B1 field map synthesis with generative deep learning used in the design of parallel-transmit RF pulses for ultra-high field MRI

Boris Eberhardt[1,2], **Benedikt A. Poser**[3], **N. Jon Shah**[1,4,5,6], **Jörg Felder**[1,2,*]

[1] Institute of Neuroscience and Medicine 4, Forschungszentrum Jülich, Jüich, Germany
[2] RWTH Aachen University, Aachen, Germany
[3] Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands
[4] Institute of Neuroscience and Medicine 11, Forschungszentrum Jülich, Jülich, Germany
[5] Department of Neurology, RWTH Aachen University, Aachen, Germany
[6] JARA-BRAIN, Translational Medicine, Aachen, Germany

Eingegangen am 10. August 2021; akzeptiert am 23. Dezember 2021

Abstract

Spoke trajectory parallel transmit (pTX) excitation in ultra-high field MRI enables B_1^+ inhomogeneities arising from the shortened RF wavelength in biological tissue to be mitigated. To this end, current RF excitation pulse design algorithms either employ the acquisition of field maps with subsequent non-linear optimization or a universal approach applying robust pre-computed pulses. We suggest and evaluate an intermediate method that uses a subset of acquired field maps combined with generative machine learning models to reduce the pulse calibration time while offering more tailored excitation than robust pulses (RP).

The possibility of employing image-to-image translation and semantic image synthesis machine learning models based on generative adversarial networks (GANs) to deduce the missing field maps is examined. Additionally, an RF pulse design that employs a predictive machine learning model to find solutions for the non-linear (two-spokes) pulse design problem is investigated.

As a proof of concept, we present simulation results obtained with the suggested machine learning approaches that were trained on a limited data-set, acquired in vivo. The achieved excitation homogeneity based on a subset of half of the B_1^+ maps acquired in the calibration scans and half of the B_1^+ maps synthesized with GANs is comparable with state of the art pulse design methods when using the full set of calibration data while halving the total calibration time. By employing RP dictionaries or machine-learning RF pulse predictions, the total calibration time can be reduced significantly as these methods take only seconds or milliseconds per slice, respectively.

Keywords: MRI, RF pulse design, Machine learning, Parallel transmission

* Corresponding author: Jörg Felder, Forschungszentrum Jülich, 52425 Jülich, Germany.
E-mail addresses: b.eberhardt@fz-juelich.de (B. Eberhardt), benedikt.poser@maastrichtuniversity.nl (B.A. Poser), n.j.shah@fz-juelich.de (N.J. Shah), j.felder@fz-juelich.de (J. Felder).

Z Med Phys **32** (2022) 334–345
https://doi.org/10.1016/j.zemedi.2021.12.003
www.elsevier.com/locate/zemedi

1 Introduction

Magnetic resonance imaging at ultra-high field (UHF) has the potential to provide a higher signal-to-noise ratio (SNR) and increased spatial resolution [1]. However, these benefits are not without cost, as increased field strengths lead to increased RF inhomogeneities and higher energy deposition with a correspondingly increased specific absorption rate (SAR) burden [1–3]. The application of pTX [4,5] can be used to remedy these drawbacks as it offers more degrees of freedom with which to tailor the spatial excitation. This is because the "spoke" k-space trajectory [6] of the parallel transmitted pulses can compensate for inhomogeneities in both the excitation B_1^+ field [7] and the B_0 field on a slice-selective basis. A magnitude least squares (MLS) RF pulse design [8] offers additional degrees of freedom as it optimizes for a homogeneous excitation field magnitude without placing constraints on the excitation phase. However, the magnitude operation is non-linear and thus casts the optimization problem into non-convex space. Thus, the optimization becomes prone to converge into local minima that may exhibit signal voids [9] if no suitable starting points are supplied [10,11]. Safety relevant SAR and hardware constraints (e.g. maximum transmit power per channel) are often imposed in MLS-based pulse designs [12], either as regularization terms or in the form of explicit constraints. In addition to optimizing for channel weights only, joint designs of RF pulse-weights and transmit k-space trajectory are common strategies for optimization [10,13,11].

As an alternative to such a patient-specific design of pTX excitation pulses, the concept of "universal pulses" has been proposed for brain imaging [14,15]. These are pre-calculated pulses, optimized over a wide range of brain anatomies, so that tissue can be excited in a broad scope of applications without requiring the measurement of B_1^+ or B_0 maps. The universality of these pulses is offset against reduced excitation homogeneity and slightly higher SAR-burden when compared with patient-specific designs. Recently, so-called fast online-customized (FOCUS) pTX pulses have been proposed that combine universal pulses UP, based on a spiral non-selective k-space trajectory, with a subsequent subject-specific pulse optimization [16].

Deep neural networks can approximate complex, non-linear functions using a composition of simple non-linear functions [17]. Usually, in discriminative learning, a mapping from input data, $\boldsymbol{x}$, to a probability of a class label y, i.e. the conditional probability, $p(y|\boldsymbol{x})$, is learned [18]. Classification and regression problems are frequently employed applications in discriminative learning [19]. Conversely, in generative models, the joint probability, $p(\boldsymbol{x}, y)$, is learned [20,21], as well as the probability distribution $p(\boldsymbol{x})$ of the input data itself, or at least an approximation [20]. This makes it possible to generate new, yet unseen data from the learned probability distribution $p(\boldsymbol{x})$. Examples of generative models include variational autoencoders (VAEs) [22] and GANs [23].

In the context of pTX pulse design in MRI, machine learning models have been employed in prior publications. One recent study [24] describes patient-specific pulse prediction with an iterative kernel ridge regression algorithm to achieve SAR-efficient RF-shimming in the head. An example of discriminative learning applied to pTX is the k_T-points [25] pulse design algorithm presented in [26]. The method, termed "smart pulse" [26], employs a machine learning classifier to predict the most suitable pulse from population-based clustered sets of RP for calibration-free dynamic abdominal RF shimming at a static field strength of 3 T and using two transmit channels. GANs have also been previously employed in combination with MRI. A neural network-based pulses design has been investigated in [27] in which the authors train the model to predict single-transmit 2D-RF pulses with the goal to compensate for B_1^+ and B_0 inhomogeneities. In [28] (arXiv preprint) a new method for image reconstruction from sub-sampled k-space data is proposed that utilizes both GRAPPA and GANs. The GAN-enhanced method outperformed the GRAPPA-only version with improved peak signal to noise and root mean square error (RMSE) on the fastMRI dataset [29]. The benefits of GANs have also been highlighted by other studies. For example, GANs were used in [30] for the synthesis of multimodal imaging data, in [31] to predict the local SAR distribution in a subject-specific way based on B_1^+ maps and in [32] for the reconstruction of MR images. In [33] a convolutional neural network is trained for SAR reduction of adiabatic RF pulses to be used in a 2D T_2-FLAIR sequence at 7 T by predicting the required slice-specific RF pulse power and RF amplitude scaling factors.

To create patient-specific tailored RF pulses with multiple transmit channels, an initial calibration procedure that acquires subject-specific B_1^+ and B_0 fields [2] is usually required. This initial calibration procedure may take several minutes by itself [1]; therefore the use of RP [14] or "smart-pulses", which select a set of UP using a machine learning classifier, are a means to save time in the clinic. Another approach for increased time efficiency is a slab-wise pulse design, as proposed in [34].

Here we examine an intermediate approach between full calibration and no calibration as a trade-off between universality and patient-tailored pulse design. With generative image to image translation models such as pix2pixHD [35] or spatially-adaptive (de)normalization (SPADE) [36], both of which employ GANs, we propose synthesizing the missing B_1^+ field maps, which can then be input into the pTX pulse design algorithm in combination with measured maps. This only requires the measurement of a subset of all transmit channels and therefore saves time during pTX measurements.

In order to evaluate the quality of the synthesized B_1^+ maps, the sub-sampled measured maps are completed by artificial maps from the GANs and are used as the input for an MLS based two spokes pulse design that optimizes both the channel weights as well as spoke locations [8,37,11]. The performance

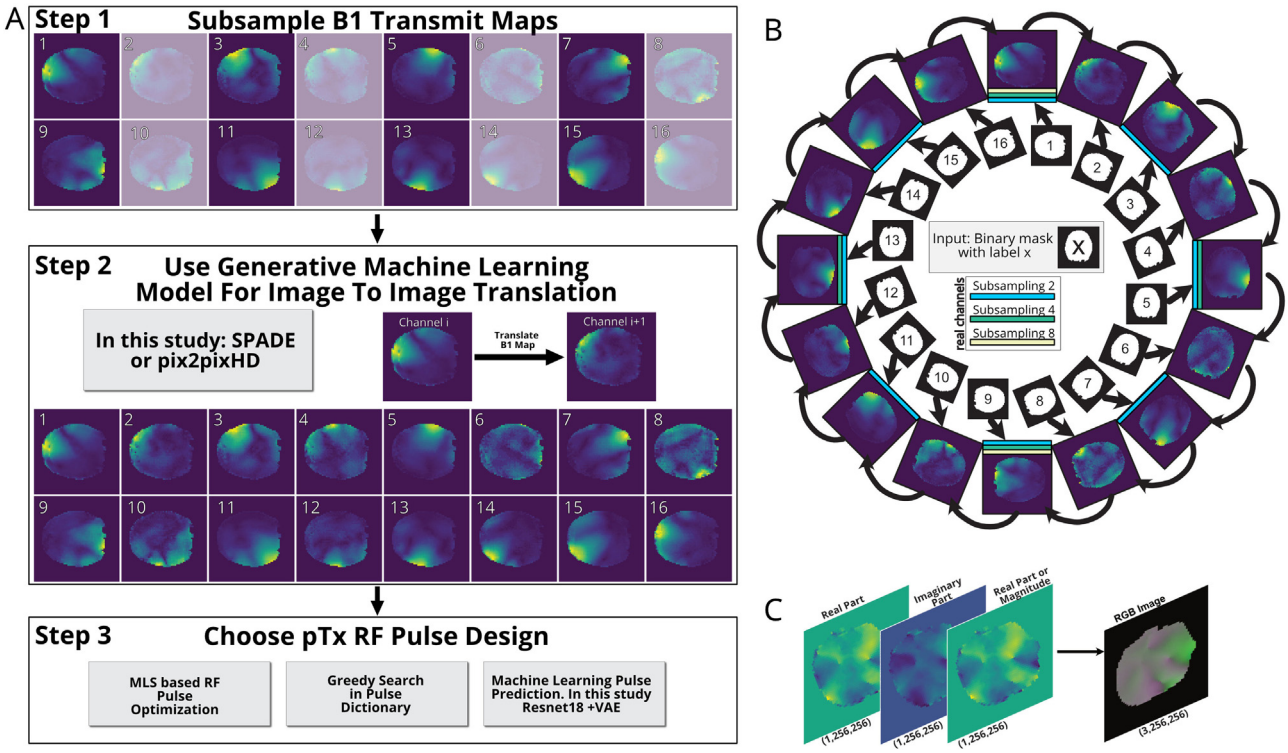

Figure 1. (A) Overview of the proposed method that consists of three main steps that include the initial subsampling (shown is a subsampling factor of 2), the subsequent second image synthesis step to generate the missing data and the third pulse design step. In this work, three possible pulse design methods are compared for the third step. (B) B_1^+ maps, masks and labels of a 16-channel pTX coil array. The arrows indicate the direction of image-to-image translation. The measured maps selected as input for the GAN are indicated by colored bars for the different sub-sampling factors. (C) Encoding of the B_1^+ maps into the RGB space required by the image-to-image translation algorithms. The real and imaginary parts of the B_1^+ maps are encoded in the R and G channel, respectively. The third (B) channel, chosen arbitrarily from multiple alternative possibilities, was also initialized with the real part (pix2pixHD) or magnitude (SPADE) of the field maps.

of the partial B_1^+ synthesis is also compared against a novel pTX RF pulse design which make use of machine learning architectures.

2 Methods

2.1 Image-to-image translation with conditional adversarial networks

GANs can be used to translate images from one domain into another. While unconditional GANs used for image generation create arbitrary samples from a noise vector, z (the latent vector that is sampled from a normal distribution) alone [23], a conditional GAN generates images both from z and an additional input image. This allows more control over the generated output, thus making it appropriate for image translation tasks [38].

Several algorithms for image-to-image translation have been proposed that are conditioned on an input image to generate a dependent output image. Amongst these, the "pix2pix" algorithm [38] is an example that uses GANs. It uses a U-NET architecture as its generator [39], proposes a GAN objective function that is very general and ensures that only results with an "authentic" appearance are generated. This is in contrast to objective functions that minimize the Euclidean distance between the images, which leads, in general, to more pronounced image blurring [38]. In pix2pixHD, [35] the pix2pix model is improved to give higher resolution results and enables a better modeling of both global and local image features in high-resolution images. The U-NET of pix2pix is replaced with a coarse-to-fine generator in pix2pixHD that outperforms the former by a wide margin as shown in [35]. With the pix2pixHD model, it is also possible to generate realistic images from label maps alone. These are masks that contain the class labels corresponding to the type of image to be generated. This method of semantic image synthesis was further improved in [36] with changes to the model architecture and the application of a novel normalization method termed spatially-adaptive (de)normalization or SPADE, which was subsequently used as the model's name.

The proposed method is summarized in Figure 1A. It consists of three main steps: The subsampling of the B_1^+ data with different subsampling factors is step one. For step two, the

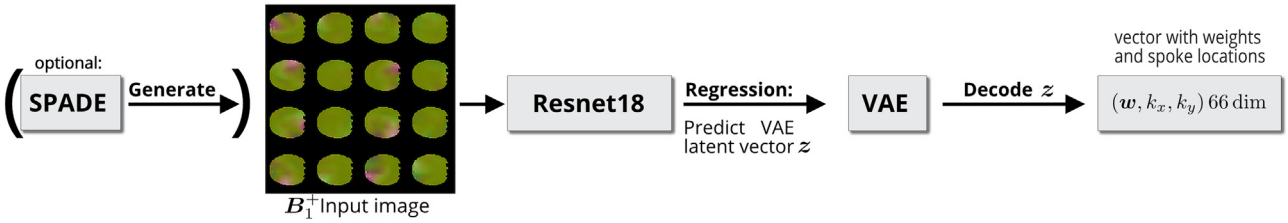

Figure 2. The proposed machine learning based RF pulse prediction algorithm. The Resnet18 model can be employed as a regression model with an MSE objective function. Its input is a PNG image that contains the real part, the imaginary part, and the magnitude of all 16 B_1^+ transmit channels, respectively. It predicts a latent vector 'z' that can be decoded by the trained VAE into weights '$\boldsymbol{w}$' and spoke locations 'k_x, k_y'. The SPADE-model, shown in parenthesis, may be plugged into the pulse design algorithm in the case that only one single B_1^+ is acquired and the others are to be generated.

image translation is carried out, and two such models to attain this are considered in the work, i.e. pix2pixHD and SPADE. In the third step, a pTX RF pulse design algorithm can be chosen that depends on a full set of B_1^+ maps. In this work, MLS based RF pulse algorithms from the literature are considered. Additionally, the use of pre-computed pulse dictionaries and machine learning pulse prediction models are investigated. Figure 1B illustrates the image-to-image translation task for B_1^+ map augmentation from a subset of measured maps. The generation of neighboring channel maps is proposed using a measured map as input (pix2pixHD) or a measured map and a channel mask + label (channel number) for the SPADE version of the algorithm. Several sub-sampling factors using only every 2nd, 4th or 8th measured B_1^+ map of a 16-channel pTX coil to generate the missing maps were investigated.

As the training of neural networks requires a large data set (with input and corresponding output) in order to achieve good performance existing maps from four healthy subjects were artificially augmented. For this purpose, the measured in vivo B_1^+ maps, acquired with the DREAM sequence [40,41], and the B_0 maps acquired with the double-echo method [42], were linearly interpolated to 5360 B_1^+ maps per channel for a total of 85,760 images for 16 transmit channels. As image-to-image translation is based on real valued images but the B_1^+ maps are complex valued; these maps were separated into their real and imaginary parts. The parts were used to compose a single RGB image as shown in Figure 1C. Finally, all RGB images were normalized to unit intensity prior to training the neural networks.

The published code repositories of the original pix2pixHD and SPADE authors were used for the implementation of the image translation [43] and the models were trained from scratch. The pix2pixHD model was trained for 24 epochs on a paired dataset compromising 85,760 B_1^+ maps stored in images of size 256×256 pixels and encoded in the portable network graphics (PNG)-format as this is the minimum resolution supported by the model. Training was carried out for image translation from transmit channel i to its nearest neighbor $i+1$. The default model settings of the code repository were not changed and the batch size was set to 8. The SPADE model was trained for 23 epochs on the same paired dataset with additional corresponding label maps. The number of filters in the SPADE generator was reduced from the default 128 to 64. Its activation norm type and the global adaptive norm type were set to batch instead of sync_batch and the batch size was set to 4 to enable training on a single memory limited GPU (11 GB) (RTX 2080ti GPU, Nvidia Corp., Santa Clara, California, USA).

2.2 Spoke pulse design algorithm

The effect of sub-sampled mapping data and its augmentation to a complete set of maps using GANs must be evaluated against a spoke design that utilizes all 16 measured B_1^+ maps. This was carried out and is discussed later. However, since multiple deep learning approaches have already been employed during the B_1^+ map generation process, an extension of the same approach to create spoke pulses and suitable transmit k-space locations is straightforward and was also included in the evaluation process of the image-to-image translation method presented above.

A machine learning based two-spokes RF pulse prediction algorithm, as visualized in Figure 2 is investigated. The algorithm employs the "Resnet18" model [44], a convolutional neural network, as implemented in the Torchvision library of Pytorch [45] to predict suitable channel weights and spoke locations.

Resnet18 is trained as a regression model with MSE loss to learn pulse data. The vector of complex channel weights $\boldsymbol{w}$ for two spokes has a length of 32 complex values or a length of 64 when decomposed into real and imaginary parts. Together with the two spoke scalar values for the second spoke, k_x, k_y, the dimension of the desired weights-spokes vectors is 66. However, here, to reduce the dimension of the vectors, the data were compressed into a 16-dimensional latent representation, z, that becomes the Resnet18 prediction target. The motivation for reducing the target dimensionality is to possibly improve the Resnet18 prediction quality and enable exploitation of a disentangled latent space representation of the VAE [20]. The decoding of the compressed data back into full dimensionality was accomplished with a VAE trained to encode pulse data into the latent representation z and to decode z back into pulses.

Once the VAE has been trained, the compressed latent representations can be decoded back into pTX channel weights and spoke locations, $\boldsymbol{w}, k_x, k_y$.

To create the $\boldsymbol{w}, k_x, k_y$ data for training the Resnet18 model, an evolution strategy (ES)-based optimization [11] was carried out for each of the 5360 slices, and the resulting optimized $\boldsymbol{w}, k_x, k_y$ data was saved as the training data set. The model takes PNG images that contain all 16 B_1^+ transmit channel maps as input, with the real part, the imaginary part, and the magnitude of the B_1^+ map encoded in the three color channels, compare Figure 2. The Resnet18 training targets were the 5360 latent vectors $\boldsymbol{z}$ that are VAE encodings of the optimized weights and spoke locations, $\boldsymbol{w}, k_x, k_y$. These VAE encodings $\boldsymbol{z} = \mathrm{VAE}(\boldsymbol{w}, k_x, k_y)$ were predicted by the Resnet18 model and decoded into $(\boldsymbol{w}, k_x, k_y)$ by the same trained VAE to make them actually useable.

The VAE was implemented as described in [46]. To create optimized pulse data for the VAE training, pulses were calculated on the same 5, 360 slices. However, rather than saving the single minimum RMSE pulse per slice, multiple pulse sets that were close to the minimum RMSE solution were saved. By employing multiple pulse algorithms as described in [11], a larger variety of locally optimized pulse data can be created. The training set, therefore, consists of optimal channel weights $\boldsymbol{w}$ and optimal spoke locations (k_x, k_y) obtained from each of the conventional design algorithms. As multiple locally optimal solutions exist per slice, a total of a million pulses were created with this method. The VAE was trained on this training dataset until its objective function had converged to a local minimum after 160 epochs.

For the sake of simplicity, training data were encoded as small PNG images with RGB color channel dimensions 16×16, obtained by reshaping the $64 = 8 \times 8$ weight vectors into dimensions of 16×16 weight vectors for the R-channel and setting the G channel pixels equal k_x and the B channel pixels equal to k_y. The dimension of this training data is a hyper-parameter which was chosen to accommodate all RF pulse weights and spoke locations. The output of the Resnet18 model needs to match the dimensions of the VAE latent-space.

If the SPADE module depicted in parenthesis in Figure 2 is added into the pulse generation workflow, it can generate the missing 15 B_1^+ channel maps required for any spatial domain based pulse design approach [47] from a single measured B_1^+ channel map with the help of the 15 label maps (sub-sampling factor of 16). The label maps are all based on the same binary mask required for the pulse design. Therefore, the creation of the label maps is trivial once a single binary mask is available, as visualized in Figure 1. The SPADE module is identical to that described in the previous section.

All networks were specifically trained on data from the selected coil array and the associated number of spoke pulses. Although a new model would need to be trained if a different coil array and different spoke designs were used, it would not be problematic as the training of the GANs can be carried out with a limited number of in vivo data and pre-trained networks for B_1^+ data augmentation could be supplied from the coil array manufacturers. As is the case for the image translation algorithms, all algorithms described in this section were trained on the same single RTX 2080ti GPU described above. All models were trained for the 16-channel pTX coil [48–50] mentioned above operating in a Magnetom Siemens 9.4 T MRI system (Siemens Healthineers, Erlangen, Germany).

2.3 Creation of robust pulses

In order to create robust pulses, data comprising all interpolated single-slice B_1^+ maps, as described above, were utilized. Random subsets of the slices were chosen to either optimize 2, 3, or 4 slices simultaneously using the method described in [51] and in conjunction with an MLS ES-based RF pulse optimization [11] until a set of 1000 robust pulses was created. From the set of 1000 pulses, the top 10, 20, 50 and 100 best-performing pulses (based on their RMSE values on the interpolated B_1^+ slices) were taken to create four robust pulse dictionaries. The process of robust pulse creation and sorting into dictionaries is visualized in Figure 3.

2.4 Performance evaluation

To evaluate the effect of the pix2pixHD and SPADE data augmentation on the performance of the pTX pulses, different sub-sampling factors were investigated: 1 – all 16 B_1^+ maps were measured so that no sub-sampling was employed, 2 – eight measured B_1^+ maps were augmented to the full set of 16 maps resulting in a factor of two sub-sampling, 4 – four measured maps were used to generate the 16-channel B_1^+ maps of the coil which were equal to a sub-sampling factor of four, 8 – sub-sampling factor equal to eight, and 16 – sub-sampling factor equal to 16. Three cases were considered:

(i) For all considered sub-sampling factors a pulse optimization was performed with B_1^+ synthesis of the missing transmit channels in order to give a full set of 16 B_1^+ maps, as required by any MLS [8] based pulse optimization, such as the spatial domain method [47]. The pulse optimization employs MLS-based pulse designs that utilize ES to search for optimized solutions as described in [11] to optimize both channel weights and spoke locations simultaneously.

(ii) For a sub-sampling factor of 16 and without any pulse optimization: A full B_1^+ synthesis of all remaining transmit channels with SPADE was performed and one robust pulse from the full dictionary of saved pulses, as described in section "Creation of Robust Pulses", was chosen by evaluating all pulses contained in the dictionary and choosing the minimum RMSE pulse.

(iii) The proposed RF machine learning pulse design method predicts latent vectors, $\boldsymbol{z}$. A trained VAE then decodes these into RF pulses without further optimization.

The reported RMSE and SAR values for all pulses described above were calculated using the full set of 16 B_1^+ maps

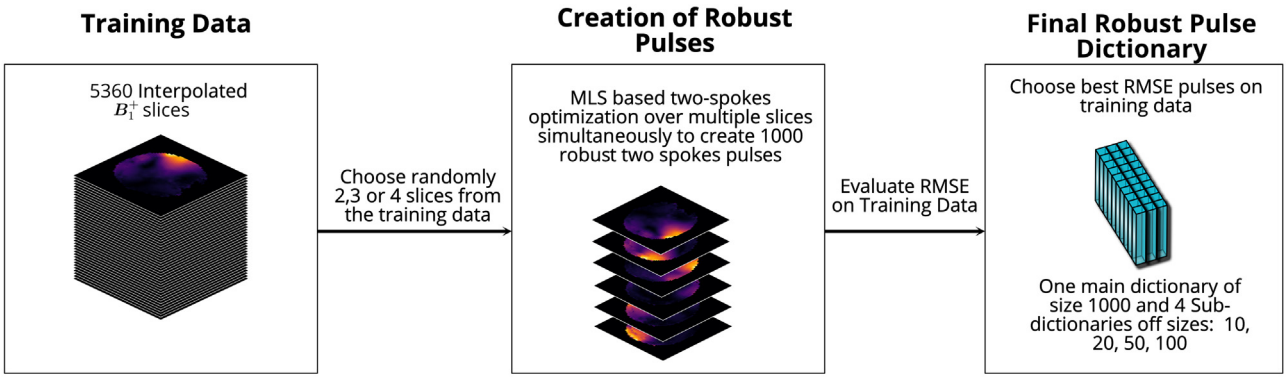

Figure 3. Creation of RPs: pulses were simultaneously optimized with MLS [8] two-spoke RF pulse designs [11] on multiple, randomly chosen slices from the training data using the method described in [51]. They were then evaluated with Bloch simulations to create a set of 1000 pulses. From this set, four additional pulse dictionaries were created by evaluating all pulses based on their RMSE performance on the training data, and the top 10, 20, 50 and 100 best performing pulses were chosen.

measured in vivo on a fifth subject, i.e. without sub-sampling. In this way, the pulses that were calculated with sub-sampled data are evaluated against the full dataset that represents the ground truth. The data of the fifth subject was not used in training any machine-learning model and was only used to evaluate the performance of the pulse design algorithms and B_1^+ map augmentation. Evaluation of RMSE and the SAR burden was carried out as described in [11].

All B_0 and B_1^+ maps were measured on the 9.4 T MAGNETOM system mentioned above using a 16-channel transceive coil array described in [48]. A dual refocusing acquisition mode (DREAM) sequence [40] with transmit channel Fourier phase encoding [41] was used to acquire B_1^+ maps. The sequence parameters were as reported in [11]: FoV = 256 mm × 224 mm, 17 slices, slice thickness = 4 mm, TR/TE1/TE2/TI = 6.8 ms/2.22 ms/4.44 ms/7.1 ms, flip angle = 7° and matrix 64 × 64. A slab selective double-echo gradient-echo sequence was used to acquire the static field maps (FoV = 200 mm × 200 mm, 44 slices, slice thickness = 4 mm, TR/TE1/TE2 = 30 ms/1 ms/3.21 ms, flip angle = 8° and matrix 50 × 50). In vivo data were acquired following approval of the study by the local ethics committee and after having obtained written informed consent.

3 Results

3.1 B1-map generation

Figure 4 visualizes all 16 measured in vivo B_1^+ maps with a normalized magnitude of one subject and compares these with versions generated from the SPADE and pix2pixHD models, respectively. One of the transmit channels is shown in greater detail in the bottom part of Figure 4, displaying both magnitude and phase. All generated images appear qualitatively close to the real measured ones; however, differences in detail can be observed. The quantitative differences between the generated images are evaluated in Table 1.

Table 1
Numerical RMSE values for several 2 spoke RF pulse designs on new subject not in the training data.

RF pulse design	RMSE ± std.(%)	Local SAR ± std. ($W\,kg^{-1}$)
RF optimization: different sub-sampling B_1^+ factors		
Real B1	4.65 ± 1.48	2.90 ± 0.47
pix2pixHD	2 5.01 ± 1.40	2.77 ± 0.71
pix2pixHD 4	8.52 ± 2.15	2.91 ± 0.90
pix2pixHD 8	9.30 ± 2.48	2.84 ± 1.22
SPADE 2	8.59 ± 3.35	3.08 ± 1.25
SPADE 4	10.40 ± 4.16	4.41 ± 2.38
SPADE 8	10.32 ± 3.61	4.22 ± 2.62
SPADE 16	10.87 ± 3.72	4.00 ± 1.96
SPADE 16 regr	13.19 ± 3.74	2.30 ± 1.06

3.2 RF pulse design

Figure 5 shows the RMSE results when performing a full RF optimization for different B_1^+ sub-sampling factors. The sub-sampling factor of measured B_1^+ transmit channel maps compared to the full set of maps is given below each bar, and it can be seen that the RMSE performance generally improves with lower sub-sampling factors. The pix2pixHD model with a sub-sampling factor of two performs almost as well as a full set of measured data. The RF prediction based on the Resnet18 model is also shown based on a sub-sampling factor of 16 and with a synthesis of the missing B_1^+ maps with SPADE. The RF prediction performs worse compared to an optimization on the same synthesized data (SPADE with sub-sampling factor 16), however, its execution is significantly faster. While the VAE compression of the weights does improve the pulse prediction of Resnet18 slightly compared to a non-compressed Resnet18 prediction without a VAE (compare Figure S5), it is expected that improvements in model architectures, objective functions and compression schemes will lead to further improvements.

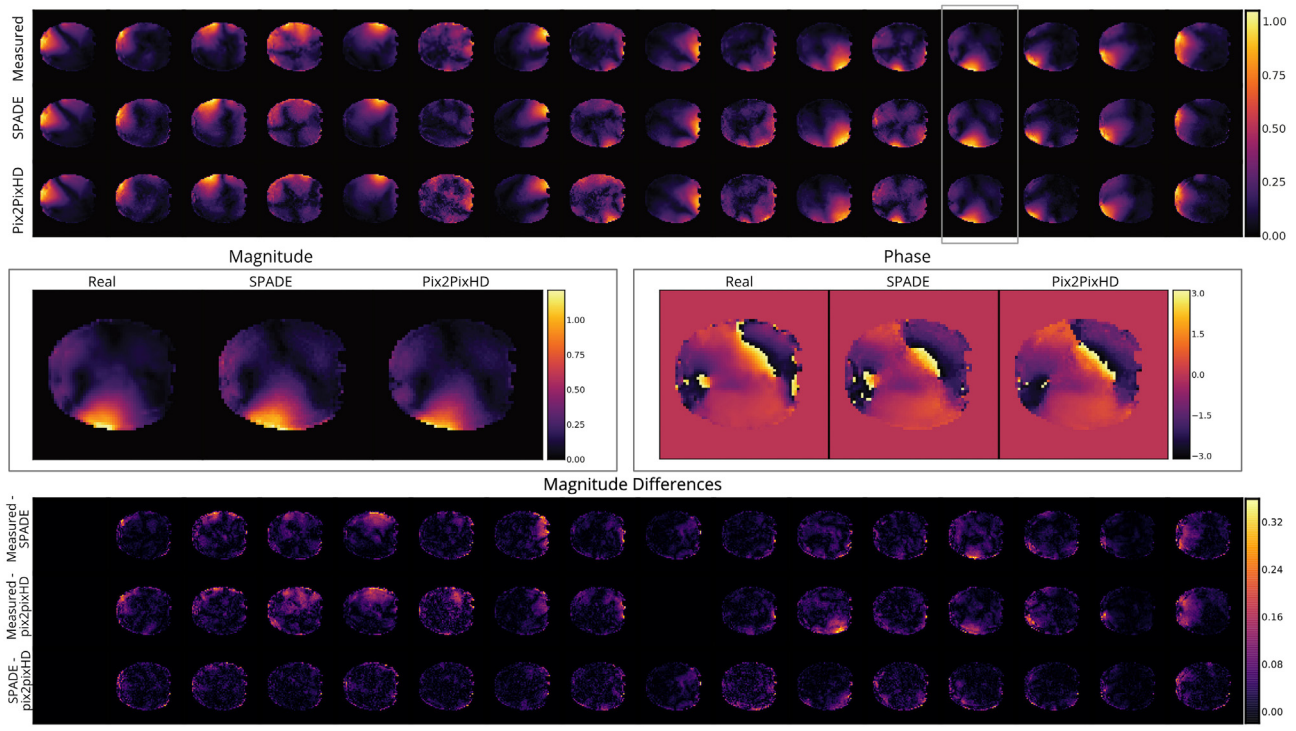

Figure 4. The top image shows the normalized magnitude-images of the measured, SPADE-generated and pix2pixHD-generated B_1^+ maps for all 16 transmit channels of an example slice, respectively. The transmit channel indicated by the box is shown in greater detail in the middle row, both in magnitude and phase. The difference in magnitude of the predicted images is given in the bottom row.

Figure 6 visualizes the flip angle patterns of the results shown in Figure 5 following Bloch simulation for a set of representative slices. Consistent with Figure 5, the excitation patterns of the pix2pixHD case with sub-sampling factors of two are very close to the case that employs the full dataset. With higher sub-sampling factors, the slice homogeneity decreases in terms of RMSE. One alternative to a pulse optimization is the evaluation of pulses in a dictionary, as shown in the bottom row, based on the SPADE generated data with a sub-sampling factor of 16. Compared to the optimization of RF pulses in the row "SPADE 16", the performance was slightly improved for some slices and slightly reduced for others. Table 1 summarizes the findings in numerical form.

Figure 7 shows the RMSE after evaluating all robust pulses contained in the five dictionaries of the respective sizes, 1000, 100, 50, 20 and 10, for different B_1^+ sub-sampling factors and with the minimum RMSE pulse from the dictionary following Bloch simulations. It can be seen that using a larger dictionary size tends to improve the resulting pulse performance compared to the smaller pulse dictionaries, especially for lower sub-sampling factors. With SPADE, no improvement with larger dictionary sizes can be observed for the higher sub-sampling factors of 8 and 16 as the greedy pulse choice on synthesized data does not necessarily translate to the best choice on the real data. As expected, the best results were achieved by utilizing all 16 real B_1^+ maps to choose the optimal pulse. However, comparable results can be achieved when utilizing a sub-sampling factor of two both for SPADE and pix2pixHD.

Due to the fact that the Bloch simulation of 1000 pulses takes more time compared to smaller dictionary sizes, there is a tradeoff between dictionary size and longer evaluation times. Consequently, for the "SPADE 16" case, even a dictionary size of 20 results in a pulse performance that is comparable to the larger dictionary sizes of 50, 100 or 1000 and might be preferred for the increased speed. Evaluation of pre-computed pulses is, in general, still faster than a subject-specific pulse optimization. Figure 7 demonstrates another use-case for which B_1^+ synthesis can be beneficial for an excitation scheme based on robust pulse dictionaries.

Figure S1 in the supplementary material shows a performance comparison for several MLS RF two spokes pulse design algorithms from the literature on the full set of measured B_1^+ maps without any sub-sampling. This represents the best possible excitation case since only real data is employed in the pulse design. Consistent with results presented in [11], the ES-based approach shows the best RMSE performance with results comparable with the grid-search and sequential quadratic programming (SQP)-based designs [11,52]. Using only pre-determined spoke-locations in the region-growing approach of [50] resulted in sub-optimal solutions. Searching for optimal latent vectors, z, of our trained VAE with an

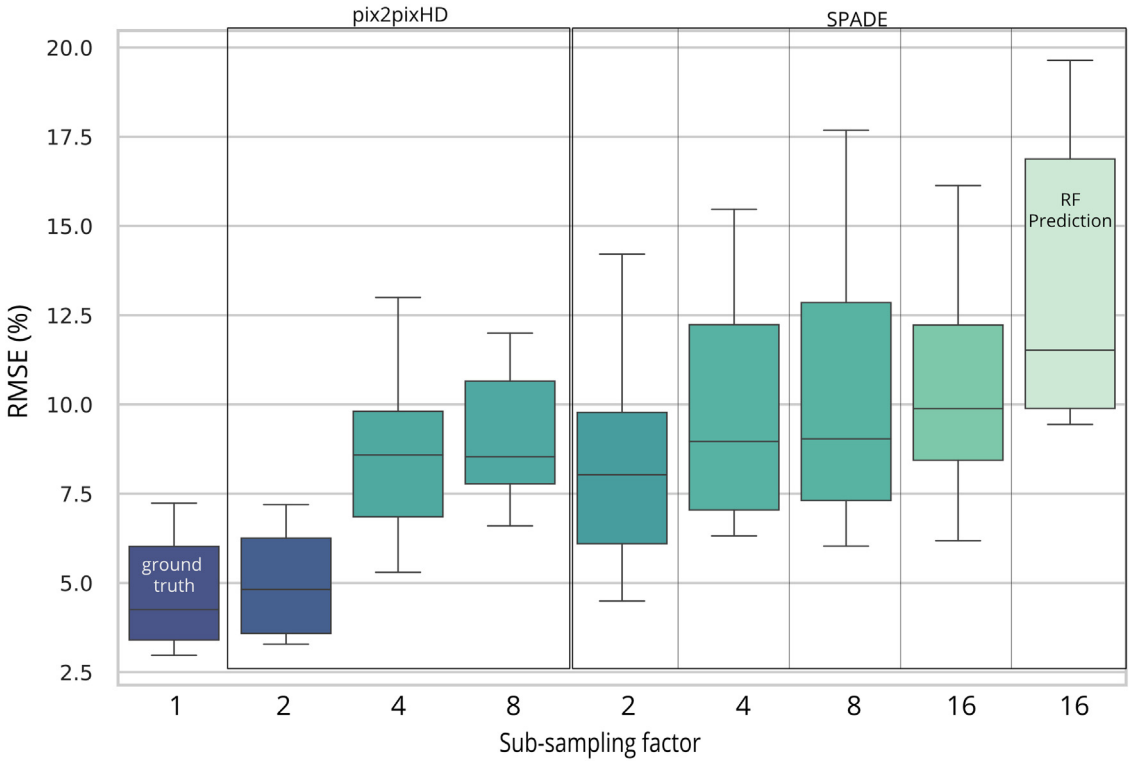

Figure 5. RF pulse RMSE performance after full optimization of RF pulses with different B_1^+ sub-sampling factors. The investigated machine learning RF algorithm (Figure 2) is shown on the right. Pulse optimizations were performed with an MLS and ES based pulse design [11] once all 16 B_1^+ were synthesized.

ES-based optimization [53] shows comparable results to the best ES -based optimization [11] in the common $\boldsymbol{w}, k_x, k_y$ optimization parameter space. Performing a grid-search as a means of initialization reduces the interquartile range of the box plots slightly.

Figure S2 shows the results of algorithms from the literature mentioned in the previous paragraph on all considered subsampling cases. It can be observed that the ES optimization performs best when the synthesized B_1^+ maps are close to the original data. However, this advantage vanishes for higher subsampling factors, and a grid-search can be a viable alternative as it is a fast and robust method [11]. It is worth optimizing the (k_x, k_y) location for all subsampling factors since, in most cases, a fixed, pre-determined location, as used in [50], is sub-optimal.

Figures S3 and S4 show the results of local SAR optimized RF pulses on sub-sampled B_1^+ data, as proposed in [11]. Instead of optimizing the RMSE, the local SAR becomes the main objective with the worst possible RMSE as an inequality constraint that must not be exceeded. By allowing a higher RMSE trade-off percentage, and thus potentially reducing the excitation homogeneity, the local SAR can be reduced, as shown in Figure S4, for all subsampling factors. While the RMSE degrades significantly with a large chosen RMSE trade-off percentage of 50%, choosing a small percentage such as 5% or 10% only leads to minor RMSE degradation (Figure S3).

4 Discussion

We demonstrated an intermediate method for RF-pulse generation in UHF-MRI as a combination of calibration free-methods like UP or "smart pulses" and methods employing a full set of per subject measured field maps. Using the described approach, it was shown that a sub-sampled set of RF-field maps can be artificially augmented to a full set using machine learning methods. While this is only one of numerous possible implementations, in a next step, a proof of concept could be carried out for a 16-channel transceiver coil used in 9.4 T neuroimaging applications. As expected, the results show that reducing the amount of calibration data to a sub-sampled set of RF field maps incurs a performance penalty against methods employing the full set of acquired data, Figure 5. However, depending on the sub-sampling factor, it can regain excitation fidelity and a SAR burden that is comparable to calibration-free methods (see Table 1 in the supplementary material).

The machine learning approaches employed were all trained on a limited set of field maps acquired from four healthy volunteers. In order to augment the limited training data, the

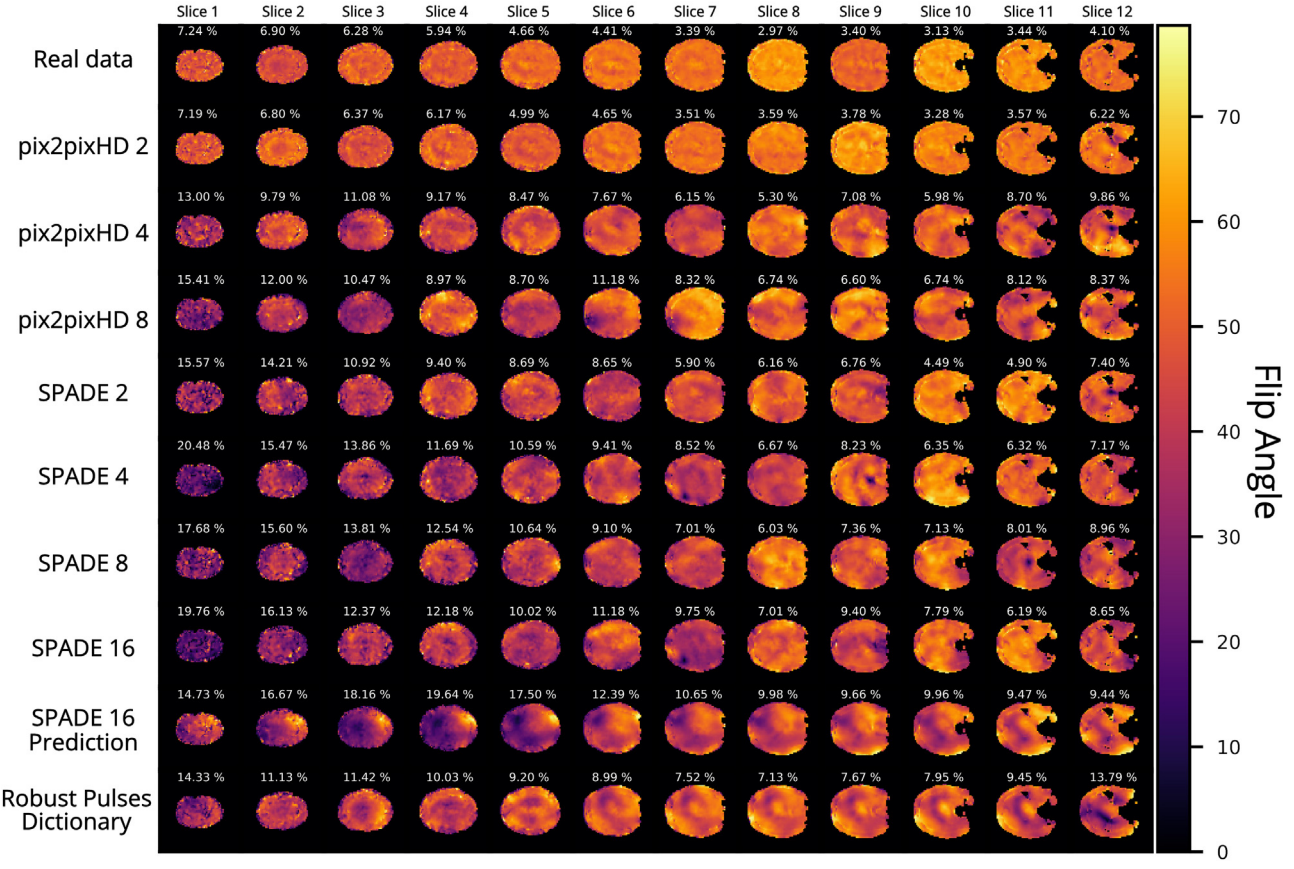

Figure 6. Bloch simulated flip-angle maps after RF pulse optimization based on either real B_1^+ data (first row) or partially sub-sampled B_1^+ data in the remaining rows. The RMSE percentage is given for each individual case. The sub-sampling factor is given at the beginning of each row next to the name of the generative machine learning model employed to synthesize the missing data. The second to last row shows the predicted excitation pattern of the employed machine learning regression model. The last row shows the excitation patterns after taking the lowest RMSE pulse from the dictionary of robust pulses of size 10 based on SPADE 16 generated B_1^+ data.

measured maps were linearly interpolated on a very fine grid. Despite only small variations in the generally smooth field maps, this approach yielded sufficient training data for the machine learning algorithms to converge during training – an observation concordant with [54], where data augmentation is described as a necessity in medical imaging as it is commonly plagued with a limited amount of training data. It should also be noted that training in the case presented here is specific for the RF coil employed and would need to be repeated for other coil systems. A possible future research direction is to investigate the robustness of the trained generative models with respect to changes in the measurement configuration such as the pTX coil array, magnetic field strength or B_1^+ mapping sequence. Whether a single model can be employed to robustly generate the missing B_1^+ maps for a cohort of coil arrays, e.g. configurations that use the same type of elements and the same geometrical arrangement of those, remains to be seen and requires further investigation. It might also be feasible to test translation in field strength, e.g. from 7 to 9.4 T, once the above prerequisites are fulfilled. An alternative would be to train a separate model for each measurement configuration. A pre-trained configuration-specific model could be supplied from the coil manufacturer or a once trained configuration-specific model could be easily loaded. An interesting extension of this work would be to apply the strategies investigated here to the most commonly used commercial pTX coils at 7 T, which would make them applicable to many 7 T sites, systems and even across MR vendors.

Concerning the synthesis of missing B_1^+ maps, as shown in Figure 4, we chose to investigate two paired image translation methods. While unpaired image translation is more common, the pairing gives additional control over the generated maps, which is desirable. Both methods investigated performed well, although slightly better results were obtained with the pix2pixHD method, which showed especially promising performance with a sub-sampling factor of two (compare Figures 4 and 5). However, the performance benefit of the pix2pixHD method compared to SPADE might also be a result of the limited GPU memory, as SPADE generally requires more memory for training purposes. Consequently, SPADE

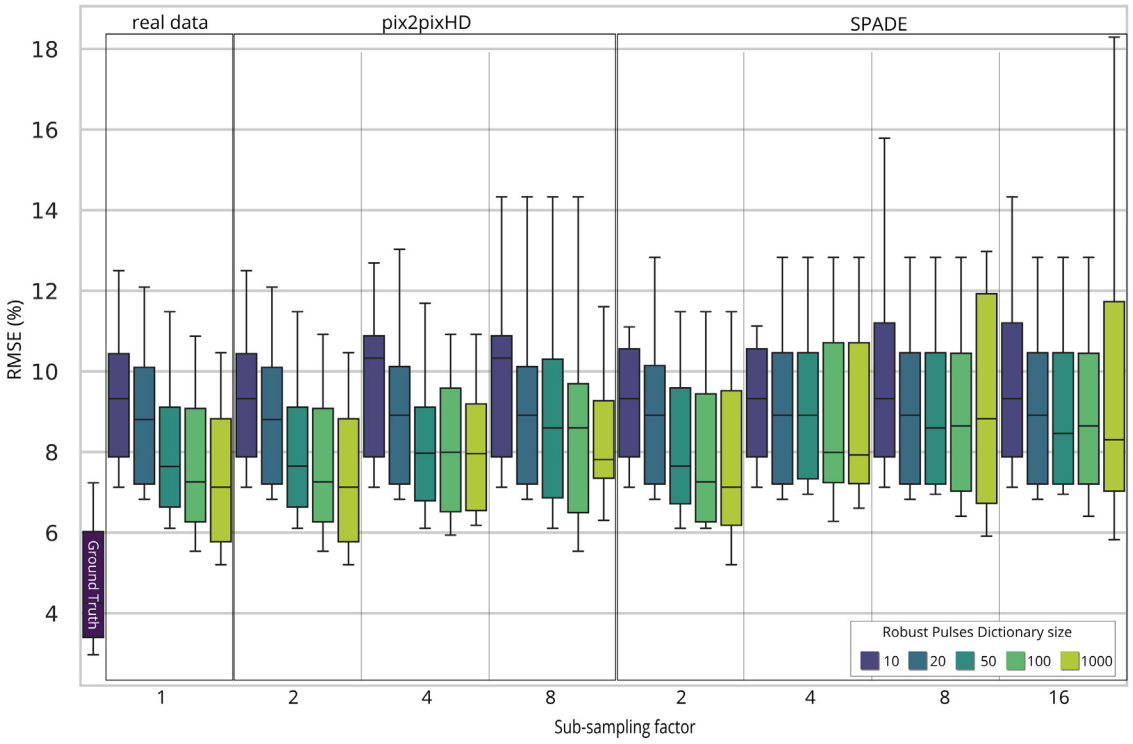

Figure 7. RMSE comparison of pre-computed robust pulses dictionaries. All pulses in the dictionary were evaluated, and the lowest RMSE pulse was chosen as the final pulse. Results for different sub-sampling factors for both image translation models, SPADE and pix2pixHD, are shown. In general, larger dictionary sizes of robust pulses perform better as there are more pulses available to choose from. The ground truth (real data with full RF optimization) is shown on the left.

might benefit from more computational resources [36]. However, due to the GPU memory limitation, we could not train with the full default model size.

Again, the investigated implementation did not cover all available paired image translation methods, used a very limited training data set, and was trained on a single GPU with limited memory resources. Therefore, it is possible that these factors impaired performance, and future improvements could be made in terms of model architectures and computing hardware. These could then be traded for a smaller number of calibration scans (faster measured preparation) using higher sub-sampling factors or to improve excitation fidelity. From a practical point, we found the SPADE method to be slightly simpler to implement as it is trained as a VAE [36] and allows the generation of all images in one go with a single seed image and the label maps that correspond to each transmit channel as visualized in Figure 1B. In contrast to SPADE, we performed image generation with pix2pixHD, which is conditioned on an input image alone. For this reason, missing images first need to be generated one at a time as input for the next channel images, slightly increasing generation time in total.

As the methods for augmenting the calibration data are already based on machine learning methods, a straightforward implementation might condense the steps of calibration data generation and RF spoke generation in a single approach. While we did not implement this method directly, we chose to add additional machine learning algorithms for pulse generation on top of the data augmentation. This was also done in order to assess the performance penalties of all steps individually. Pulse generation was performed using a Resnet18 regression model in combination with a trained VAE to reduce the dimensionality of the data. Again, the implementation only shows one feasible solution from numerous possible implementations.

As shown in Figure 5, the pulses predicted with the machine learning model were less optimal compared to the other investigated variants; however, in some cases, they showed a comparable performance to other considered algorithms, as seen in Figure 6. The near instant prediction time makes the approach interesting as an alternative to other pulse designs. Overall, the investigation of improved model architectures and different data encoding schemes is a promising research direction for the future.

In light of this, several alternative machine-learning RF pulse designs are conceivable. A classification of pre-computed pulses similar to that in [25] can be easily implemented with our employed Resnet18 model by simply changing the training objective, and a potential clustering of the target data could provide the class labels. Additionally, pulse weight generation with generative models such

as SPADE represents an interesting alternative as they do not utilize an MSE objective, which could lead to blurring of the generated data to reduce the objective loss. Several encoding schemes with VAE to aid the training process of any machine learning regression or classification model are also conceivable.

Using either pre-computed RF pulses or machine-learning predicted RF pulses without subject-dependent optimization may still result in pulse solutions that exhibit signal voids as these solutions could correspond to local minima in the optimization space of pulse parameters, as shown in Figure 6. However, it is expected that this problem can be addressed using published pulse design algorithms, e.g. [10,55,56,52,11], that specifically avoid signal voids. These methods could either be run in an additional design step starting with the pre-computed pulses from the trained algorithms presented here, or, to avoid solutions with signal voids from the beginning, might be incorporated into the training.

Finally, the combination of the sub-sampling factor and size of dictionaries must be observed. In general, a higher sub-sampling factor allows a decrease in pulse dictionary size without significant performance penalties, as shown in Figure 7. Certainly, an in-depth investigation of this interplay could give a better indication of the optimal topology for the algorithms employed.

5 Conclusion

By substituting missing data with synthesized data generated by machine learning models, it is possible to save time on the initial calibration and to either optimize RF pulses or make improved choices when pre-calculated pulses are used.

Conflict of interest

None declared.

Acknowledgements

The authors thank D.H.Y. Tse for providing several scripts used in this work. They acknowledge C. Rick for proofreading the manuscript.

Appendix A Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.zemedi.2021.12.003.

References

[1] Deniz CM. Parallel transmission for ultrahigh field MRI. Top Magn Reson Imaging 2019;28(3):159–71.
[2] Van de Moortele PF, Akgun C, Adriany G, Moeller S, Ritter J, Collins CM, et al. B1 destructive interferences and spatial phase patterns at 7 T with a head transceiver array coil. Magn Reson Med 2005;54(6):1503–18.
[3] Vaughan T, DelaBarre L, Snyder C, Tian J, Akgun C, Shrivastava D, et al. 9.4T human MRI: preliminary results. Magn Reson Med 2006;56(6):1274–82.
[4] Katscher U, Bornert P, Leussler C, van den Brink JS. Transmit SENSE. Magn Reson Med 2003;49(1):144–50.
[5] Zhu Y. Parallel excitation with an array of transmit coils. Magn Reson Med 2004;51(4):775–84.
[6] Saekho S, Yip CY, Noll DC, Boada FE, Stenger VA. Fast-kz three-dimensional tailored radiofrequency pulse for reduced B1 inhomogeneity. Magn Reson Med 2006;55(4):719–24.
[7] Grissom WA, Setsompop K, Hurley SA, Tsao J, Velikina JV, Samsonov AA. Advancing RF pulse design using an open-competition format: report from the 2015 ISMRM challenge. Magn Reson Med 2017;78(4):1352–61.
[8] Setsompop K, Wald LL, Alagappan V, Gagoski BA, Adalsteinsson E. Magnitude least squares optimization for parallel radio frequency excitation design demonstrated at 7 Tesla with eight channels. Magn Reson Med 2008;59(4):908–15.
[9] Paez A, Gu C, Cao Z. Robust RF shimming and small-tip-angle multispoke pulse design with finite-difference regularization. Magn Reson Med 2021.
[10] Dupas L, Massire A, Amadon A, Vignaud A, Boulant N. Two-spoke placement optimization under explicit specific absorption rate and power constraints in parallel transmission at ultra-high field. J Magn Reson 2015;255:59–67.
[11] Eberhardt B, Poser BA, Shah NJ, Felder J. Application of evolution strategies to the design of SAR efficient parallel transmit multispoke pulses for ultra-high field MRI. IEEE Trans Med Imaging 2020;39(12):4225–36.
[12] Guerin B, Gebhardt M, Cauley S, Adalsteinsson E, Wald LL. Local specific absorption rate (SAR), global SAR, transmitter power, and excitation accuracy trade-offs in low flip-angle parallel transmit pulse design. Magn Reson Med 2014;71(4):1446–57.
[13] Gras V, Vignaud A, Amadon A, Mauconduit F, Le Bihan D, Boulant N. In vivo demonstration of whole-brain multislice multispoke parallel transmit radiofrequency pulse design in the small and large flip angle regimes at 7 Tesla. Magn Reson Med 2017;78(3):1009–19.
[14] Gras V, Vignaud A, Amadon A, Le Bihan D, Boulant N. Universal pulses: a new concept for calibration-free parallel transmission. Magn Reson Med 2017;77(2):635–43.
[15] Gras V, Mauconduit F, Vignaud A, Amadon A, Le Bihan D, Stöcker T, et al. Design of universal parallel-transmit refocusing kT-point pulses and application to 3D T2-weighted imaging at 7T. Magn Reson Med 2018;80(1):53–65.
[16] Herrler J, Liebig P, Gumbrecht R, Ritter D, Schmitter S, Maier A, et al. Fast online-customized (FOCUS) parallel transmission pulses: a combination of universal pulses and individual optimization. Magn Reson Med 2021;85(6):3140–53.
[17] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–44.
[18] Ng AY, Jordan MI. On discriminative vs. generative classifiers: a comparison of logistic regression and Naive Bayes. In: Dietterich TG, Becker S, Ghahramani Z, editors. Advances in neural information processing systems 14. MIT Press; 2002. p. 841–8.
[19] Lathuilière S, Mesejo P, Alameda-Pineda X, Horaud R. A comprehensive analysis of deep regression. IEEE Trans Pattern Anal Mach Intell 2020;42(9):2065–81.
[20] Kingma DP, Welling M. An introduction to variational autoencoders. Found Trends® Mach Learn 2019;12(4):307–92.
[21] Dupont E, The YW, Doucet A. Generative models as distributions of functions; 2021 arXiv:2102.04776 [cs, stat].

[22] Kingma DP, Welling M. Auto-Encoding variational Bayes. In: 2nd international conference on learning representations, ICLR 2014. 2014.
[23] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks; 2014 arXiv:1406.2661 [cs, stat].
[24] Ianni JD, Cao Z, Grissom WA. Machine learning RF shimming: prediction by iteratively projected ridge regression. Magn Reson Med 2018;80(5):1871–81.
[25] Cloos MA, Boulant N, Luong M, Ferrand G, Giacomini E, Le Bihan D, et al. kT-points: short three-dimensional tailored RF pulses for flip-angle homogenization over an extended volume. Magn Reson Med 2012;67(1):72–80.
[26] Tomi-Tricot R, Gras V, Thirion B, Mauconduit F, Boulant N, Cherkaoui H, et al. SmartPulse, a machine learning approach for calibration-free dynamic RF shimming: preliminary study in a clinical environment. Magn Reson Med 2019.
[27] Vinding MS, Aigner CS, Schmitter S, Lund TE. DeepControl: 2DRF pulses facilitating inhomogeneity and B0 off-resonance compensation in vivo at 7 T. Magn Reson Med 2021;85(6):3308–17.
[28] Tavaf N, Torfi A, Ugurbil K, Van de Moortele P-F. GRAPPA-GANs for parallel mri reconstruction; 2021 arXiv:2101.03135.
[29] Zbontar J, Knoll F, Sriram A, Murrell T, Huang Z, Muckley MJ, et al. fastMRI: an open dataset and benchmarks for accelerated MRI; 2019.
[30] Lan H, Toga AW, Sepehrband F. Three-dimensional self-attention conditional GAN with spectral normalization for multimodal neuroimaging synthesis. Magn Reson Med 2021;86(3):1718–33.
[31] Meliadò EF, Raaijmakers AJE, Sbrizzi A, Steensma BR, Maspero M, Savenije MHF, et al. A deep learning method for image-based subject-specific local SAR assessment. Magn Reson Med 2020;83(2):695–711.
[32] Lei K, Mardani M, Pauly JM, Vasanawala SS. Wasserstein GANs for MR imaging: from paired to unpaired training. IEEE Trans Med Imaging 2021;40(1):105–15.
[33] Abbasi-Rad S, O'Brien K, Kelly S, Vegh V, Rodell A, Tesiram Y, et al. Improving FLAIR SAR efficiency at 7T by adaptive tailoring of adiabatic pulse power through deep learning estimation. Magn Reson Med 2021;85(5):2462–76.
[34] Wu X, Schmitter S, Auerbach EJ, Uğurbil K, Van de Moortele P-F. A generalized slab-wise framework for parallel transmit multiband RF pulse design. Magn Reson Med 2016;75(4):1444–56.
[35] Wang T-C, Liu M-Y, Zhu J-Y, Tao A, Kautz J, Catanzaro B. High-resolution image synthesis and semantic manipulation with conditional GANs. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
[36] Park T, Liu M-Y, Wang T-C, Zhu J-Y. Semantic image synthesis with spatially-adaptive normalization. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2019.
[37] Kassakian PW. Convex approximation and optimization with applications in magnitude filter design and radiation pattern synthesis [PhD thesis]. In: Engineering-electrical engineering and computer sciences. Berkeley: University of California; 2006.
[38] Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks; 2018 arXiv:1611.07004 [cs].
[39] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation; 2015 arXiv:1505.04597 [cs].
[40] Nehrke K, Bornert P. DREAM – a novel approach for robust, ultrafast, multislice B(1) mapping. Magn Reson Med 2012;68(5):1517–26.
[41] Tse DH, Poole MS, Magill AW, Felder J, Brenner D, Jon Shah N. Encoding methods for B1(+) mapping in parallel transmit systems at ultra high field. J Magn Reson 2014;245:125–32.
[42] Kanayamay S, Kuhara S, Satoh K. In vivo rapid magnetic field measurement and shimming using single scan differential phase mapping. Magn Reson Med 1996;36(4):637–42.
[43] NVlabs. Imaginaire: NVIDIA PyTorch GAN library with distributed and mixed precision support, GitHub repository; 2020.
[44] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR). 2016. p. 770–8.
[45] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: an imperative style, high-performance deep learning library. In: Wallach H, Larochelle H, Beygelzimer A, d'Alché-Buc F, Fox E, Garnett R, editors. Advances in neural information processing systems 32. Curran Associates, Inc.; 2019. p. 8024–35.
[46] Zhao S, Song J, Ermon S. InfoVAE: balancing learning and inference in variational autoencoders. In: Proceedings of the AAAI conference on artificial intelligence, vol. 33, no. 1. 2019. p. 5885–92.
[47] Grissom W, Yip CY, Zhang Z, Stenger VA, Fessler JA, Noll DC. Spatial domain method for the design of RF pulses in multicoil parallel excitation. Magn Reson Med 2006;56(3):620–9.
[48] Shajan G, Kozlov M, Hoffmann J, Turner R, Scheffler K, Pohmann R. A 16-channel dual-row transmit array in combination with a 31-element receive array for human brain imaging at 9.4 T. Magn Reson Med 2014;71(2):870–9.
[49] Tse DHY, Wiggins CJ, Ivanov D, Brenner D, Hoffmann J, Mirkes C, et al. Volumetric imaging with homogenised excitation and static field at 9.4 T. Magn Reson Mater Phys Biol Med 2016;29(3):333–45.
[50] Tse DHY, Wiggins CJ, Poser BA. High-resolution gradient-recalled echo imaging at 9.4T using 16-channel parallel transmit simultaneous multislice spokes excitations with slice-by-slice flip angle homogenization. Magn Reson Med 2017;78(3):1050–8.
[51] Schmitter S, Wu X, Uğurbil K, de Moortele P-FV. Design of parallel transmission radiofrequency pulses robust against respiration in cardiac MRI at 7 Tesla. Magn Reson Med 2015;74(5):1291–305.
[52] Hoyos-Idrobo A, Weiss P, Massire A, Amadon A, Boulant N. On variant strategies to solve the magnitude least squares optimization problem in parallel transmission pulse design and under strict SAR and power constraints. IEEE Trans Med Imaging 2014;33(3):739–48.
[53] Hansen N, Arnold DV, Auger A. Evolution strategies. In: Handbook of computational intelligence. 2015.
[54] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J. Big Data 2019;6(1):60.
[55] Vinding MS, Guérin B, Vosegaard T, Nielsen NC. Local SAR, global SAR, and power-constrained large-flip-angle pulses with optimal control and virtual observation points. Magn Reson Med 2017;77(1):374–84.
[56] Yoon D, Fessler JA, Gilbert AC, Noll DC. Fast joint design method for parallel excitation radiofrequency pulse and gradient waveforms considering off-resonance. Magn Reson Med 2012;68(1):278–85.