

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

MRSaiFE: An AI-based Approach Towards the Real-Time Prediction of Specific Absorption Rate

Sayim Gokyar¹, Member, IEEE, Fraser J. L. Robb², Wolfgang Kainz³, Member, IEEE, Akshay Chaudhari⁴, and Simone Angela Winkler¹, Senior Member, IEEE

¹Department of Radiology, Weill Cornell Medicine, New York City, NY 10065 USA

²GE Healthcare Coils, 1515 Danner Drive, Aurora, OH 44202 USA

³Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, MD, USA

⁴Integrative Biomedical Imaging Informatics at Stanford (IBIIS), James H. Clark Center, 318 Campus Drive, S255 Stanford, CA 94305 USA

Corresponding author: Sayim Gokyar (e-mail: sag4008@med.cornell.edu).

This work was supported by the National Institutes of Health (NIH) under NIH R00EB024341, and GE Healthcare. Authors would like to acknowledge the valuable assistance of Cynthia Fox in the proofreading of the manuscript and Jinwei Zhang for fruitful discussions.

ABSTRACT The purpose of this study is to investigate feasibility of estimating the specific absorption rate (SAR) in MRI in real time. To this goal, SAR maps are predicted from 3T- and 7T-simulated magnetic resonance (MR) images in 10 realistic human body models via a convolutional neural network. Two-dimensional (2-D) U-Net architectures with varying contraction layers and different convolutional filters were designed to estimate the SAR distribution in realistic body models. Sim4Life (ZMT, Switzerland) was used to create simulated anatomical images and SAR maps at 3T and 7T imaging frequencies for Duke, Ella, Charlie, and Pregnant Women (at 3, 7, and 9 month gestational stages) body models. Mean squared error (MSE) was used as the cost function and the structural similarity index (SSIM) was reported. A 2-D U-Net with 4 contracting (and 4 expanding) layers and 64 convolutional filters at the initial stage showed the best compromise to estimate SAR distributions. Adam optimizer outperformed stochastic gradient descent (SGD) for all cases with an average SSIM of $90.5 \pm 3.6\%$ and an average MSE of $0.7 \pm 0.6\%$ for head images at 7T, and an SSIM of $>85.1 \pm 6.2\%$ and an MSE of $0.4 \pm 0.4\%$ for 3T body imaging. Algorithms estimated the SAR maps for 224x224 slices under 30 ms. The proposed methodology shows promise to predict real-time SAR in clinical imaging settings without using extra mapping techniques or patient-specific calibrations.

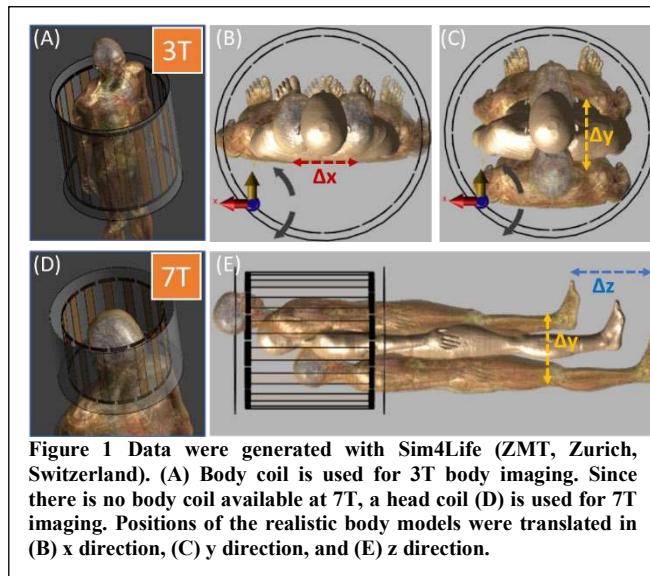
INDEX TERMS Image processing, Magnetic resonance imaging, Machine learning, Specific absorption rate, Supervised learning

I. INTRODUCTION

Modern neuroscience relies on understanding the brain in health and disease. Thus, the availability of technology that can significantly increase the spatial resolution and sensitivity achievable with magnetic resonance (MR) neuroimaging at the current used field strength of 3T and emerging ultra-high field (UHF) strengths of 7T and higher, consistent with safety, would offer the potential to advance our understanding of brain structure and function by enabling investigations with greater specificity and granularity [1].

A key limitation to the high potential of UHF magnetic resonance imaging (MRI) in neuroscience research and clinical or diagnostic applications [2, 3] is the safety concern

related to the nonuniform deposition of radiofrequency (RF) power in the body [4], quantified by the specific absorption rate (SAR), which can lead to dangerous tissue heating and damage [5]. Not only does the average SAR possess a quadratic dependence on the static magnetic field strength (B_0) [6], increasing 4-fold from 3T to 7T, but due to the higher Larmor frequency and thus shortened in-tissue wavelength, it also exhibits a spatial variation that can lead to “local SAR” patterns or “hotspots” of focused high RF power deposition and localized tissue heating [7-13]. This results in significant SAR nonuniformity at anatomical dimensions similar to the brain at 7T (wavelength ~ 11 cm in tissue) as well as the body at 3T (wavelength ~ 26 cm in tissue) [14].



Moreover, parallel transmit (pTx) technology with multiple independent transmit RF channels [15-17] is now common in UHF applications and can lead to even stronger hotspots because of potential constructive interference of the electric fields [18, 19].

While a small portion of UHF MRI has received Food and Drug Administration (FDA) approval [20, 21], most clinical imaging has been performed at 3T to date. This is because substantial safety and technological hurdles still need to be surmounted before the potential benefits of higher sensitivity and spatial resolution can be fully realized. Specifically, there is a lack of technology that can measure local SAR due to anatomical and positional variations between patients, and between transmit coils. Current technology is not equipped to measure spatially varying local SAR; the only quantity that can be determined in vivo is the overall average, or global, SAR, delivered to the entire anatomy under investigation.

Local SAR variation is highly difficult to predict due to anatomical and positional variations between patients, as well as variations in transmit coils. Many institutions use a conservative estimate of the peak local SAR via its ratio to the measurable global SAR; typically, $\sim 20:1$ [19], thereby severely limiting the applied transmit power and thus the imaging performance achievable by UHF MRI, in particular resolution and/or scan time. This critical barrier is one of the main reasons 7T has not yet reached the patient in its full capacity and holds back its success as an extremely powerful imaging modality with unprecedented ability to decipher fine structures.

MR thermometry as an alternative approach suffers from a coarse temperature resolution [22]. The advent of artificial intelligence (AI) and machine learning in MRI [23-36] has opened up new avenues for the prediction of various imaging characteristics, among them the recent prediction of local SAR in prostate imaging [37-40], as well as the prediction of temperature rise in the brain for 33 different tissue types [41].

In this paper, we expand these pioneering techniques by use of a comprehensive training set of anatomical models and predict local SAR both at the 3T and 7T field strengths, for a large variety of landmarks and thus imaging anatomies. We propose MRSaiFE [42-44], which is eventually expected to become an AI-based, exam-integrated, MRI safety prediction software facilitating the safe generation of 3T and 7T images. In this work, we demonstrate proof of concept and feasibility for MRSaiFE using a database of 10 human body models. . The algorithm is trained and evaluated for use at (1) 3T body dimensions and (2) 7T brain dimensions – significant SAR non-uniformity can occur at both field strengths and provides good basis to train the network on a variety of scenarios. Our results suggest SAR-monitoring with $>90\%$ structural similarity index (SSIM) [45] and $<1\%$ mean squared error (MSE) for high resolution imaging at UHF.

II. METHODS

A. DATA GENERATION

The input data for this study were acquired from Sim4Life simulations (ZMT, Zurich, Switzerland) using Virtual Population (IT'IS, Zurich, Switzerland) body models (Figure 1). Figure 2 shows an illustration of the proposed neural network. An anatomical image is used as the input to a neural network and predicts a SAR map in image form at the output.

The anatomical input image that would come from an MRI scanner in a real experiment was approximated by using black/white images of the body models. Image data sets were synthesized for 1 mm anatomical resolution and down sampled to $224 \times 224 \times 224$ pixels for a field of view (FoV) of $30 \text{ cm} \times 30 \text{ cm} \times 30 \text{ cm}$ for head imaging (1.3 mm isotropic resolution) and for a FoV of $60 \text{ cm} \times 60 \text{ cm} \times 60 \text{ cm}$ for body imaging (2.7 mm isotropic resolution). In order to encode landmark information as part of a neural network with a single input and output image, these anatomical images were weighted by the normalized B_1^+ field of the unloaded RF coils [43].

For 3T body MRI SAR prediction, a 16-rung body coil model (65 cm diameter, 67 cm length) made for a standard bore size of 60 cm was used in conjunction with the body models Ella (26y, body mass index (BMI) 21.6 kg/m^2), Duke (34y, BMI 22.4 kg/m^2), and pregnant women at gestational stages of 3, 7, and 9 months (26y, BMI not available). Duke was positioned at 225 different imaging landmarks spanning $\pm 40 \text{ mm}$, $\pm 60 \text{ mm}$, and $\pm 60/-100 \text{ mm}$ along the x-, y-, and z-axes (axial: xy-plane, coronal: yz-plane, sagittal: xz-plane). Similarly, Ella and the pregnant women were positioned at 125 different imaging landmarks varying between $\pm 40 \text{ mm}$ in x- and y-axes, and $\pm 200 \text{ mm}$ in z-axis. The 1-gram averaged peak local SAR outputs were evaluated for 1 W of simulated input power and coronal SAR slices were extracted. Here, the iterations where the body models intersected with the coil, and slices containing no information, were discarded from the dataset. This resulted in

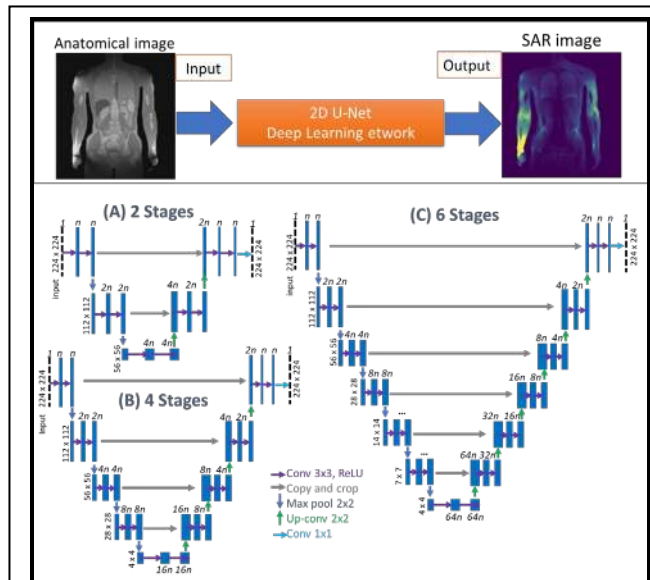


Figure 2. 2-D U-Net architectures used in this study. (A) 2 stages, (B) 4 stages, and (C) 6 stages of encoder layers were implemented. Length of feature vectors was defined by n and varied between 32 to 256 as powers of 2. Increasing n improved the network capacity, with costs of increased memory

a set of 67,861 input anatomical and 67,861 output SAR images at 3T.

For 7T head MRI SAR prediction, a 16-rung birdcage head coil model (30 cm diameter, 25 cm length) was used with the body models Ella (26y female, BMI 21.6 kg/m²), Charlie (8w female, weight 4.3 kg, BMI not available), and pregnant women at different gestational stages (3, 7, and 9 months, 26y, BMI not available). All models were positioned at the same imaging landmarks within the coil ranging from ± 20 mm in x- and y-axes, and ± 60 mm in z-axis. Input anatomical image and output SAR image generation followed the same steps as for the 3T analysis, resulting in a set of 52,806 input anatomical and 52,806 output SAR images at 7T.

The data were split into three categories for each body model separately with an approximate ratio of 80, 10, and 10% for training, validation, and testing, respectively. A summary of the dataset is provided in Table I. Algorithms were trained and tested for each body model separately.

B. NEURAL NETWORK ARCHITECTURE

A 2D U-Net architecture [46] estimated SAR maps (y) of a given MRI image (x) by learning a transformation, $y=f(x)$, between the two. Here, y is the simulated SAR map of a given slice, which was also used as the ground truth (GT).

We implemented a set of U-Net architectures using a cascade of 2D convolutional filters (Figure 2). Convolutional layers were paired with nonlinear rectified linear unit (ReLU) activation functions [47], batch normalization [48], and He initialization [49]. The input image size was set to 224 x 224

pixels, zero padding was used for convolutions, and the output was cropped to the size of the input image. In order to analyze the effect of different hyperparameters on the network performance, we designed 3 U-Nets with varying number of contraction (encoder) and expansion (decoder) layers, using different filter feature map lengths (n) and learning rates (LR). The first, second, and third architectures used 2, 4, and 6 encoder and decoder stages, respectively (Figure 2. A, B, and C). Each stage contained 2 cascaded convolutional filters [36] with the first layer feature map lengths varied from 32 to 256 and orders of 2 (i.e., $n=32, 64, 128$, and 256). These filter lengths decreased to 16 and 32 for the 6-stage network due to limited memory. The feature maps of the first layers were incremented by a factor of 2 for the subsequent encoding stages and decremented by a factor of 2 for the consequent decoding stages. All networks used 3x3 convolution filters, a 1x1 stride length, 2x2 max pooling, and up-convolutions. The final layer of each architecture included 1x1 convolutional filters with a sigmoid activation function.

Table I. Dataset generated in this study.

| Field | Body | Total # | Training | Validation | Testing |
|----------|---------|---------|----------|------------|---------|
| Strength | Model | Images | | | |
| 3T | Ella | 10,340 | 8272 | 1034 | 1034 |
| | Duke | 16,766 | 13332 | 1616 | 1818 |
| | Preg3M | 11,070 | 8820 | 1080 | 1170 |
| | Preg7M | 14,605 | 11684 | 1397 | 1524 |
| | Preg9M | 15,080 | 12035 | 1450 | 1595 |
| 7T | Ella | 12,600 | 10080 | 1296 | 1224 |
| | Charlie | 12,075 | 9660 | 1242 | 1173 |
| | Preg3M | 12,702 | 10147 | 1241 | 1314 |
| | Preg7M | 7,548 | 6052 | 748 | 748 |
| | Preg9M | 7,881 | 6319 | 781 | 781 |
| TOTAL | | 120,667 | 88,129 | 10,851 | 11,347 |

C. TRAINING AND VALIDATION

We used a 3T body imaging data set generated by using Ella with the 3 different U-Net configurations to determine the best performing architecture and optimized the hyperparameters using SSIM, MSE, and training time per image as performance metrics.

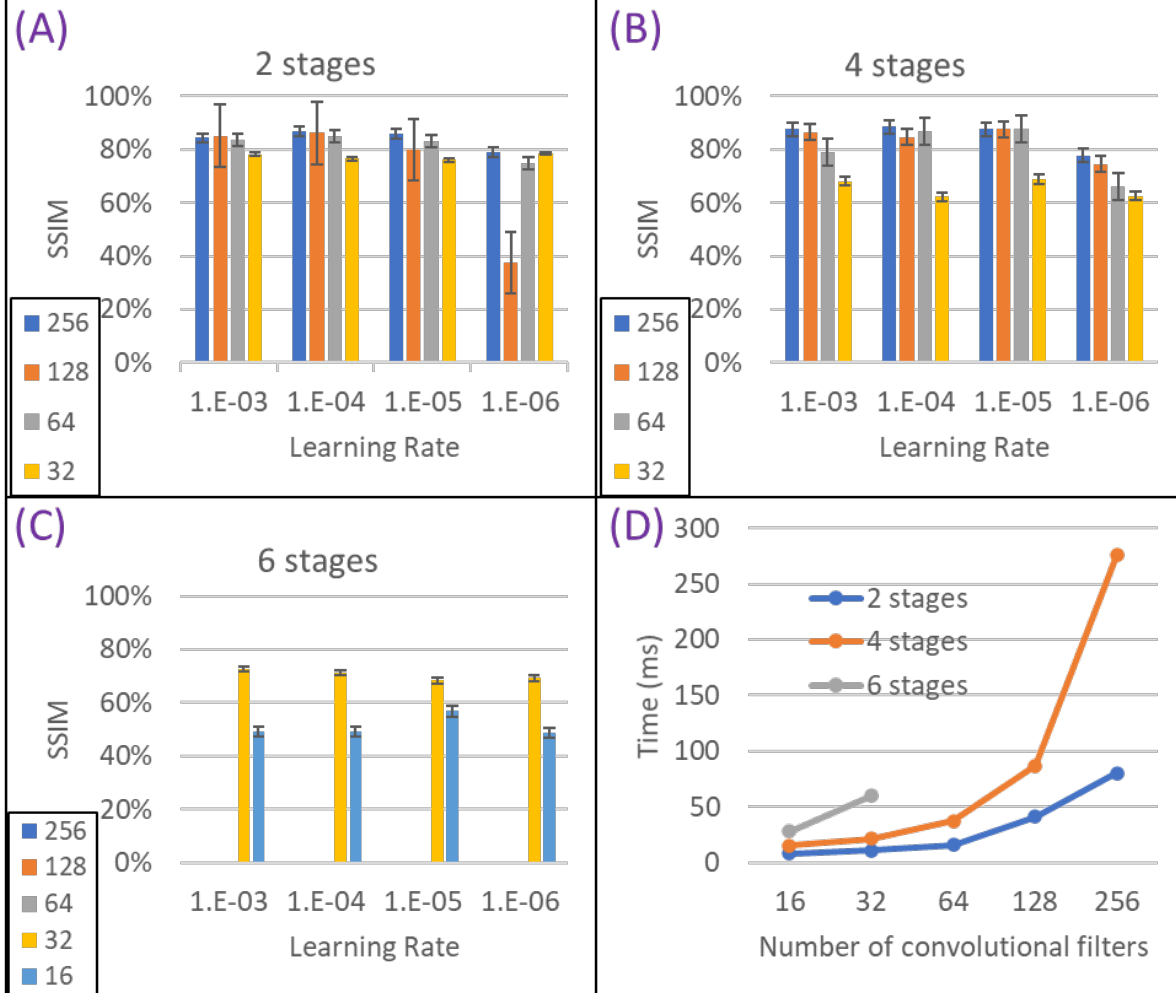


Figure 3. Performance comparison of three U-Nets due to different hyperparameters and time budget. All structures were trained and tested on the same realistic body model (Ella) for 3T body imaging. SSIM of (A) 2-stage, (B) 4-stage, and (C) 6-stage U-Nets are depicted. (D) Time to train each image showed exponential increase among different U-Nets. Increasing the length of feature maps showed linear increase on training time for the same network up to $n=128$.

Many deep learning image processing algorithms use MSE and SSIM as a basis of comparison between the ground truth and predicted images. The most traditional comparison uses the MSE and represents the cumulative mean squared error between ground truth and predicted images. The SSIM is often used in addition to measure the perceived change in structural information from ground truth to prediction. Given our need to predict anatomical outlines and SAR hotspots, this metric is a good indicator of general structural correctness of the predicted image.

Training was performed using adaptive moment estimation (Adam) [50] and stochastic gradient descent (SGD) [51] optimizers with constant exponential decay rates and momentum values (i.e., $\beta_1=0.90$, $\beta_2=0.99$, and $\epsilon=10^{-7}$ for Adam, as well as a momentum=0.9 for SGD). To optimize the hyperparameters, the MSE between the simulated and predicted SAR maps was minimized on the validation dataset by using (1).

$$MSE = 1/N * \sum (y - \hat{y})^2 \quad (1)$$

Here N is the total number of pixels in the images. A dynamic learning rate (LR) approach was used with initial LRs of 10^{-3} , 10^{-4} , 10^{-5} , and 10^{-6} , and these rates were decreased to 80% at every 5th epoch. The maximum number of epochs was set to 200, and the best model was determined and saved using the Keras and Tensorflow backends (Google, Mountain View, CA). All training algorithms were run twice for validation purposes, and the batch size for training, validation, and test datasets was set to 1. All algorithms were run on an NVIDIA Titan RTX (NVIDIA, Santa Clara, CA) Graphics Processing Unit (GPU).

D. TESTING

Testing was performed on the testing datasets prepared for each body model. Quantitative image comparisons were performed between the GT images (simulated SAR) and the

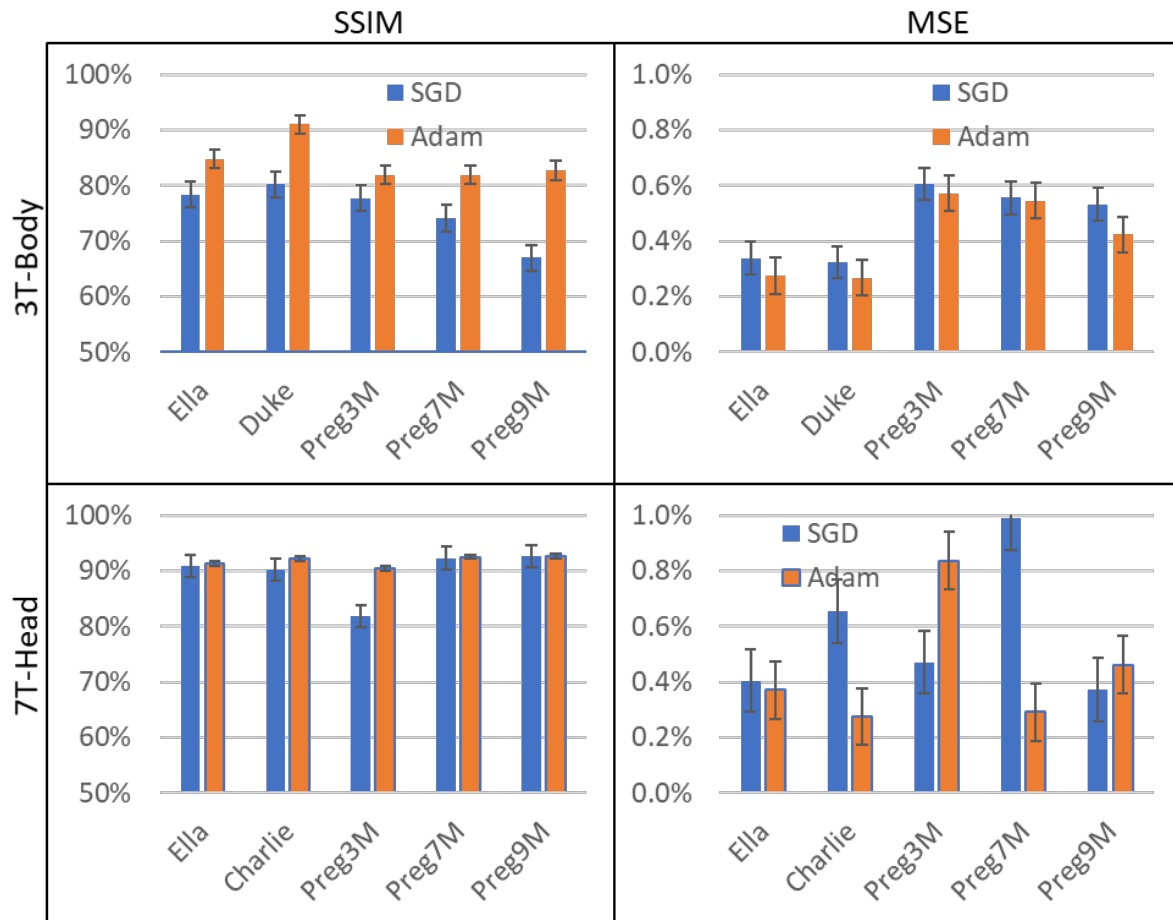


Figure 4. Inter-subject performance analysis of the 4-stage U-Net with 64 convolutional filters at the initial layer. Adam optimizer worked better than SGD in terms of SSIM for all scenarios. (Top Row) For 3T Body SAR prediction, an average SSIM of $85.1 \pm 6.2\%$ and an average MSE of $0.4 \pm 0.4\%$ was observed (Adam), as opposed to a SSIM of $69.3 \pm 4.5\%$ and an MSE of $0.5 \pm 0.4\%$ (SGD). (Bottom Row) In the 7T brain SAR prediction, we observed an average SSIM of $90.5 \pm 3.6\%$ and an average MSE of $0.7 \pm 0.6\%$ (Adam), compared to a SSIM of $81.4 \pm 2.6\%$ and an MSE of $0.5 \pm 0.5\%$ (SGD). The higher spatial resolution of the 7T data (ie, isotropic pixel size of ~ 1.33 mm compared to 2.67 mm at 3T) resulted in improved SSIM and MSE.

predicted SAR maps using MSE. In addition, SSIM was evaluated to provide a perceptual image quality comparison.

We also introduced a hotspot analysis by masking the images only using the pixels that fell within a percentage of highest local SAR values. We evaluated the similarity in location between ground truth and prediction by use of SSIM between ground truth and predicted hotspots. We also evaluated the maximum SAR values for all hotspots and calculated the relative error between prediction and ground truth.

III. RESULTS

A. HYPERPARAMETER ANALYSIS

SSIM results and the training time of the 3 U-Nets for different hyperparameters are given in Figure 3 (Adam) for 3T body MRI. Increasing the learning rate of the 2-stage network (Figure 3A) from 10^{-6} to a higher level (eg, 10^{-3})

increased the average SSIM from $68.9 \pm 17.6\%$ to $81.1 \pm 4.8\%$. LRs chosen over a range from 10^{-3} to 10^{-5} resulted in similar SSIM. Additionally, increasing the feature map length from 32 to 64 improved the SSIM from $77.3 \pm 1.2\%$ to SSIM of $81.7 \pm 1.0\%$. The use of 128-layer feature maps led to the highest variation in SSIM (ie, $\pm 23.2\%$) among all LRs and feature map sizes for all body models. Increasing the feature map length from 16 to 256 increased the time required to train each slice from 8 to 80 ms with limited improvement in SSIM.

The architecture containing 4 stages exhibited more variation for different LRs and feature map lengths (Figure 3B). The use of 32 convolution filter layers at the initial stage generated the lowest SSIM (ie, $65.4 \pm 3.5\%$), which clearly demonstrates the importance of the length of the feature map at the initial layer. The best average SSIM, ie $83.0 \pm 9.4\%$, was reached when the LR was 10^{-5} . Although changing the initial layer feature map length from 64 to 256 increased the

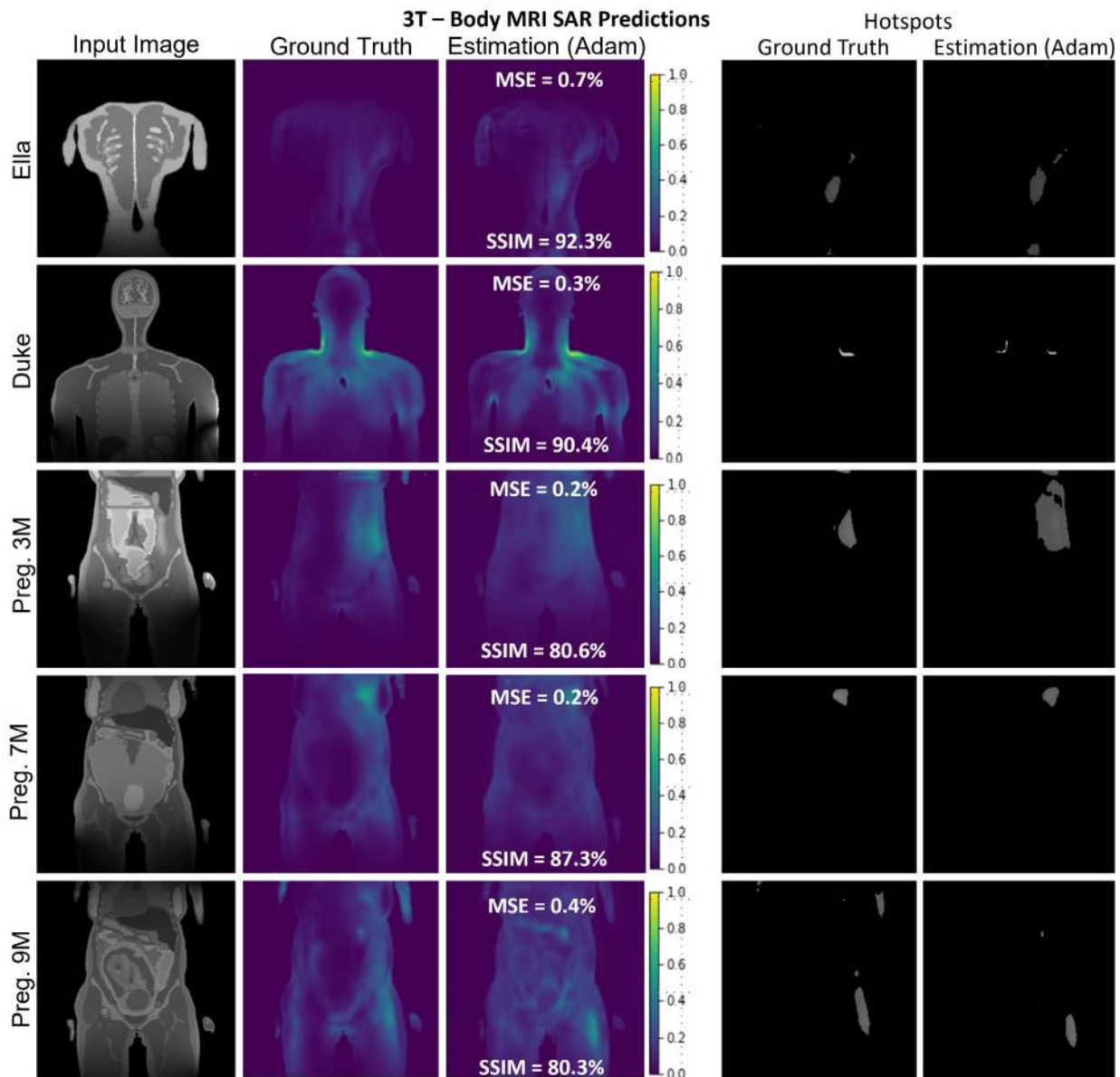
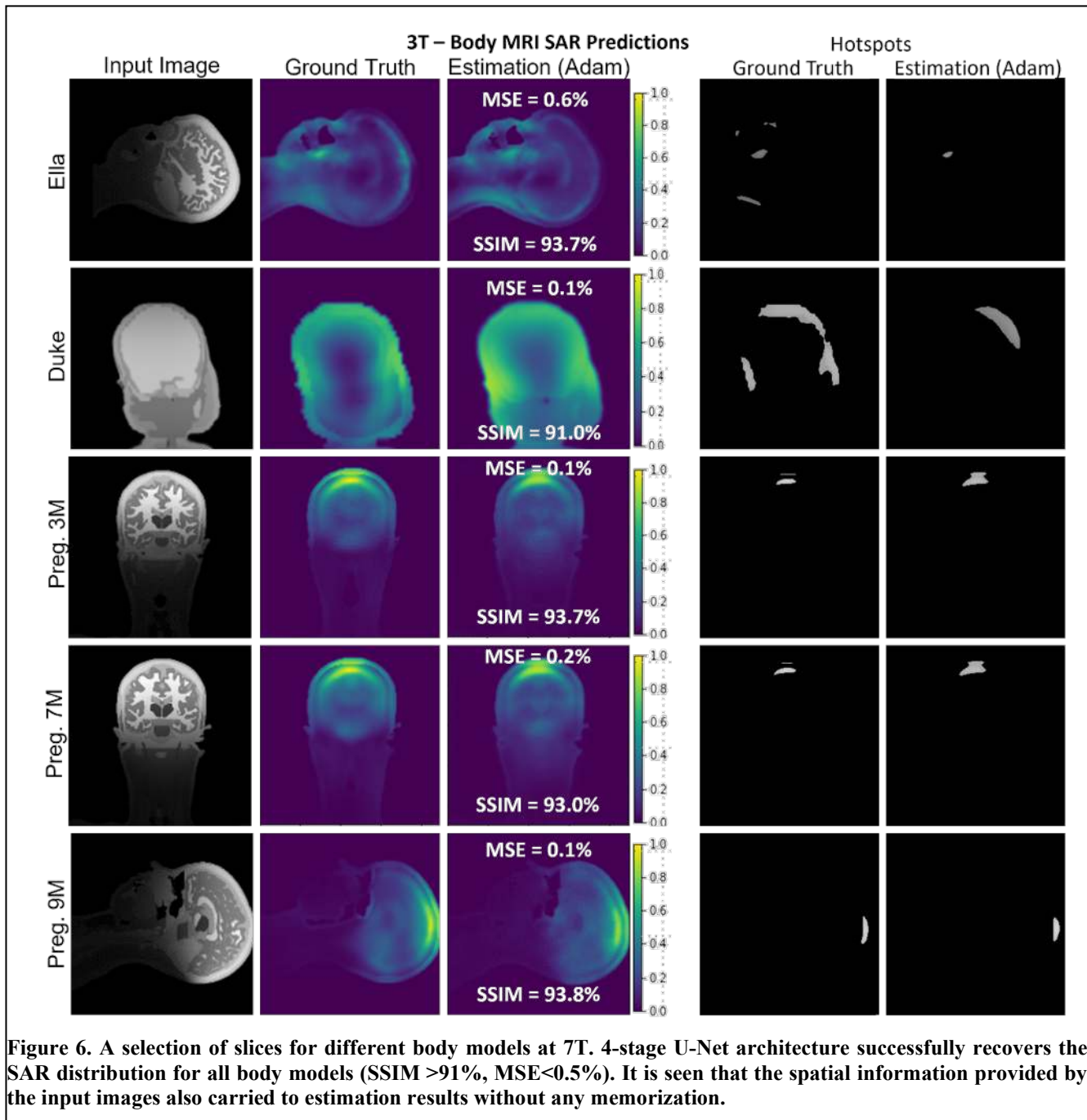


Figure 5. A selection of slices for Ella, Duke, and Pregnant Women of different gestational stages at 3T body imaging. Despite very high variation of SAR maps and input images, the proposed 4-stage U-Net architecture successfully recovered the distribution with better than 80% average SSIM and less than 0.5% average MSE for all body models.

capacity of the network, it improved the average SSIM only slightly (ie, $79.9 \pm 10.2\%$ to $85.4 \pm 5.1\%$) but decreased the dependence on hyperparameter variations significantly (ie, the variation of SSIM decreased from 10.2% to 5.1%). Additionally, the training time for the 256-layer feature map was increased by a factor >2 compared to 128 layers. Thus, among the investigated 4-stage U-Net architectures, the use of either 64 or 128 feature maps at the initial stage produced the best trade-off for all LRs.

The results for the 6-stage U-Net architecture are reported in Figure 3C. Using 32 feature map layers at the initial stage

improved the average SSIM to $70.3 \pm 2.0\%$ compared to the case with 16 layers (ie, $51.0 \pm 4.0\%$), again highlighting the importance of the initial layer size. This result was also slightly improved compared to the 4-stage U-Net (ie, $65.4 \pm 3.5\%$), emphasizing the merely incremental contribution of higher number of stages for this problem. Similarly, although the number of layers and networks capacity is much higher than in the 2-stage architecture, the 6-stage network with 32 feature maps (SSIM = $70.3 \pm 2.0\%$) was not able to outperform the 2-stage architecture with 64 feature maps (SSIM = $81.7 \pm 4.6\%$).



In summary, the initial layer feature map length was a key factor contributing to model performance and led us to the selection of a 4-stage U-Net architecture with 64 initial layer convolutional filters, using a LR of 10^{-5} .

B. PERFORMANCE EVALUATION OF DIFFERENT BODY MODELS AND OPTIMIZERS

We used the optimized 4-stage U-Net architecture to analyze performance metrics for different body models and different field strengths (Figure 4).

Adam outperformed SGD for all body models and all field strengths according to SSIM metrics (Figure 4). For 3T body SAR prediction (top row), an average SSIM of $85.1 \pm 6.2\%$ and an average MSE of $0.4 \pm 0.4\%$ was observed (Adam), as opposed to a SSIM of $69.3 \pm 4.5\%$ and an MSE of $0.5 \pm 0.4\%$ (SGD). Specifically, for Ella, we observed a SSIM of $87.9 \pm 8.0\%$ and an MSE of $0.2 \pm 0.2\%$. At the same field strength for Duke, we observed a $91.0 \pm 3.7\%$ SSIM and a $0.3 \pm 0.2\%$ MSE. Similarly, SSIM results of $81.9 \pm 4.9\%$, $81.9 \pm 7.8\%$, and $82.7 \pm 6.6\%$ were obtained for the 3-, 7-, and 9-month pregnant women, respectively. An average MSE of

0.5% with a standard deviation of $<0.5\%$ was observed for the pregnant women.

In the 7T brain SAR prediction (Figure 4, bottom row), we observed similar performance for all body models, with an average SSIM of $90.5 \pm 3.6\%$ and an average MSE of $0.7 \pm 0.6\%$ (Adam), compared to a SSIM of $81.4 \pm 2.6\%$ and an MSE of $0.5 \pm 0.5\%$ (SGD). The higher spatial resolution of the 7T data (ie, isotropic pixel size of ~ 1.33 mm compared to 2.67 mm at 3T) resulted in improved SSIM and MSE. For the baby body model Charlie, an SSIM of $92.2 \pm 1.7\%$ and an MSE of $0.8 \pm 0.5\%$ was achieved. Similarly, SSIM results of $91.3 \pm 3.6\%$, $84.1 \pm 5.8\%$, $92.5 \pm 3.6\%$, and $92.6 \pm 3.4\%$ were obtained for Ella as well as the 3-, 7-, and 9-month pregnant women, respectively.

B. HOTSPOT ANALYSIS

Visual representation showed good agreement of ground truth and predicted hotspots in 10 subjects (Figure 5 and 6, hotspot columns). Using the top 15% of local SAR values, we obtained a mean SSIM of $97.3\% \pm 2.5\%$ denoting good agreement on hotspot locations, as well as a relative mean error of $1.5\% \pm 1.1\%$ of maximum SAR values when the SAR map was normalized to the mean of the ground truth. This corresponds to normalizing the predicted SAR map to the, measurable, global SAR at the time of the patient exam. Mean errors without normalization were found to be 11%, still reasonable given the small training dataset. An added underestimation constraint that will be part of the full MRSaiFE algorithm is proposed in the discussion and is expected to help with hotspot prediction accuracy.

C. PERCEPTUAL IMAGE QUALITY ANALYSIS

Representative 3T body imaging slices randomly selected from each body model are depicted in Figure 5. Qualitatively, the 4-stage U-Net designed in this work successfully estimated the SAR distributions for the given slices.

Similarly, a selection of slices and performance metrics for 7T head imaging are depicted in Figure 6. A representative anatomical input image for Ella with very fine detail is shown and yielded a predicted SAR with low spatial detail as expected from the GT SAR simulation (top row). The proposed U-Net successfully (ie, with 93.7% SSIM and with 0.6% MSE) estimated this distribution. Similarly, a slice taken from Charlie simulations with low anatomical detail (second row) was used. The U-Net architecture successfully estimated its SAR distribution with a SSIM of 91.0% and an MSE of 0.1%. For the high-resolution head imaging simulations conducted on the 3 pregnant body models (rows 3, 4, and 5), the proposed U-Net yielded an SSIM of $\geq 93\%$. Regardless of the spatial information encoded in the anatomical input slices, the 4-stage U-Net architecture successfully recovered the SAR maps of the corresponding slices.

III. DISCUSSION

The U-Net architecture is assumed to be one of the workhorses of biomedical image segmentation and classification. In its optimization for the purposes of our application, we found that the number of convolutional filters at the input layer is tremendously important for overall network performance. From the MRI perspective, the spatial and intensity information encoded in the MR image is limited compared to electrooptical images used in daily life. This is mainly due to lower spatial resolution (ie, ~ 1 mm) and lower image sizes (eg, 256x256 pixels) in MRI. As a result, the overall network complexity of the U-Net has to be adapted to estimate SAR maps without overfitting or underfitting. In this work, we found that using a 4-stage U-Net structure with 64 or 128 convolutional filters at the initial stage results in optimal performance and reasonable training time. Additional stages without a concurrent increase in the number of filters at the first layer did not improve performance due to the information loss incurred at the first layer.

It was also seen in Figure 4 that the prediction using whole-body anatomies yielded higher inter-subject standard deviations compared to head sized imaging. This could be attributed to the higher variation of anatomical image detail at the input level.

The relatively small database used in this proof-of-concept study in conjunction with the excellent agreement between predicted SAR and simulated SAR suggested that a large-scale SAR database along with experimental verification would lead to small SAR errors for a future real-time implementation of MRSaiFE.

One of the limitations in this work is the use of MRI-like anatomical input images. The final algorithm when used at the console will use an acquired MR image. In order to estimate whether the algorithm performs similarly we will develop our algorithm further using simulated MR images from the MRI scanner simulator platform (SYSSIM) [52] offered by Sim4Life.

Future work beyond this proof-of-concept stage will include the use of experimental in vivo and in vitro thermometry data. For patient safety reasons, we will use an underestimation bound of no greater than 5% as our metric to ensure that the actual SAR never significantly exceeds predicted SAR. To this goal, we will add a weighted L1-norm term that introduces a slight overestimation, weighting the difference between ground truth and generator output.

In this application, a 3D convolutional neural network could prove useful and will be tested as part of future work because heating from neighboring slices could prove important for training performance. We will therefore analyze the performance of both 3D networks and choose the network that provides the best training performance.

Overall, MRSaiFE could eventually provide UHF MRI with consistent tissue heating monitoring for use as a safe, practical, and non-invasive mainstream tool for clinical

understanding, diagnostics, monitoring, and treatment guidance at sub-mm resolution. In practice, the existing conservative SAR margins of 20:1 could become exchangeable with optimized, patient-specific margins, and will free up valuable transmit power that can be used towards better sensitivity, resolution, or scan time.

This work could also significantly impact the safety of scanning patients with medical implants. Implants present a great cause for tissue heating concerns [29, 53, 54]. The tissue heating incurred from implants can be introduced by modifying the predicted SAR map via the implant trajectory and transfer function [55]. This ultimately would result in such patients being able to undergo MRI exams more routinely, and not only in critical situations or not at all.

The expansion of the technique to parallel transmission is straightforward: the SAR map from each channel can be predicted separately and superimposes linearly to form the total SAR. These maps can then be used in real-time SAR monitoring directly, or as an input metric in SAR-aware pTx pulse design.

The advent of even higher field strengths such as 9.4T and 10.5T, for human MRI has brought about even greater scrutiny, and valid concerns about patient safety [30, 56]. The spatial SAR variations and average global SAR are increased compared to 7T, and MRSaiFE can be of great use in bringing these technologies to clinical practice.

The effects of B_0 inhomogeneities on SAR are negligible due to their very small magnitude compared to the B_1 inhomogeneities (variations on the order of 10^{-6} for B_0 vs 10^1 for B_1) and are thus not included in this tool.

In hyperthermia, tissue heating is directed at specific tissue regions with the goal of ablation. UHF MRI, with its intrinsic short wavelength and state-of-the-art parallel transmit capability, can be used to tailor these heating hotspots by tailoring local SAR. MRSaiFE offers the potential to enable targeted treatment planning in MR hyperthermia for cancer and other diseases in the long term.

IV. CONCLUSION

We developed a proof of concept for MRSaiFE, which is an AI-based, exam-integrated, real-time MRI safety prediction software tool that facilitates the safe generation of UHF MRI images by means of accurate local SAR-monitoring at sub-W/kg levels with less than 1% MSE error. We trained the software, with a small database of images, to perform a feasibility study and have achieved successful proof of concept for both the 3T and 7T field strengths. SAR maps were predicted with a residual MSE of <1% for both 7T head imaging and 3T body imaging. A SSIM of >90% for 7T head imaging and >80% for 3T body imaging was also achieved.

REFERENCES

- [1] J. P. Lerch *et al.*, "Studying neuroanatomy using MRI," *Nature Neuroscience*, vol. 20, no. 3, pp. 314-326, 2017, doi: 10.1038/nn.4501.
- [2] E. F. McKiernan and J. T. O'Brien, "7T MRI for neurodegenerative dementias in vivo: a systematic review of the literature," *J Neurol Neurosurg Psychiatry*, vol. 88, no. 7, pp. 564-574, 2017.
- [3] M. Moccia, S. Ruggieri, A. Ianniello, A. Toosy, C. Pozzilli, and O. Ciccarelli, "Advances in spinal cord imaging in multiple sclerosis," *Therapeutic Advances in Neurological Disorders*, vol. 12, p. 175628641984059, 2019, doi: 10.1177/1756286419840593.
- [4] O. Kraff and H. H. Quick, "7T: Physics, safety, and potential clinical applications," (in en), *Journal of Magnetic Resonance Imaging*, vol. 46, no. 6, pp. 1573-1589, 2017 2017, doi: 10.1002/jmri.25723.
- [5] K. Uğurbil *et al.*, "Brain imaging with improved acceleration and SNR at 7 Tesla obtained with 64-channel receive array," *Magnet Reson Med*, vol. 82, no. 1, pp. 495-509, 2019, doi: 10.1002/mrm.27695.
- [6] C. M. Collins and M. B. Smith, "Calculations of B1 distribution, SNR, and SAR for a surface coil adjacent to an anatomically-accurate human body model," *Magnetic Resonance in Medicine*, vol. 45, no. 4, pp. 692-699, 2001, doi: 10.1002/mrm.1092.
- [7] P. P. Winkler SA, Thornton MM, Rutt BK, "Method and System for Estimating the Specific Absorption Rate of a Tissue Region Prior to a Magnetic Resonance Imaging Scan," 2015.
- [8] S. A. Winkler, P. A. Picot, M. M. Thornton, and B. K. Rutt, "Direct SAR mapping by thermoacoustic imaging: A feasibility study," *Magnet Reson Med*, vol. 78, no. 4, pp. 1599-1606, 2017, doi: 10.1002/mrm.26517.
- [9] J. W. Hand, R. W. Lau, J. J. W. Lagendijk, J. Ling, M. Burl, and I. R. Young, "Electromagnetic and thermal modeling of SAR and temperature fields in tissue due to an RF decoupling coil," *Magnetic Resonance in Medicine*, vol. 42, no. 1, pp. 183-192, 1999, doi: 10.1002/(sici)1522-2594(199907)42:1<183::aid-mrm24>3.0.co;2-i.
- [10] O. P. Gandhi and X. B. Chen, "Specific absorption rates and induced current densities for an anatomy-based model of the human for exposure to time-varying magnetic fields of MRI," *Magnet Reson Med*, vol. 41, no. 4, pp. 816-823, 1999, doi: 10.1002/(sici)1522-2594(199904)41:4<816::aid-mrm22>3.0.co;2-5.
- [11] D. Simunic, P. Wach, W. Renhart, and R. Stollberger, "Spatial distribution of high-frequency electromagnetic energy in human head during MRI: numerical results and measurements," *IEEE Transactions on Biomedical Engineering*, vol. 43, no. 1, p. 88, 1996, doi: 10.1109/10.477704.
- [12] S. Oh, Y.-C. Ryu, G. Carluccio, C. T. Sica, and C. M. Collins, "Measurement of SAR-induced temperature increase in a phantom and in vivo with comparison to numerical simulation," *Magnetic Resonance in Medicine*, vol. 71, no. 5, pp. 1923-1931, 2014, doi: 10.1002/mrm.24820.
- [13] U. Katscher, T. Voigt, C. Findekklee, P. Vernickel, K. Nehrke, and O. Dossel, "Determination of Electric Conductivity and Local SAR Via B1 Mapping," *IEEE Transactions on Medical Imaging*, vol. 28, no. 9, pp. 1365-1374, 2009, doi: 10.1109/tmi.2009.2015757.
- [14] S. A. Winkler and B. K. Rutt, "Practical methods for improving B1+ homogeneity in 3 tesla breast imaging," *Journal of Magnetic Resonance Imaging*, vol. 41, no. 4, pp. 992-999, 2015, doi: 10.1002/jmri.24635.
- [15] Y. Zhu, "Parallel excitation with an array of transmit coils," *Magnetic Resonance in Medicine*, vol. 51, no. 4, pp. 775-784, 2004, doi: 10.1002/mrm.20011.
- [16] U. Katscher, P. Börner, C. Leussler, and J. S. Van Den Brink, "Transmit SENSE," *Magnetic Resonance in Medicine*, vol. 49, no. 1, pp. 144-150, 2003, doi: 10.1002/mrm.10353.
- [17] M. Pendse, R. Stara, M. Mehdi Khalighi, and B. Rutt, "IMPULSE: A scalable algorithm for design of minimum

- specific absorption rate parallel transmit RF pulses," *Magnet Reson Med*, vol. 81, no. 4, pp. 2808-2822, 2019.
- [18] B. Guérin, K. Setsompop, H. Ye, B. A. Poser, A. V. Stenger, and L. L. Wald, "Design of parallel transmission pulses for simultaneous multislice with explicit control for peak power and local specific absorption rate," *Magnetic Resonance in Medicine*, vol. 73, no. 5, pp. 1946-1953, 2015, doi: 10.1002/mrm.25325.
- [19] IEC60601-2-33, "Medical electrical equipment - Part 2-33: Particular requirements for the basic safety and essential performance of magnetic resonance equipment for medical diagnosis," International Standard 2010-03-10.
- [20] SiemensHealthineers. "FDA Clears MAGNETOM Terra 7T MRI Scanner from Siemens Healthineers." <https://www.siemens-healthineers.com/en-us/news/magnetomterrafdaclearance.html> (accessed).
- [21] H. GE. "Bringing Ultra-High Field MR Imaging from Research to Clinical: SIGNA 7.0T FDA Cleared." https://www.accessdata.fda.gov/cdrh_docs/pdf20/K202966.pdf (accessed).
- [22] V. Rieke and K. Butts Pauly, "MR thermometry," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 2, pp. 376-390, 2008, doi: 10.1002/jmri.21265.
- [23] Q. Tian *et al.*, "Improving in vivo human cerebral cortical surface reconstruction using data-driven super-resolution," *Cerebral Cortex*, vol. 31, no. 1, pp. 463-482, 2021, doi: 10.1093/cercor/bhaa237.
- [24] B. Sveinsson *et al.*, "Synthesizing Quantitative T2 Maps in Right Lateral Knee Femoral Condyles from Multicontrast Anatomic Data with a Conditional Generative Adversarial Network," (in En), *Radiology: Artificial Intelligence*, 2021/05/26/ 2021, doi: 10.1148/ryai.2021200122.
- [25] K. E. Keenan *et al.*, "Challenges in ensuring the generalizability of image quantitation methods for MRI," *Med Phys*, 2021, doi: 10.1002/mp.15195.
- [26] A. S. Chaudhari *et al.*, "Prospective Deployment of Deep Learning in MRI : A Framework for Important Considerations, Challenges, and Recommendations for Best Practices," *Journal of Magnetic Resonance Imaging*, vol. 54, no. 2, pp. 357-371, 2021, doi: 10.1002/jmri.27331.
- [27] A. S. Chaudhari *et al.*, "Low-count whole-body PET with deep learning in a multicenter and externally validated study," *npj Digital Medicine*, vol. 4, no. 1, 2021, doi: 10.1038/s41746-021-00497-2.
- [28] A. S. Chaudhari *et al.*, "Diagnostic Accuracy of Quantitative Multicontrast 5-Minute Knee MRI Using Prospective Artificial Intelligence Image Quality Enhancement," *Am J Roentgenol*, vol. 216, no. 6, pp. 1614-1625, 2021, doi: 10.2214/ajr.20.24172.
- [29] P. R. Stijnman *et al.*, "Accelerating implant RF safety assessment using a low-rank inverse update method," *Magnet Reson Med*, vol. 83, no. 5, pp. 1796-1809, 2020.
- [30] A. Sadeghi-Tarakameh *et al.*, "In vivo human head MRI at 10.5T: A radiofrequency safety study and preliminary imaging results," *Magnet Reson Med*, vol. 84, no. 1, pp. 484-496, 2020, doi: 10.1002/mrm.28093.
- [31] J. Kemnitz *et al.*, "Clinical evaluation of fully automated thigh muscle and adipose tissue segmentation using a U-Net deep learning architecture in context of osteoarthritic knee pain," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 33, no. 4, pp. 483-493, 2020, doi: 10.1007/s10334-019-00816-5.
- [32] D. Fürst, W. Wirth, A. Chaudhari, and F. Eckstein, "Layer-specific analysis of femorotibial cartilage t2 relaxation time based on registration of segmented double echo steady state (dess) to multi-echo-spin-echo (mese) images," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 33, no. 6, pp. 819-828, 2020, doi: 10.1007/s10334-020-00852-6.
- [33] F. Eckstein *et al.*, "A Deep Learning Automated Segmentation Algorithm Accurately Detects Differences in Longitudinal Cartilage Thickness Loss – Data from the FNIH Biomarkers Study of the Osteoarthritis Initiative," *Arthritis Care & Research*, 2020, doi: 10.1002/acr.24539.
- [34] A. S. Chaudhari *et al.*, "Utility of deep learning super-resolution in the context of osteoarthritis MRI biomarkers," *Journal of Magnetic Resonance Imaging*, vol. 51, no. 3, pp. 768-779, 2020, doi: 10.1002/jmri.26872.
- [35] Arjun *et al.*, "The International Workshop on Osteoarthritis Imaging Knee MRI Segmentation Challenge: A Multi-Institute Evaluation and Analysis Framework on a Standardized Dataset," *arXiv pre-print server*, 2020-05-26 2020, doi: None arxiv:2004.14003.
- [36] A. S. Chaudhari *et al.*, "Super-resolution musculoskeletal MRI using deep learning," *Magnet Reson Med*, vol. 80, no. 5, pp. 2139-2154, 2018, doi: 10.1002/mrm.27178.
- [37] E. F. Meliàdo *et al.*, "A deep learning method for image-based subject-specific local SAR assessment," (in en), *Magnet Reson Med*, vol. 83, no. 2, pp. 695-711, 2020 2020, doi: 10.1002/mrm.27948.
- [38] E. F. Meliàdo, C. A. T. Van Den Berg, P. R. Luijten, and A. J. E. Raaijmakers, "Intersubject specific absorption rate variability analysis through construction of 23 realistic body models for prostate imaging at 7T," *Magnet Reson Med*, vol. 81, no. 3, pp. 2106-2119, 2019, doi: 10.1002/mrm.27518.
- [39] E. F. Meliàdo *et al.*, "A deep learning method for image-based subject-specific local SAR assessment," *Magnet Reson Med*, vol. 83, no. 2, pp. 695-711, 2020, doi: 10.1002/mrm.27948.
- [40] S. Mandija, E. F. Meliàdo, N. R. F. Huttinga, P. R. Luijten, and C. A. T. Van Den Berg, "Opening a new window on MR-based Electrical Properties Tomography with deep learning," *Scientific Reports*, vol. 9, no. 1, 2019, doi: 10.1038/s41598-019-45382-x.
- [41] J. M. Kabil, S. Geethanath, and J. T. Vaughan, "Intra- and inter- brain RF heating prediction with the Non-Invasive Temperature Estimation (NITE) method," 2020, 4191. [Online]. Available: <https://archive.ismrm.org/2020/4191.html>.
- [42] S. A. Winkler, I. Saniour, A. Chaudhari, F. Robb, and J. T. Vaughan, "MRSaiFE: Towards the Real-Time Prediction of Tissue Heating in MRI – a Feasibility Study," *arXiv pre-print server*, 2021-02-01 2021, doi: None arxiv:2102.01023.
- [43] S. A. Winkler, E. Motovilova, S. Gokyar, I. Saniour, F. Robb, and A. Chaudhari, "MRSaiFE: towards the real-time prediction of tissue heating in MRI - a feasibility study," in *Proc Intl. Soc. Mag. Reson. Med.*, May 15-20 2021, vol. 2486: ISMRM, p. 2486.
- [44] S. A. Winkler, I. Saniour, F. Robb, Chaudhari, and J. T. Vaughan, "MRSaiFE: towards the real-time prediction of tissue heating in MRI - a feasibility study.,," in *Proc. IEEE IMBIOC*, Toulouse, France, 2020.
- [45] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, 2004, doi: 10.1109/tip.2003.819861.
- [46] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *arXiv pre-print server*, 2015-05-18 2015, doi: None arxiv:1505.04597.
- [47] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *ICML*, 2010.
- [48] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *arXiv pre-print server*, 2015-03-02 2015, doi: None arxiv:1502.03167.
- [49] K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification," *arXiv pre-print server*, 2015-02-06 2015, doi: None arxiv:1502.01852.
- [50] Diederik and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv pre-print server*, 2017-01-30 2017, doi: None arxiv:1412.6980.

- [51] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998, doi: 10.1109/5.726791.
- [52] Z. Cao *et al.*, "Bloch-based MRI system simulator considering realistic electromagnetic fields for calculation of signal, noise, and specific absorption rate," *Magnet Reson Med*, vol. 72, no. 1, pp. 237-247, 2014, doi: 10.1002/mrm.24907.
- [53] M. Kozlov, W. Kainz, and L. Daniel, "Influence of a Metallic Shield on RF-Induced Heating of a Lead with Straight and Helical Wires," 2019 2019: IEEE, doi: 10.1109/nemo.2019.8853734. [Online]. Available: <https://dx.doi.org/10.1109/nemo.2019.8853734>
- [54] S. Gokyar, A. Alipour, E. Unal, E. Atalar, and H. V. Demir, "Magnetic Resonance Imaging Assisted by Wireless Passive Implantable Fiducial e-Markers," *IEEE Access*, vol. 5, pp. 19693-19702, 2017.
- [55] S.-M. Park, R. Kamondetdacha, and J. A. Nyenhuis, "Calculation of MRI-induced heating of an implanted medical lead wire with an electric field transfer function," *Journal of Magnetic Resonance Imaging*, vol. 26, no. 5, pp. 1278-1285, 2007, doi: 10.1002/jmri.21159.
- [56] T. Vaughan *et al.*, "9.4T human MRI: Preliminary results," *Magnet Reson Med*, vol. 56, no. 6, pp. 1274-1282, 2006, doi: 10.1002/mrm.21073.