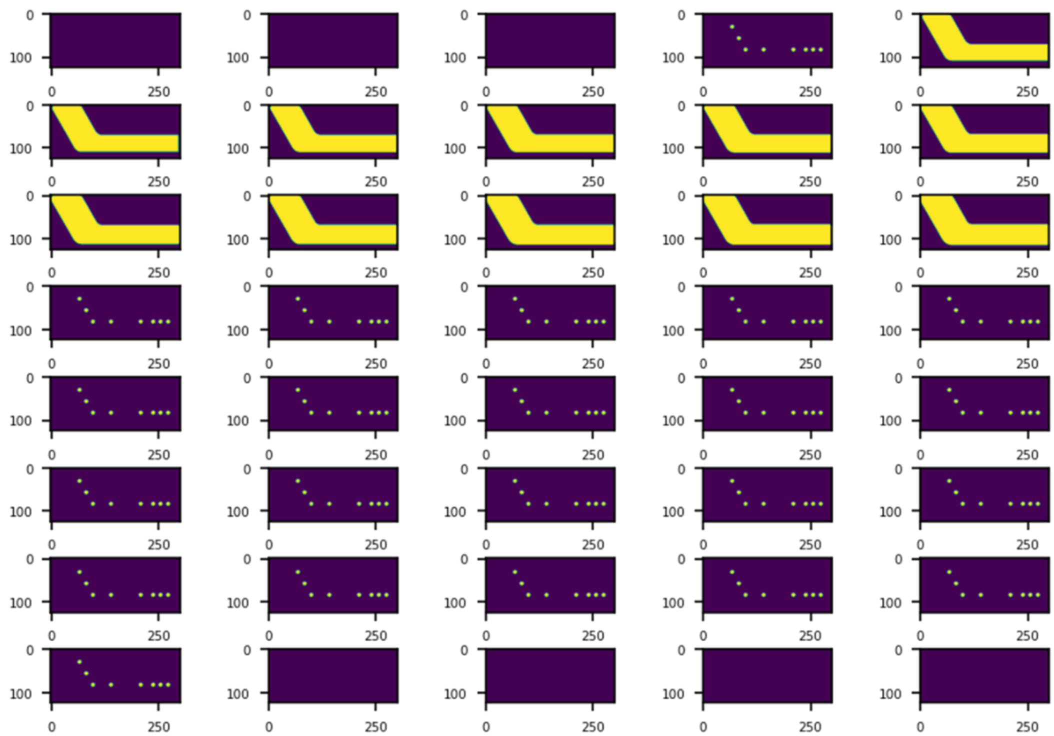


Figure 71. Forty-layers of images in y-direction for each complex shape device.

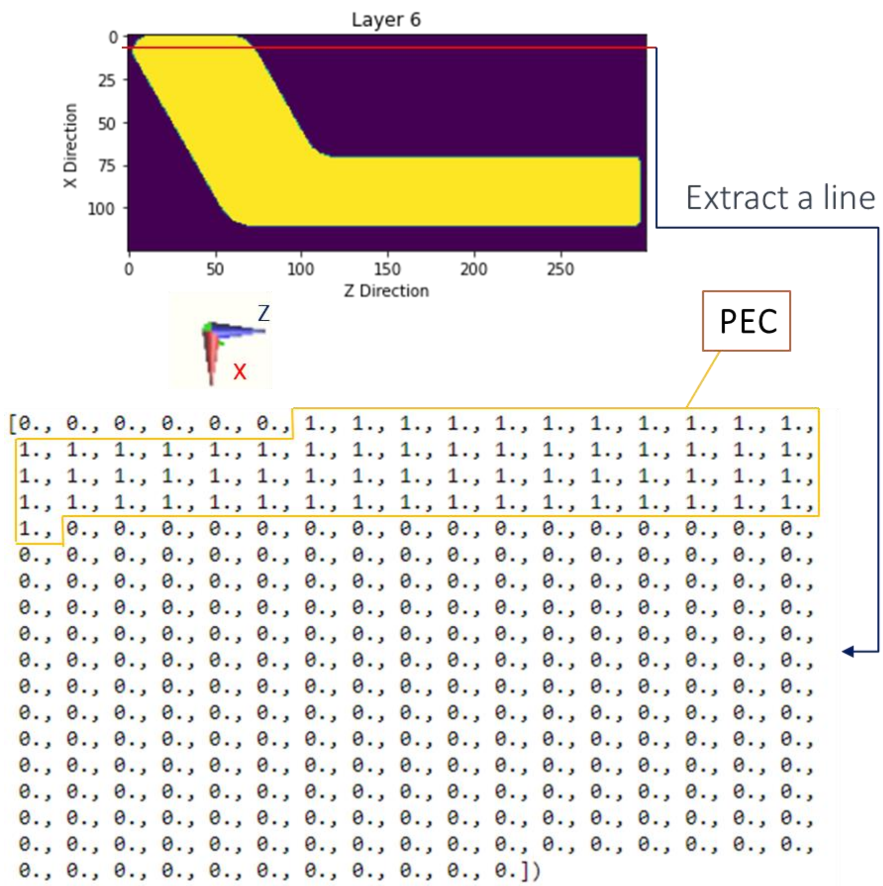


Figure 72. The details in a sliced layer of the complex shape device using mesh representation.

The details in a sliced layer of the complex shape device using the mesh representation were shown in Figure 72. The materials type of PEC will be defined as 1 and other materials will be defined as 0. Thus, the geometrical dimension, shape, orientation, and material type information were included and used as the input of NN.

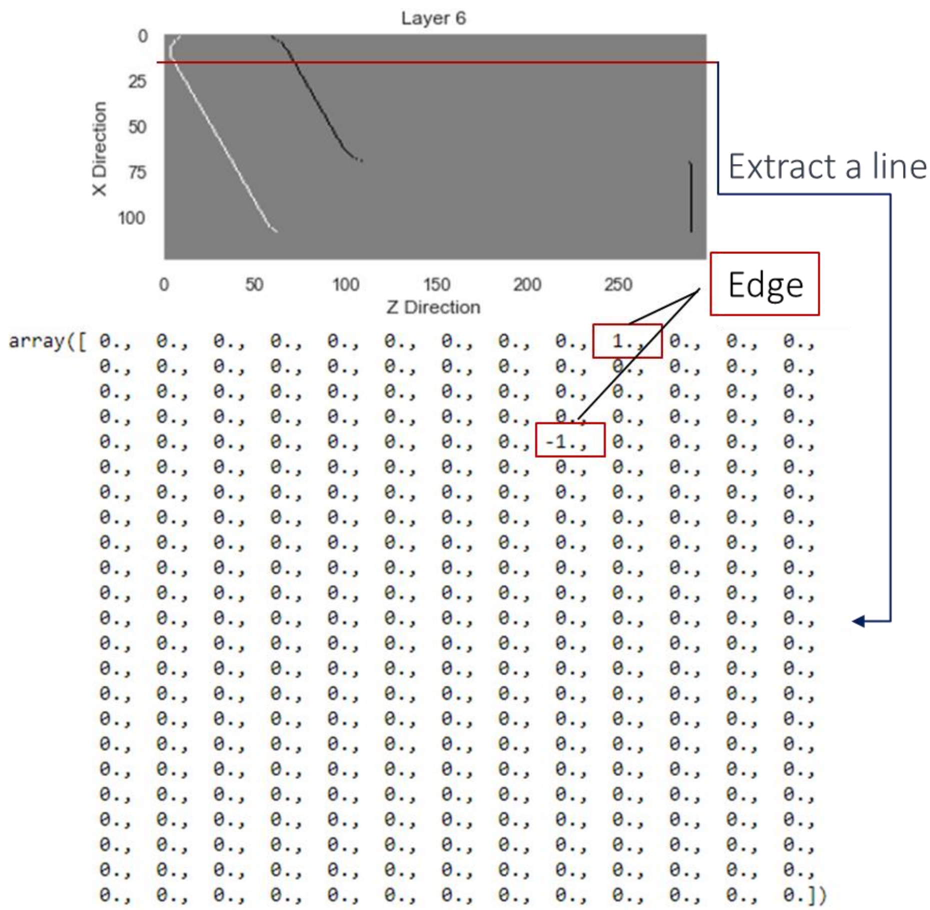


Figure 73. The details in a sliced layer of the complex shape device using the edge representation.

Another type of input representation can be extracted by getting the difference between two points in each layer of the slices. The edge information would be identified as 1 or -1 in each layer as shown in Figure 73. Each device with the layers of edge representations will be used as the input for the CNN.

8.3.3 Results

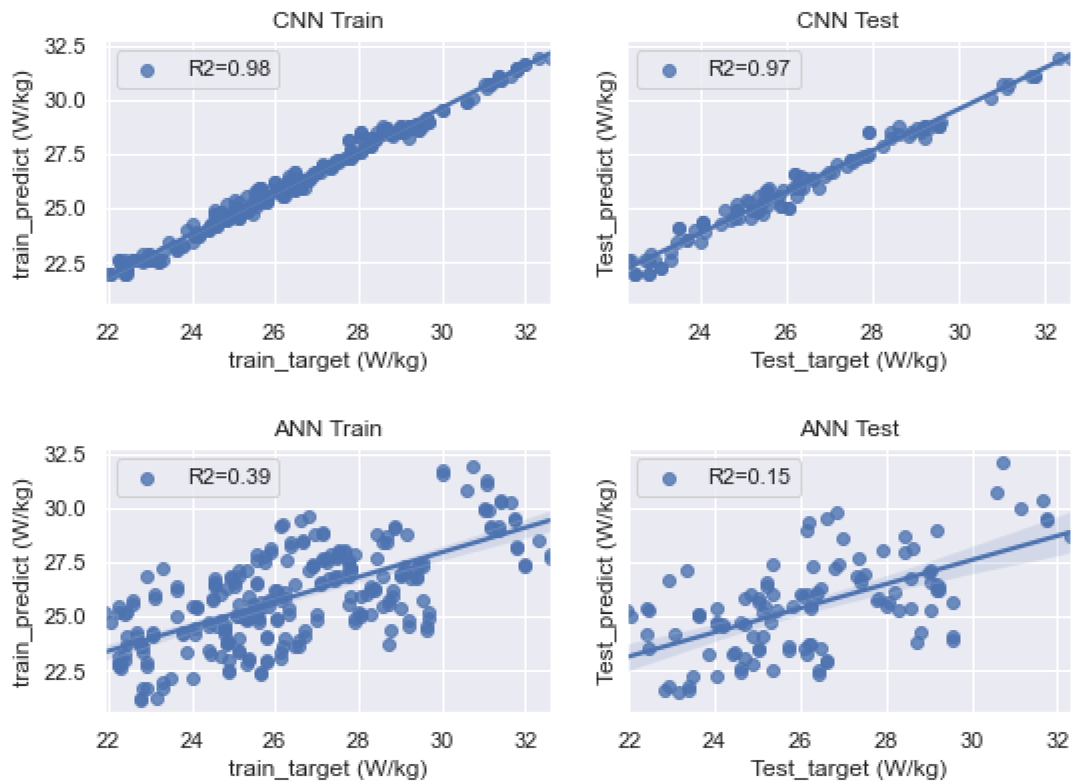


Figure 74. The correlation coefficient of the CNN and ANN results.

Simple geometrical features, such as the plate length, plate width, screw length, and screw diameter, etc. were used as the input for ANN which has lower-dimensional data representation. For the higher dimensional input which contains 40 layers of images will be used as the input for CNN. The output is one-dimensional psSAR10g value for the complex shape plate system. The 384 configurations of the complex shape plate devices were randomly divided, 70% of which were used for training, while the residuals were used for testing. The training and testing correlation coefficient results for both ANN and CNN are shown in Figure 74. The training and testing results of the ANN are divergent which has a correlation coefficient of less than 0.39. However, CNN training and testing results which have the correlation coefficient were larger than 0.97. The correlation coefficient results of the ANN indicated that the neural network cannot learn the non-linear relationship between the simple geometrical features due to the loss of detailed structure information in the complex shape implants. The detailed 3D geometrical representations were included in the CNN, so the performance of this network will be better than ANN.

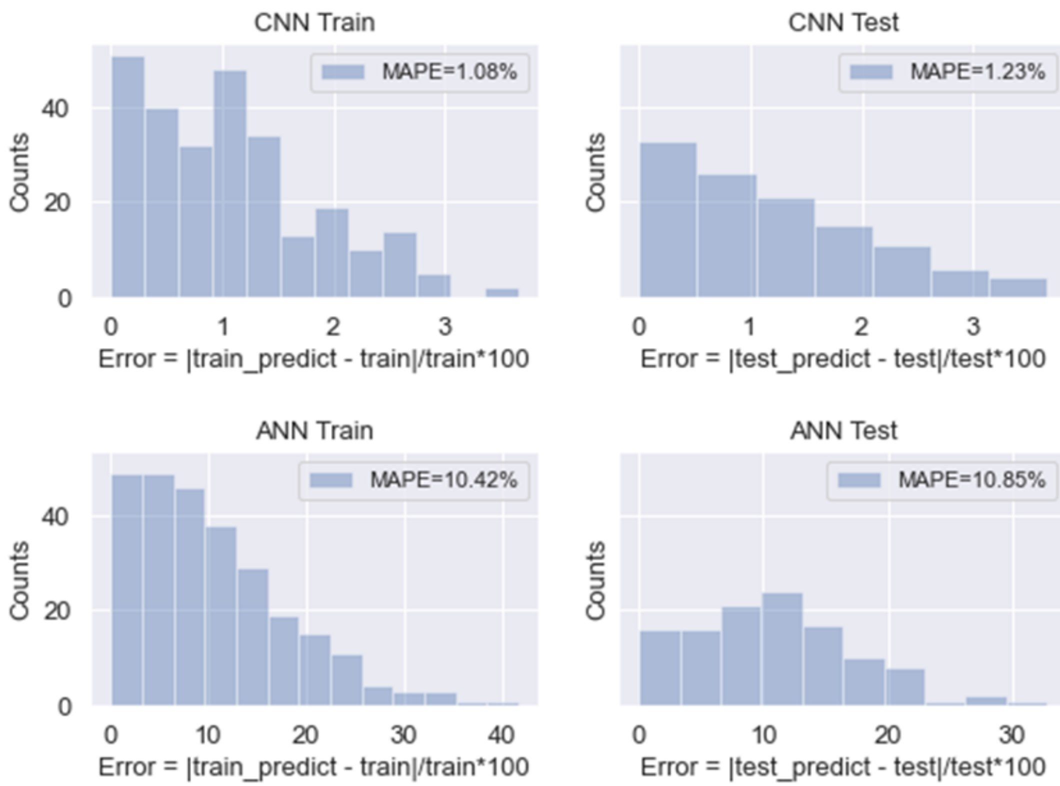


Figure 75. Error histogram of the CNN and ANN results.

The error histogram of the ANN and CNN are shown in Figure 75. The MAPE for the training of CNN was less than 1.08%, while it was less than 1.23% for testing. The error for both training and testing under the ANN model was much larger than the CNN model with a MAPE of 10.42%. The CNN can be used to predict the RF-induced heating after the training and testing process which has a relatively small error rate. However, it's hard to predict the RF-induced heating by using the simple geometrical features for complex shape medical implants as the testing correlation coefficient was less than 0.15 and MAPE was larger than 10.85%.

8.3.4 Minimum Resolution (Mesh Size) Study

The larger mesh size (resolution) could be used to reduce the training time of the CNN. CNN results for volume-based estimation with different mesh sizes were shown in Figure 76. The time cost for the CNN using a 1mm mesh size both in x, y, and z-direction will be 912s, however, the time will be greatly decreased to 37s when using a larger mesh size of 10 mm. The MAPE is still relatively small with a mesh size of 10 mm which is close to 5.77% and the worst-case error is less than 7.26%. Similarly, CNN results for boundary edge-based estimation with different mesh sizes were shown in Figure 77. The time cost for a mesh size of 10 mm is 31s and the MAPE is 12.76%. In such a mesh size, the worst-case prediction error is less than 10.42%.

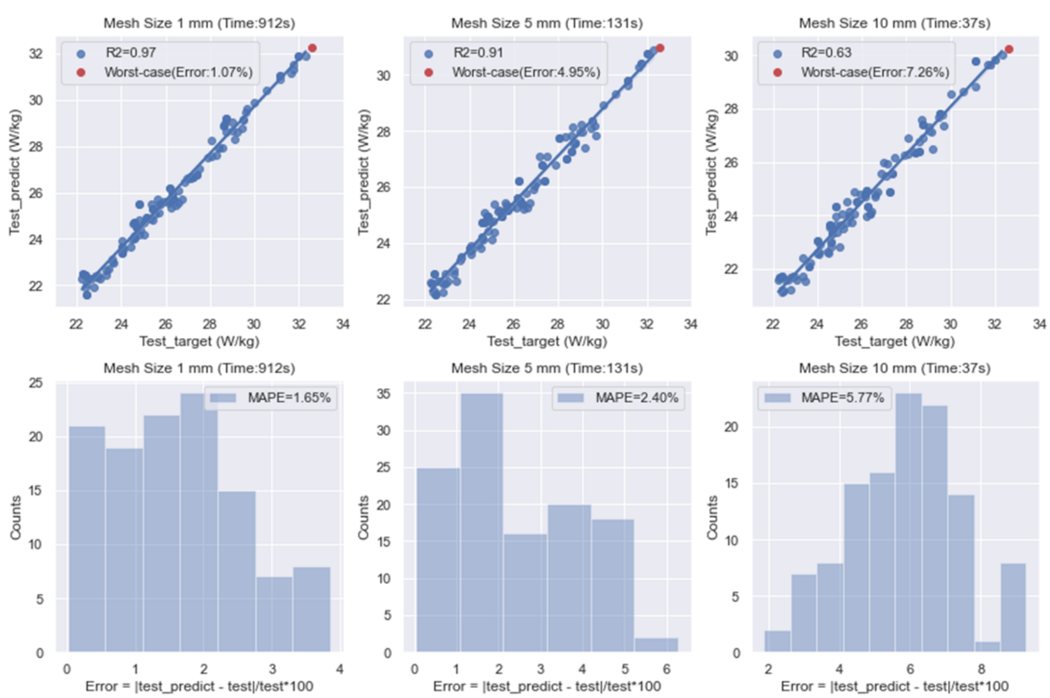


Figure 76. CNN results for volume-based estimation with different mesh sizes.

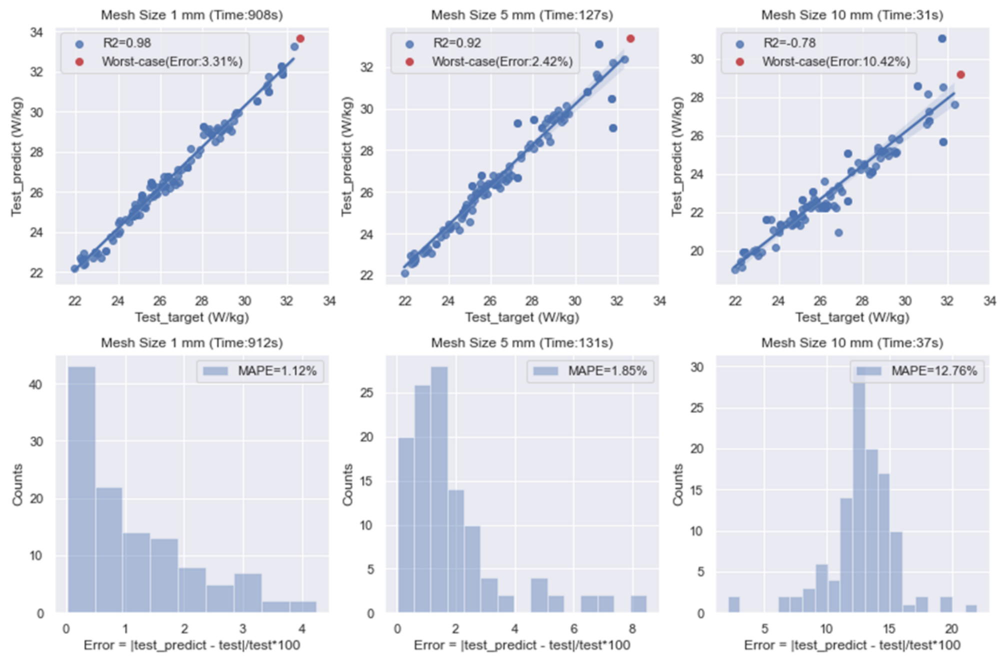


Figure 77. CNN results for boundary edge-based estimation with different mesh sizes.

8.3.5 CNN Performance using Different Training Set

The performance of the CNN using different training sets both for volume-based estimation and edge-based estimation was shown in Figure 78. The worst-case prediction error and the MAPE was less than 5% if the CNN trained by more than 20% of the data set. The CNN will not converge if using a too-small number of the random training set ($< 20\%$) as the MAPE will be larger than 13% for volume-based estimation. The performance of the CNN was not guaranteed if the training data set was less than 20%. Thus, it's recommended to use a training data set that was larger than 20% for the RF-induced heating fast evaluation of complex plate systems. 80% of configurations will no longer be needed for numerical simulations if the CNN only takes 20% training data. Therefore, the gain of the CNN for the complex plate system is $Gain_{CNN} = 80\%$.

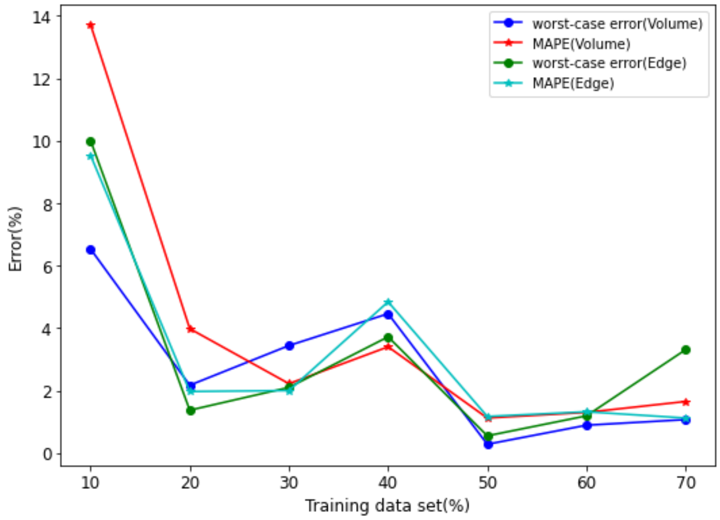


Figure 78. Performance of CNN using the different training set.

8.4 Time Analysis for Convolutional Neural Network

The upper bound of time complexity of the network for external fixation fast evaluation can be expressed as,

$$O(T_{ANN} \times T_{CNN}) = O(T_{ANN} \times \sum_{l=1}^{N} n_{l-1} \cdot s_l^2 \cdot n_l \cdot m_l^2), \tag{16}$$

where T_{ann} is the time complexity of all the dense layers, N denotes the total number of convolutional layers, l is the index of the convolutional layer, n_l is the number of input channels (slice of images in the x-z plane) of the l-th layer, s_l is the size of the filter, and m_l is the output size of kernel convolution from the filter. In this study, the time for $T_{ann} \approx 2s$. For the 3-layer convolutional network, the approximate time close to 460 s using a mesh size of 1 mm, thus the total cost time for the whole CNN is therefore close to 920s. This is much less than the full-wave modeling of the devices. Because each device inside the phantom will cost more than 2 hours by the numerical simulation using an NVIDIA C2075 high-performance graphics processing unit (GPU) which had a normal 1150 MHz clock rate of 448 CUDA cores. Once the network has been trained, the CNN network predictive models are capable of producing hundreds of RF-induced heating predictions within minutes. The training time is roughly three times the testing time.

9 Uncertainty Evaluation

The uncertainty process evaluation follows ISO/TS 10974 Ed. 1 Annex R. According to this document, the result of a simulation value is a function of N parameters $x_1, x_2, \dots, x_N$. The combined standard uncertainty $u_c(y)$ can be derived from individual uncertainty components u_{x_i}. The total uncertainty can be calculated from the root sum square of the uncertainty of the individual components

$$u_c(y) = \sqrt{\sum_{i=1}^{m} c_i^2 \cdot u^2(x_i)}. \tag{17}$$

In this formula c_i is the sensitivity coefficient calculated by $\partial y / \partial x_i$. $u(x_i)$ is the standard deviation of each term and $u_c(y)$ is the combined uncertainty. In this study, the uncertainties of the phantom position, the device position, the gel dielectric properties, and the grid resolution were included.

9.1 The Sensitivity Coefficient Evaluation

Typically, individual uncertainty can be evaluated in two steps. First, the sensitivity coefficient c_i needs to be identified. Then the standard deviation of each term $u(x_i)$ should be evaluated.

The sensitivity coefficients of the phantom position, device position, gel dielectric properties, and grid resolution are estimated. The temperature rise y will be used as the target value. Then the relationship between y and the temperature of the gel T_{gel} can be expressed as

$$y = T_{gel}. \tag{18}$$

In this case, the T_{gel} is a function of the phantom position, rod position, gel dielectric properties, and grid resolution. It is a linear equation when the variation of these parameters is small. The sensitivity coefficient can be calculated as $\Delta y/\Delta x_i$.

To calculate the sensitivity coefficients regarding the phantom position, the phantom position is moved ±5 mm in the x, y, z direction from the standard position where the isocenter of the gel at is at the middle of the coil in x, z-direction, and the phantom is placed in the y-direction, 420 mm from the top of the coil. The sensitivity coefficient is calculated for each direction. The relative percentage R_c is obtained by dividing the sensitivity coefficient by the calculated temperature rise y when the phantom is at the standard position. The sensitivity coefficient of the phantom position is shown in Table 21.

Table 21. Sensitivity coefficient of the phantom position.

Position	$\boldsymbol{\Delta y}$ ($\circ$ $\boldsymbol{C}$)	$\boldsymbol{\Delta x}$ ($\boldsymbol{mm}$)	$\boldsymbol{\Delta y/\Delta x}$	$\boldsymbol{R_c}$**(%)**
X direction	1	10	0.1	0.625
Y-direction	0.7	10	0.07	0.437
Z direction	0.2	10	0.02	0.125
Combination	NA	NA	NA	0.773

To calculate the sensitivity coefficients of the device position, the position of the simplest rod is moved ±5 mm in the x, y, z direction from the standard position (4.5 cm under the gel and 2 cm from the boundary of the gel, center of z-direction). The sensitivity coefficient of the device position is shown in Table 22.

Table 22. Sensitivity coefficient of the device position.

Position	$\Delta y\ (\circ C)$	$\Delta x\ (mm)$	$\Delta y/\Delta x$	$R_c(\%)$
X-direction	0.3	10	0.03	0.187
Y direction	4.4	10	0.44	2.750
Z direction	0.2	10	0.02	0.125
Combination	NA	NA	NA	2.759

To calculate the sensitivity coefficient regarding the phantom properties (conductivity and permittivity), the value of conductivity and permittivity are set to be ±10% deviation from the original value separately. The relative percentage R_c is obtained by dividing the sensitivity coefficient by the calculated temperature rise y when the conductivity and permittivity are the original value. The sensitivity coefficient of the phantom properties was shown in Table 23.

Table 23. Sensitivity coefficient of the phantom properties.

Properties	$\Delta y\ (\circ C)$	Δx	$\Delta y/\Delta x$	$R_c(\%)$
Conductivity	5.1	0.114 S/m	44.737	279.61
Permittivity	0.1	16.06	0.006	0.039

To calculate the sensitivity coefficient regarding the grid resolution, the max step of the mesh is set to be 0.5*0.5*1 mm and 0.25*0.25*0.5 mm. The sensitivity coefficient of the grid resolution was shown in Table 24.

Table 24. Sensitivity coefficient of the grid resolution.

Properties	Δy ($\circ C$)	Δx **(mm)**	$\Delta y / \Delta x$	$R_c (\%)$
Grid resolution	0.3	0.25	1.2	7.5

The summary of sensitivity coefficients for each parameter was shown in Table 25.

Table 25. Sensitivity coefficient for each parameter.

Parameter	$R_c (\%)$
Phantom position	0.773
Device position	2.759
Gel conductivity	279.605
Gel permittivity	0.038
Grid resolution	7.5

9.2 The Evaluation of the Individual Uncertainty Components

Based on the measurements, the estimated standard deviation of the phantom position is 5 mm and the estimated standard deviation of the device position is 5 mm. The relative permittivity of the water over the range of 0.1ºC to 99 ºC can be calculated using the following equation

$$\varepsilon = 87.740 - 0.4008t + 9.398(10^{-4})t^{2} - 1.410(10^{-6})t^{3}. \tag{19}$$

The measurement temperature usually ranges from 18 ºC to 45 ºC. The relative permittivity of the liquid gel under various temperatures is calculated by the above equation and the uncertainty of the permittivity is the standard deviation of 2.85. The uncertainty of gel conductivity is assessed by mixing the gel according to ASTM F-2182 standard 10 times and using a conductivity meter to measure liquid conductivity. The measured results are shown in Table 26.

Table 26. The measure results of the conductivity.

Measurement	**Conductivity (S/m)**
1	0.47
2	0.48
3	0.52
4	0.50
5	0.42
6	0.41
7	0.48
8	0.46
9	0.45
10	0.50

The standard deviation of these ten observations from Table 26 is 0.033 S/m. Furthermore, the uncertainty of the grid resolution is estimated to be 0.25 mm. Therefore, the summary of the individual uncertainty component was shown in Table 27.

Table 27. Summary of the individual uncertainty component.

Component	**Standard deviation**
Phantom position	10 mm
Device position	10 mm
Gel conductivity	0.033 S/m
Gel permittivity	2.85
Grid resolution	0.25 mm

Based on the sensitivity coefficient evaluation results and the individual component evaluation results, the uncertainties introduced by each source ($c_i u(x_i)$) can be calculated and shown in Table 28.

Table 28. The uncertainty caused by each source.

Source	**Uncertainty (%)**
Phantom position	7.73082
Device position	27.59218
Gel conductivity	9.2269749
Gel permittivity	0.11091345
Grid resolution	1.875

In summary, the combined uncertainty from the phantom position, device position, phantom properties, and grid resolution is 30.162%.

10 Discussion

For the parameterized ANN, the minimum step size of lambda/10 can be used as the training set which can reduce the numerical calculation burden. The worst-case errors were less than 6% for all the study cases including three fully implanted devices and one partially implanted device. The worst-case prediction error is not guaranteed as more devices should be investigated in future studies. The critical devices which can induce invertible damages to the human should follow standard guidance to do full investigations and shouldn't use the neural networks for heating fast evaluations.

The worst-case prediction using the neural network will induce large errors if the configurations corresponding to the worst-case SAR were not included within the cover range in the training set. The errors will increase if the cover range of the parameter in the training set was gradually moving far away from the worst-case configuration as shown in Figure 79.

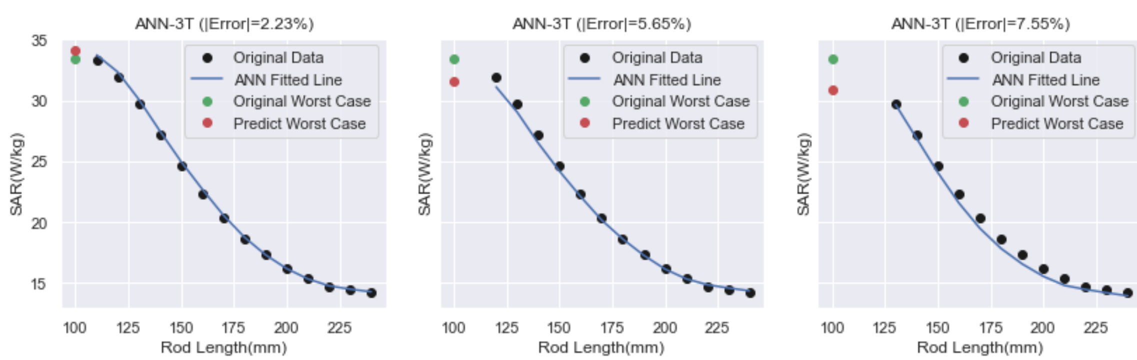


Figure 79. Worst-case prediction results outside the cover range of training set by studying the rod at 3 T system.

In this case, the worst-case configuration was with a rod length of 100 mm at 3 T system. The training set will gradually exclude 10 mm of rod length away from the worst-case configuration. Thus, the cover range of the training set will be smaller. As shown in the figure, the predicted worst-case SAR from the neural network still can be good if the cover range was not far away from the worst-case configuration as it was still higher than the true worst-case SAR (error < 2.23%). However, the predicted worst-case SAR will be lower than the true worst-case SAR and the error will be increased to 7.55% as the cover range of the training set decrease to a range [130~250] mm.

Therefore, to be able to conservatively predict the worst-case SAR using a neural network, the training set is better to cover the range of parameters necessary to include the configuration that corresponds to the worst-case SAR. For example, the worst-case rod length is 100 mm at the 3 T system, then the training set is better to cover this worst-case rod length in the range of the training set or at least not far away from the worst-case rod length.

For future study, the network can be investigated to estimate the RF-induced heating for more complex and realistic implantable devices. There are three ways to further improve neural network’s applicability: Firstly, enlarge the number of labeled devices by providing simulation results of different types of implants; Secondly, extract common features or complex features for different implants that can accurately predict the RF exposure; and Thirdly, consider other critical factors related to RF-induced heating, such as loading position, devices orientation, RF coil type, etc.

11 Conclusion

In this paper, a fast prediction method to evaluate RF-induced heating of PIMDs using NNs was proposed. Previously, it has shown the validity of ANN to predict RF-induced heating of simple plate devices in a homogeneous ASTM phantom. However, the criterion to determine the minimum training data set was not identified. Moreover, the devices that cannot be parameterized by simple geometrical features have not been taken into consideration, which is critical for standard RF-induced heating evaluation. Some of the small geometrical changes play important roles in MRI RF-induced heating evaluation.

Two types of NN were proposed to predict the RF-induced heating in-vitro conditions, the parameterized NN and the mesh-based CNN. The criterion for the minimum training data set was first identified with the simplest device for each type of network, then several complex devices were included to validate the criterion. The training data set was selected based on the wavelength under the 1.5 T system and 3 T system. For the parameterized NN, the fully implanted and partially implanted devices were fully studied. For the mesh-based CNN, two different devices were investigated, the impacts of the mesh size were also studied.

The results indicate that the parameterized NN can be used to predict the RF-induced heating with a small error rate using the minimum step size at both 1.5 T and 3 T systems. The worst-case also can be accurately predicted by using the appropriate minimum step size. Similarly, the parameterized CNN can be used to predict the RF-induced heating for the devices that cannot be parameterized. It is recommended to use a small mesh size to ensure the convergence of the CNN. It has shown that CNN can fast predict the RF-induced heating for the complex shape medical implants with layers of image representations. The NNs can work as the surrogate model to predict the non-linear relationship between input parameters and the RF-induced heating for PIMDs in *in-vitro* evaluations.

References

[1] E. Mattei, G. Calcagnini, F. Censi, M. Triventi, and P. Bartolini, "Numerical model for estimating RF-induced heating on a pacemaker implant during MRI: experimental validation," vol. 57, no. 8, pp. 2045-2052, 2010.

[2] S. Park, J. Nyenhuis, C. Smith, E. Lim, K. Foster, K. Baker, G. Hrdlicka, A. Rezai, P. Ruggieri, and A. Sharan, "Gelled versus nongelled phantom material for measurement of MRI-induced temperature increases with bioimplants," vol. 39, no. 5, pp. 3367-3371, 2003.

[3] S. Oh, A. G. Webb, T. Neuberger, B. Park, and C. Collins, "Experimental and numerical assessment of MRI-induced temperature change and SAR distributions in phantoms and in vivo," vol. 63, no. 1, pp. 218-223, 2010.

[4] R. Guo, M. Chen, J. Zheng, R. Yang, J. Chen, and W. Kainz, "Comparison of in-vivo and in-vitro MRI RF heating for orthopedic implant at 3 tesla." pp. 123-128.

[5] P. Nordbeck, I. Weiss, P. Ehses, O. Ritter, M. Warmuth, F. Fidler, V. Herold, P. M. Jakob, M. E. Ladd, and H. Quick, "Measuring RF-induced currents inside implants: impact of device configuration on MRI safety of cardiac pacemaker leads," vol. 61, no. 3, pp. 570-578, 2009.

[6] Y. Liu, J. Chen, F. G. Shellock, and W. Kainz, "Computational and experimental studies of an orthopedic implant: MRI-related heating at 1.5-T/64-MHz and 3-T/128-MHz," vol. 37, no. 2, pp. 491-497, 2013.

[7] A. 2182-11a, "Standard test method for measurement of radio frequency induced heating on or near passive implants during magnetic resonance imaging," *American Society for Testing and Materials International*, 2011.

[8] R. Luechinger, V. Zeijlemaker, E. M. Pedersen, P. Mortensen, E. Falk, F. Duru, R. Candinas, and P. Boesiger, “In vivo heating of pacemaker leads during magnetic resonance imaging,” vol. 26, no. 4, pp. 376-383, 2005.

[9] M. Shinkai, K. Ueda, S. Ohtsu, H. Honda, K. Kohri, and T. Kobayashi, “Effect of functional magnetic particles on radiofrequency capacitive heating: an in vivo study,” vol. 93, no. 1, pp. 103-108, 2002.

[10] T. Sommer, C. Vahlhaus, G. Lauck, A. v. Smekal, M. Reinke, U. Hofer, W. Block, F. Träber, C. Schneider, and J. Gieseke, “MR imaging and cardiac pacemakers: in vitro evaluation and in vivo studies in 51 patients at 0.5 T,” vol. 215, no. 3, pp. 869-879, 2000.

[11] X. Huang, Z. Wang, J. Chen, and J. Zheng, "Numerical Study on MRI RF Heating for Circular External Fixators under 1.5 T MRI." pp. 280-285.

[12] X. Ji, J. Zheng, R. Yang, W. Kainz, and J. Chen, “Evaluations of the MRI RF-induced heating for helical stents under a 1.5 T MRI system,” vol. 61, no. 1, pp. 57-64, 2018.

[13] J. Zheng, D. Li, J. Chen, and W. Kainz, "Numerical study of SAR for multi-component orthopaedic hip replacement system during MRI." pp. 116-120.

[14] R. Guo, J. Zheng, Y. Wang, Q. Zeng, Q. Wang, R. Yang, W. Kainz, and J. Chen, “Computational and experimental investigation of RF-induced heating for multiple orthopedic implants,” *Magn Reson Med*, vol. 82, no. 5, pp. 1848-1858, Nov, 2019.

[15] Q. Zeng, Q. Wang, J. Zheng, W. Kainz, J. Chen, and Biology, “Evaluation of MRI RF electromagnetic field induced heating near leads of cochlear implants,” vol. 63, no. 13, pp. 135020, 2018.

[16] FDA-2015-D-2104, “Assessment of radiofrequency-induced heating in the magnetic resonance (MR) environment for multi-configuration passive medical devices,” *Center for Devices and Radiological Health*, 2016.

[17] X. Ji, M. Xia, J. Zheng, J. Chen, and W. Kainz, "Study on the searching strategies of assessing the MRI RF-induced heating for an implantable plate and screw system." pp. 932-935.

[18] J. Zheng, X. Ji, W. Kainz, and J. Chen, “Study on Search Strategies for Assessing the Worst Case RF-Induced Heating for Multi-Configuration Implant System Under MRI,” 2018.

[19] J. Zheng, R. Yang, and J. Chen, "Fast prediction of MRI RF-induced heating for implantable plate devices using neural network." pp. 1011-1012.

[20] J. Zheng, Q. Lan, X. Zhang, W. Kainz, and J. Chen, “Prediction of MRI RF exposure for implantable plate devices using artificial neural network,” 2019.

[21] M. Xia, J. Zheng, R. Yang, S. Song, J. Xu, Q. Liu, W. Kainz, S. A. Long, and J. Chen, “Effects of patient orientations, landmark positions, and device positions on the MRI RF-induced heating for modular external fixation devices,” 2020.

[22] R. Guo, J. Zheng, and J. Chen, "MRI RF-Induced Heating in Heterogeneous Human Body with Implantable Medical Device," *High-Resolution Neuroimaging-Basic Physical Principles and Clinical Applications*: IntechOpen, 2017.

[23] D. Li, X. Ji, J. Zheng, J. Chen, and C. Pan, "A novel design of implantable medical stent for reducing the MRI RF-induced heating." pp. 1945-1946.

[24] H. Muranaka, T. Horiguchi, Y. Ueda, S. Usui, N. Tanki, and O.Nakamura, “Evaluation of RF heating on hip joint implant in phantom during MRI

examinations," vol. 66, no. 7, pp. 725-733, 2009.

[25] H. Muranaka, T. Horiguchi, S. Usui, Y. Ueda, O. Nakamura, F. Ikeda, K. Iwakura, and G. Nakaya, "Evaluation of RF heating on humerus implant in phantoms during 1.5 T MR imaging and comparisons with electromagnetic simulation," vol. 5, no. 2, pp. 79-88, 2006.

[26] J. Powell, A. Papadaki, J. Hand, A. Hart, and D. McRobbie, "Numerical simulation of SAR induced around Co-Cr-Mo hip prostheses in situ exposed to RF fields associated with 1.5 and 3 T MRI body coils," vol. 68, no. 3, pp. 960-968, 2012.

[27] L. Winter, E. Oberacker, C. Özerdem, Y. Ji, F. von Knobelsdorff-Brenkenhoff, G. Weidemann, B. Ittermann, F. Seifert, and T. Niendorf, "On the RF heating of coronary stents at 7.0 Tesla MRI," vol. 74, no. 4, pp. 999-1010, 2015.

[28] R. Yang, J. Zheng, W. Kainz, and J. Chen, "Numerical investigations of MRI RF-induced heating for external fixation device in TEM and birdcage body coils at 3 T," vol. 60, no. 3, pp. 598-604, 2017.

[29] J. Zheng, Q. Lan, W. Kainz, S. A. Long, and J. Chen, "Genetic algorithm search for the worst-case MRI RF exposure for a multiconfiguration implantable fixation system modeled using artificial neural networks," 2020.

[30] J. C. Rayan, N. Reddy, J. H. Kan, W. Zhang, and A. Annapragada, "Binomial classification of pediatric elbow fractures using a deep learning multiview approach emulating radiologist decision making," *Radiology: Artificial Intelligence*, vol. 1, no. 1, pp. e180015, 2019.

[31] Y. L. Thian, Y. Li, P. Jagmohan, D. Sia, V. Chan, and R. Tan, "Convolutional neural networks for automated fracture detection and localization on wrist

radiographs," *Radiology: Artificial Intelligence,* vol. 1, no. 1, pp. e180001, 2019.

[32] Y. Ohta, H. Yunaga, S. Kitao, T. Fukuda, and T. Ogawa, "Detection and Classification of Myocardial Delayed Enhancement Patterns on MR Images with Deep Neural Networks: A Feasibility Study," *Radiology: Artificial Intelligence,* vol. 1, no. 3, pp. e180061, 2019.

[33] K. Wang, A. Mamidipalli, T. Retson, N. Bahrami, K. Hasenstab, K. Blansit, E. Bass, T. Delgado, G. Cunha, and M. S. Middleton, "Automated CT and MRI Liver Segmentation and Biometry Using a Generalized Convolutional Neural Network," *Radiology: Artificial Intelligence,* vol. 1, no. 2, pp. 180022, 2019.

[34] F. Liu, B. Guan, Z. Zhou, A. Samsonov, H. Rosas, K. Lian, R. Sharma, A. Kanarek, J. Kim, and A. Guermazi, "Fully Automated Diagnosis of Anterior Cruciate Ligament Tears on Knee MR Images by Using Deep Learning," *Radiology: Artificial Intelligence,* vol. 1, no. 3, pp. 180091, 2019.

[35] K. Liu, Q. Li, J. Ma, Z. Zhou, M. Sun, Y. Deng, W. Tu, Y. Wang, L. Fan, and C. Xia, "Evaluating a Fully Automated Pulmonary Nodule Detection Approach and Its Impact on Radiologist Performance," *Radiology: Artificial Intelligence,* vol. 1, no. 3, pp. e180084, 2019.

[36] R. Meier, P. Lux, S. Jung, U. Fischer, J. Gralla, M. Reyes, R. Wiest, R. McKinley, and J. Kaesmacher, "Neural Network–derived Perfusion Maps for the Assessment of Lesions in Patients with Acute Ischemic Stroke," *Radiology: artificial intelligence,* vol. 1, no. 5, pp. e190019, 2019.

[37] A. Parakh, H. Lee, J. Eisner, D. V. Sahani, and S. Do, "Urinary Stone Detection on CT Images Using Deep Convolutional Neural Networks: Evaluation of Model

Performance and Generalization," *Radiology: Artificial Intelligence,* vol. 1, no. 4, pp. e180066, 2019.

[38] A. Agarap, "Deep learning using rectified linear units (relu)," 2018.

[39] V. Dumoulin, and F. Visin, "A guide to convolution arithmetic for deep learning," 2016.

[40] D. P. Kingma, and J. Ba, "Adam: A method for stochastic optimization," 2014.

[41] Z.MedTech, "Sim4Life 4.0 User Manual," *ZMT Zurich MedTech, Zurich, Switzerland*, 2018.

[42] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems,* vol. 25, pp. 1097-1105, 2012.

[43] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, and M. Isard, "Tensorflow: A system for large-scale machine learning." pp. 265-283.

[44] P. D. Hajek, H. Bicknell Jr, W. E. Bronson, J. A. Albright, S. Saha, and j. A. volume, "The use of one compared with two distal screws in the treatment of femoral shaft fractures with interlocking intramedullary nailing. A clinical and biomechanical analysis," vol. 75, no. 4, pp. 519-525, 1993.