

DeepSPIO: Super Paramagnetic Iron Oxide Particle Quantification using Deep Learning in Magnetic Resonance Imaging

Gabriel della Maggiora^{*†¶}, Carlos Castillo-Passi^{*†‡¶}, Wenqi Qiu[§], Shuang Liu[§], Carlos Milovic^{*†¶}, Masaki Sekino[§], Cristian Tejos^{*†¶}, Sergio Uribe^{*†‡¶} and Pablo Irarrazaval^{*†‡¶}

Email: {gedellamaggiora, cncastillo, pim}@uc.cl

^{*}Department of Electrical Engineering, Pontificia Universidad Católica de Chile, Santiago, Chile

[†]Biomedical Imaging Center, Pontificia Universidad Católica de Chile, Santiago, Chile [‡]Institute for

Biological and Medical Engineering, Pontificia Universidad Católica de Chile, Santiago, Chile [§]Department of Bioengineering, School of Engineering, University of Tokyo, Tokyo, Japan, [¶]Millennium Nucleus for Cardiovascular Magnetic Resonance, Santiago, Chile

Abstract—The susceptibility of Super Paramagnetic Iron Oxide (SPIO) particles makes them a useful contrast agent for different purposes in MRI. These particles are typically quantified with relaxometry or by measuring the inhomogeneities they produced. These methods rely on the phase, which is unreliable for high concentrations. We present in this study a novel Deep Learning method to quantify the SPIO concentration distribution. We acquired the data with a new sequence called View Line in which the field map information is encoded in the geometry of the image. The novelty of our network is that it uses residual blocks as the bottleneck and multiple decoders to improve the gradient flow in the network. Each decoder predicts a different part of the wavelet decomposition of the concentration map. This decomposition improves the estimation of the concentration, and also it accelerates the convergence of the model. We tested our SPIO concentration reconstruction technique with simulated images and data from actual scans from phantoms. The simulations were done using images from the IXI dataset, and the phantoms consisted of plastic cylinders containing agar with SPIO particles at different concentrations. In both experiments, the model was able to quantify the distribution accurately.

Index Terms—Machine Learning, Deep Learning, Neural Networks, MRI, quantification, susceptibility, QSM, SPIO.



1 INTRODUCTION

Superparamagnetic iron oxide (SPIO) nanoparticles produce considerable changes in the main magnetic field of a Magnetic Resonance Imaging (MRI) acquisition. These field distortions change the signal’s phase, which in turn manifests as geometric distortion and signal attenuation. In general, the signal dropout is an artifact to be avoided, but it also can be used as a contrast agent [1].

In particular, SPIO particles have been used to know if cancer cells have metastasized from sentinel lymph nodes, after tumor removal, by using their concentration as a biomarker. Furthermore, SPIO particles can be used in liver imaging, MR angiography, and tracking of labeled cells in vivo. Quantification of SPIO density or total quantity is important for clinicians since it could provide information for diagnosis and evaluation of treatment [1], [2]. The distribution of particles is commonly measured by using two types of methods: relaxometry-based or model-based.

The relaxometry methods use the magnitude of the MR image and quantify the SPIO particles within an area of interest by its relaxation rate. The effective transverse relaxation R_2^* is assumed to vary linearly with the SPIO concentration: $R_2^* = R_{20}^* + r_2^*c$ [3]. Where R_{20}^* is the intrinsic relaxation rate, r_2^* denotes the relaxation rate of the contrast agent, and c is the concentration of the particles.

The model-based methods solve a – somewhat difficult – inverse problem. This problem consists of finding the particle’s concentration, or susceptibility, from the measured signal. Solving this problem requires first finding the field map from the signal’s phase and then solving the corresponding inverse model. The equations for the direct model are well known and are typically based on spherical dipoles for each particle. [1], [4].

Both types of methods cannot be applied in cases where the concentration of SPIO is high since the signal is significantly attenuated or entirely lost. The loss of signal is particularly relevant for model-based methods since they rely on the phase, which becomes extremely unreliable.

In this work, we propose to use Deep Learning to quantify the SPIO concentration from an MRI acquisition. As will be described later, the MRI sequence has the acquisition plane perpendicular to the excitation, such that the excitations’ profile and location are easily seen.

For the reconstruction, we only use the geometric distortions, making the phase irrelevant. The idea is to use the shape of these profiles and their centerlines’ curvature to infer the SPIO distribution. The model training is done by solving the forward model in a simulator that generates synthetic images. Previous studies in quantitative susceptibility mapping (QSM) [5], [6] have successfully estimated the susceptibility using Deep Learning with an encoder-

decoder architecture. QSM Deep Learning models are based mainly on U-Net [7] and SegNet [8].

These networks were designed for image segmentation [7], [8]. They transfer features from the encoder to the decoder to enforce a similarity with the output. Furthermore, these architectures generate a latent representation of the problem to reconstruct the desired output. These models have been successfully used in QSM and several other image-to-image tasks such as MRI image reconstruction [9], image denoising [10], etc.

Our approach is different than previous Deep Learning methods for two reasons. Firstly, because in our network, the input and output are visually distinct and, therefore, not directly related, as opposed to segmentation problems. Secondly, our approach is based on geometric distortions. Thus, it is less sensitive to voids in the signal.

Our encoder network is topologically similar to that of U-Net. Nevertheless, we replaced the pooling layers by convolutional layers, and we added four residual blocks at the end of the encoder. The latter extends the representation capabilities of the model. The essential component of DeepSPIO is its parallel decoder network. Each decoder uses learnable up-sampling operations through layers of transposed convolutions. The reason for using a parallel decoder is that at the output, we work in the wavelet domain with one decoder for each of the discrete wavelet channels. This output coding improves the gradient flow through the network since the wavelet channels are relatively independent of each other, and so are their gradients at each decoder. Together with the residual bottleneck, these changes allow for a much better gradient flow to the encoder, thus improving the network convergence rate.

The two main contributions of this work are a new technique for quantifying SPIO concentrations suited for high concentrations, and a new network architecture that reconstructs well these concentrations in a situation where a U-Net cannot converge.

In the following sections, we describe the main architectures previously used for this kind of image processing; we describe the MR acquisition and the simulations; we describe the architecture employed in DeepSPIO. Finally, we discuss the results and potential applications of this technique.

2 RELATED WORK

The concentration of SPIO is directly related to its susceptibility and the changes it produces in the magnetic field. Thus, it is natural to look into existing techniques that measure susceptibility from field distortions. Susceptibility estimation is an active topic of research with challenging problems. Usually, the susceptibility is estimated from phase variations introduced by the magnetic field deviations and their corresponding off-resonance effects. The image's magnitude also suffers distortions, but they are frequently neglected, which is valid for weak off-resonance conditions. That is not the case for strong susceptibility conditions, which is the problem addressed in this work.

The field map distortion introduced by a single point can be modeled as a magnetic dipole [11]. A convolution

with this dipole kernel estimates the effect of more complex susceptibility distributions.

Most QSM methods estimate the susceptibility source distribution by a spatial deconvolution of the field map with the dipole kernel. And the field map is typically measured from two or more phase images acquired with gradient-echo sequences.

These techniques present three crucial challenges. The first one is unwrapping the phase data, the second is removing the external magnetic field contributions (or background), and the third is solving an ill-posed deconvolution. [12], [13]. Unwrapping is usually performed by best path algorithms or Laplacian-based operators. External contributions to the magnetization field may include imperfect shimming, magnetic susceptibility sources outside the region of interest, etc. [13]. Background magnetization fields are usually decoupled from local magnetization fields by exploiting properties of harmonic functions such as orthogonality, the spherical mean value, or solving the Poisson equation.

The main difficulty for performing the local field's deconvolution is that the Fourier transform of the dipole kernel contains a double-cone zero-valued surface. If these zeros are not taken into consideration, they produce streaking artifacts. Several solutions have been proposed, such as truncating the dipole in frequency space [12] or enforcing priors [14]–[17].

These methods cannot be applied directly to quantify SPIO concentrations. They rely on the phase of the signal, which is either lost or very unreliable due to the strong distortion in the field. Furthermore, most unwrapping algorithms are unable to handle larger than 2π jumps and rely on smooth data.

Another approach to handle strong susceptibility effects in the surrounding tissue is to design sequences that are less sensitive to metal objects. In this case, the objective is to remove the artifacts when imaging close to metallic implants. Several techniques have been proposed, such as SEMAC [18], MAVRIC [19] and their modifications. They provide a robust encoding against metal-induced field inhomogeneities within a feasible scan time. They are spin-echo (SE) sequences that minimize the intra-voxel dephasing due to field variations based on view-angle-tilting (VAT) compensation gradients to suppress in-plane distortions. The VAT gradient is applied simultaneously with the readout gradient, equal to the one used during the excitation, effectively tilting the readout direction and limiting the in-plane distortion to less than a pixel [20].

Inspired by SEMAC, we developed the View Line sequence [21] to encode the SPIO induced distortions in the image. Changes in the field map are reflected as displacements of the voxels in the z direction. We use the acquired image to estimate the SPIO distribution by quantifying these displacements. This encoding is a robust manner to see the effects of the field distortions in the surrounding tissue when the susceptibility is so strong that most of the signal is lost in the immediate area of the particles.

All these approaches use analytical formulations to find the susceptibility. But even if they had the perfect field map in the surroundings of the particles, they cannot work correctly because of the missing information close to the particles and also because they cannot represent the com-

plete complexity of the problem. Motivated by the success in several research fields, in our approach, we use Deep Learning to solve the SPIO quantification problem in an end-to-end manner.

The encoder-decoder model has an expressive power that increases exponentially with depth. This expressiveness is originated from the combinatorial nature of the rectified linear activation unit (ReLU) used for the encoding and decoding process and the adaptive basis that the model learns as well [22]. The main two encoder-decoder networks used in the medical imaging area [23] are U-Net [7] and its variation, SegNet [8].

These architectures have been used with several degrees of success in different areas such as denoising, deconvolution, image reconstruction, super-resolution, and in particular, QSM. As mentioned before, the techniques used in QSM are not directly applicable to SPIO quantification, but it is worth mentioning two neural network-based approaches: DeepQSM [5] and QSMnet [6].

DeepQSM uses the encoder-decoder U-Net architecture. It shows results that are comparable to the best analytical models. It was trained by computing the well-posed QSM forward solution of synthetic data consisting of simple shapes.

QSMnet is also a U-Net that performs the dipole deconvolution directly. This model was trained by computing the gold standard COSMOS (Calculation Of Susceptibility through Multiple Orientation Sampling) solution on an image and then using this as a reference for a single orientation input. QSMnet accurately obtains the COSMOS solution for a single orientation image [24]. Our work is inspired by the success achieved by previous QSM networks. In those models, the architecture is an encoder-decoder network that learns hierarchical features and a smooth representation between two manifolds. [22]. The encoder consists of convolutions with a preset number of filters, ReLU as activation function, and pooling layers for sub-sampling. DeepSPIO differs from U-Net and SegNet in that it uses strided convolutions. [7], [8].

Since our image encodes the information in its spatial configuration, it is imperative to conserve the position of the features. The local invariance of max-pooling would hinder the performance of our model. U-net also concatenates each convolution block of the encoder to the respective mirror decoder block. These skip connections can transfer some features from the encoding to the decoder making up for lost details in the image. Nevertheless, this property is not useful to us since visually, our input is entirely different from our output, and these features hinder the performance, so we removed these skip connections to prioritize the construction of hierarchical features.

3 METHODS

The main idea of this work is to use the geometric distortions induced by the paramagnetic particles' strong susceptibility. The hypothesis is that the concentration of the particles can be estimated from artifacts farther away, even where the signal is lost. For this, we use the View Line sequence, which images the excitation's profile [21]. For training, we simulate the effects of the SPIO in the data acquired with that sequence.

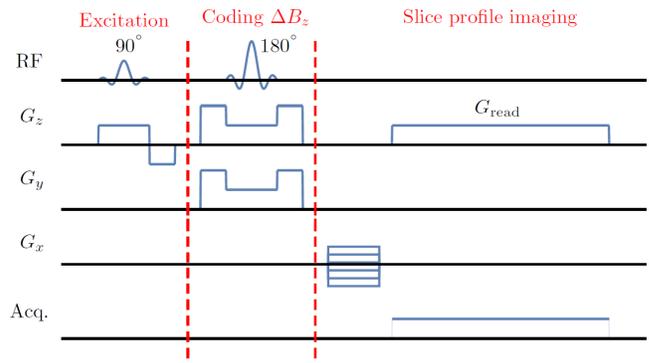


Fig. 1: View Line sequence for exciting spins in a line along the x -dimension.

3.1 Forward Model

In this section, we describe the View Line sequence shown in Figure 1. It consists of three parts: excitation, selective refocusing, and readout.

The excitation is a standard selective excitation pulse that excites a slice in z , centered at z_0 . Since we expect strong field deviations, the slice will not be planar, and its center, z'_0 will deviate from z_0 as a function of (x, y) : $z'_0(x, y)$.

The center of the excitation satisfies

$$z_0 G_z = z'_0 G_z + \Delta B_z(x, y, z'_0).$$

The left side is the intended excitation field (position), and the right side is the actual field. To avoid superposition of the signal due to different deviations at different (x, y) positions, a 180° selective refocusing pulse is used, such that spatial deviations only in the direction of phase encoding are refocused. For that, the refocusing pulse is applied in a diagonal

$$z_0 G_z + y_0 G_y = z'_0 G_z + y'_0 G_y + \Delta B_z(x, y'_0, z'_0)$$

which is satisfied for $y'_0 = y_0$. After the excitation and refocusing, the center of the excited column is at a constant $y = y_0$ and at varying z .

Finally, the readout consists of a standard imaging stage with the phase encoding in the x direction and the readout in the z direction. Our simulation considered the voxel warping effects due to the non-linear gradients introduced by the SPIO particles.

The sequence is repeated for many lines such that a collection of contiguous profiles is collected. These lines will be fed to the network in two groups, even and odd, to avoid overlapping of the profiles, and to ensure that the network will see the edges.

In our simulations, we used convolution of the susceptibility distribution with the dipole kernel to obtain the field map. The convolution in cartesian coordinates is:

$$\Delta B_z(\mathbf{x}) \propto \mathcal{F}^{-1} \left\{ \frac{1}{3} - \left(\frac{k_z}{\|\mathbf{k}\|} \right)^2 \right\} * C(\mathbf{x}). \quad (1)$$

Where $C(\mathbf{x})$ corresponds to the concentration distribution of SPIO particles.

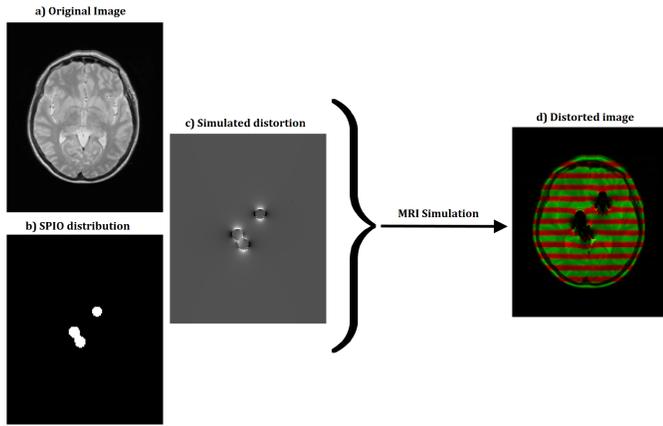


Fig. 2: Example of simulated image. (a) is distorted by (c) which is produced by (b). In the distorted image (d) the red lines correspond to the even simulations and the green lines to the odd simulations

3.2 Simulated dataset

We used the previously described convolution model to simulate the field variation as a function of the particle’s concentration. To build the dataset used for training we defined two geometric shapes: metaballs and toruses. Metaballs [25] are organic-looking three dimensional objects defined as follows:

$$M(x) = \mathbb{1}_{\sum_{i=1}^N m_i(x) \geq l}(x), \text{ with } m_i(x) = \frac{1}{\|x - x_0^i\|^2}, \quad (2)$$

where x_0^i is the center, and l is a threshold to determine its size.

The torus, centered at x_0 , radius r and revolution radius R , was defined as

$$T(x) = \mathbb{1}_{(\sqrt{(x-x_0)^2+(y-y_0)^2}-R)^2+(z-z_0)^2 \leq r^2}(x). \quad (3)$$

We generated multiple objects and assigned a uniform concentration of c to each one, and the total concentration $C(x)$ is the superposition of all of them. This concentration was convolved with the three-dimensional dipole kernel to obtain the desired field deviations $\Delta B_z(x)$.

To obtain the final image, we choose the central slice of ΔB_z and simulate multiple excitations perpendicular to that slice. In this way, we get multiple strips that encode the magnetic field distortions in the z direction. We arrange the acquired lines in a two-channel image, such that even lines are the first channel and odd lines the second, as shown in Fig. 2.

The input-output training sample is the pair formed by the image (made by the even and odd lines) and the one level wavelet coefficients (Daubechies 4) of the concentration $C(x)$. To build the training set we sampled c linearly, with 800,000 values between 0.22 and 17.5 $\mu\text{g}/\text{mm}^3$ using the previously described shapes. For each c value, we generated an SPIO distribution mixing the geometric shapes from 1 to 5 figures chosen randomly. We randomized l between 0.01 and 0.05 (Eq. 2), r between 2 and 4 and R between 5 and 10 (Eq. 3). We used the proton density

images provided in the IXI¹ dataset as the image to distort and generated 4 million samples.

4 ARCHITECTURE

DeepSPIO is composed of an encoder, a bottleneck in the middle, and multiple decoders emerging from this bottleneck, as shown in Figure 3.

The encoder is formed by 12 convolutional layers, similar to the first half of a U-Net [7], where convolutions replaced the pooling layers. The reason for this substitution is that the U-Net was designed for object segmentation, where spatial invariance, both locally and globally, is an important characteristic. The pooling steps provide local invariance, which is not desired in our case. We need to retain the local spatial dependency because the information is precisely encoded in space, and the voxel displacements can occur within the size of a filter. Thus, the model needs to be capable of summarizing spatial information correctly. The layers of strided convolutions serve this purpose. We used parameter sharing because the shape of the distortion is the same regardless of the position.

The bottleneck consists of 4 residual blocks. We added these blocks to increase the receptive field of the network and to improve the convergence of the model.

Finally, after the bottleneck, we have four decoders that generate the four channels of the wavelet transform of the SPIO concentration. The architecture of the decoders is similar to the one in the U-Net. In other words, our network can be thought of as four networks sharing the same encoder and bottleneck. Since the inputs are not geometrically related to the output, we decided against keeping the concatenations of U-Net.

4.1 The Encoder

Each layer of the encoder network performs convolutions with a group of filters to produce a set of feature maps. These are then batch normalized [26]. Then an element-wise ReLU ($f(x) = \max(0, x)$), is applied. Following that, we apply a 2D convolution with a stride of 2 with linear activation. This layer downsamples the output by a factor of 2. We used convolutions instead of pooling layers to retain important spatial information. After the downsampling, we increase the number of filters in the next convolution block by 2, in the same fashion as U-Net.

4.2 The Bottleneck

We built the bottleneck concatenating four residual blocks. Each block is the superposition of two paths. One is a 2D convolution followed by batch normalization and a ReLU activation. The second path is a shortcut with a linear activation. Mathematically, each residual block i with input x and output y is

$$y = F(x, \{\theta_i\}) + \theta_s x,$$

where $F(\cdot)$ represents the convolution/normalization/activation process with parameters θ_i , and the parameters of the shortcut are θ_s .

1. <https://brain-development.org/ixi-dataset/>

The purpose of having these residual blocks with shortcuts is to allow for a shallower network. The usage of these blocks has some advantages in terms of reducing overfitting and, more importantly, reducing degradation due to vanishing gradients during training [27]. We propose to use projection shortcuts [27], implemented as a 1×1 2D convolution. Projection shortcuts, unlike the identity, give the model the ability to do cross-channel learning. That is, it acts like a coordinate-dependant linear transformation of the filter space [28].

4.3 Decoder Ensemble

The decoder is composed of four decoders, one for each of the four channels of the wavelet transform. Each decoder network upsamples its input feature map from the last residual block of the bottleneck. We upsample using transposed convolutions with a stride of 2, such that the network will learn the upsampling operator. As the activation function we used a Leaky ReLU with $\alpha = 0.01$:

$$f(x) = \begin{cases} \alpha x, & x \leq 0 \\ x, & x > 0 \end{cases}.$$

Finally, the features are batch-normalized.

The Leaky ReLU activation function has shown better performance as compared to the standard definition [29], [30]. Additionally, it retains its simplicity while calculating derivatives, so it does not make the model training much slower.

The hypothesis behind using four decoders is that allowing each decoder to specialize in only one channel of the Wavelet transform will result in better convergence [31], [32]. Therefore, exploiting the statistical differences and the relative independence of the wavelet coefficients.

The output of each decoder is a single 2D convolution layer with linear activation and one filter. This layer is used to predict the corresponding wavelet coefficients. Finally, the concentration distribution is the inverse wavelet transform.

4.4 Training

We used the previously described set to train our model. All the weights were initialized using the technique described in He et al. [33]. To train the model, we use Adam [34] with a fixed learning rate of 10^{-5} and a momentum of 0.9. Before each epoch, the training set is shuffled. We chose a mini-batch size of 32, and we ensured that every image is used only once per epoch.

The loss function was the normalized absolute deviation plus a regularizing term in the network parameters:

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^F \frac{\|\alpha_i^j - \hat{\alpha}(\theta)_i^j\|_1}{\|\alpha_i^j\|_1} + \lambda \|\theta\|_2^2,$$

where N is the number of training images, F the number of channels in the wavelet transform, α is the target and $\hat{\alpha}(\theta)$, the network estimation of the wavelet coefficients. It is important to use a relative deviation to avoid images with high SPIO concentrations to dominate [35].

5 EXPERIMENTS AND RESULTS

To test the performance of our network, we did four experiments: Firstly, with synthetic images generated as in the training group. Secondly, with simulated images from an organ not used in training (breast instead of the brain). Thirdly, with simulations of a simple geometry with a nearby air sphere. And fourthly, with real data acquired with the scanner. The first three experiments allowed us to test the network in complex and different controlled situations, whereas the fourth, although simpler in geometry, includes all the complexity of the actual acquisition.

Additionally, to study the effect of noise, we tested the model simulating a cylinder with different SPIO concentrations and noise levels.

To measure the prediction error, we used the Normalized Mean Absolute Error (NMAE) defined as

$$\text{NMAE}(\hat{C}, C) = \frac{\|\hat{C} - C\|_1}{\|C\|_1}.$$

Where \hat{C} is the estimated SPIO concentration in image space, and C , is the real concentration distribution.

All the experiments were conducted on a PC with an Intel i9-7900x processor and two NVIDIA GTX 1080Ti. We implemented our model on Keras [36] with the TensorFlow backend [37]. The training took approximately 40 hours.

Experiment 1: Simulated images. To validate the network, we estimated the SPIO concentration with images and particle distributions similar to those used during training. We obtained an average NMAE of 0.17 ± 0.10 . Figure 4 shows the estimation done by DeepSPIO on synthetic data and compares it to the simulated ground truth. There is an excellent agreement.

Experiment 2: Simulated images from a different set. Since the training IXI dataset only has brain images, there is a risk for the model to be biased towards spurious artifacts only present in those images. To test for that bias, we simulated the SPIO-induced distortions in human breast images. Figure 5 shows an example of the DeepSPIO estimation and the corresponding ground-truth. We did not find any bias.

Experiment 3: Distortion from an air source. To verify the effect of other off-resonance sources, we tested our method with a simulated cylinder and a nearby sphere with an air-tissue susceptibility difference of $\Delta\chi_{(\text{cgs})} = 7.5 \times 10^{-8}$. This experiment simulates the effect of air, found in some parts of the body, such as the sinus cavities in the head. Figure 6 shows examples of DeepSPIO estimations for different concentrations and the corresponding ground truth. We used an idealistic shape (sphere) as opposed to a more complex shape. As can be appreciated in Figure 6, this is not important because the effect of the air is negligible in comparison with the high concentration SPIOs.

Experiment 4: Real images. We tested our model with actual MRI data. We constructed several phantoms consisting of plastic cylinders filled with agar, with a small cylindrical bottle of SPIO (Fig. 7).

We tested five different concentrations of SPIO, $C = \{27.9, 13.9, 9.3, 7.0, 3.5\} \mu\text{g}/\text{mm}^3$. We also made another

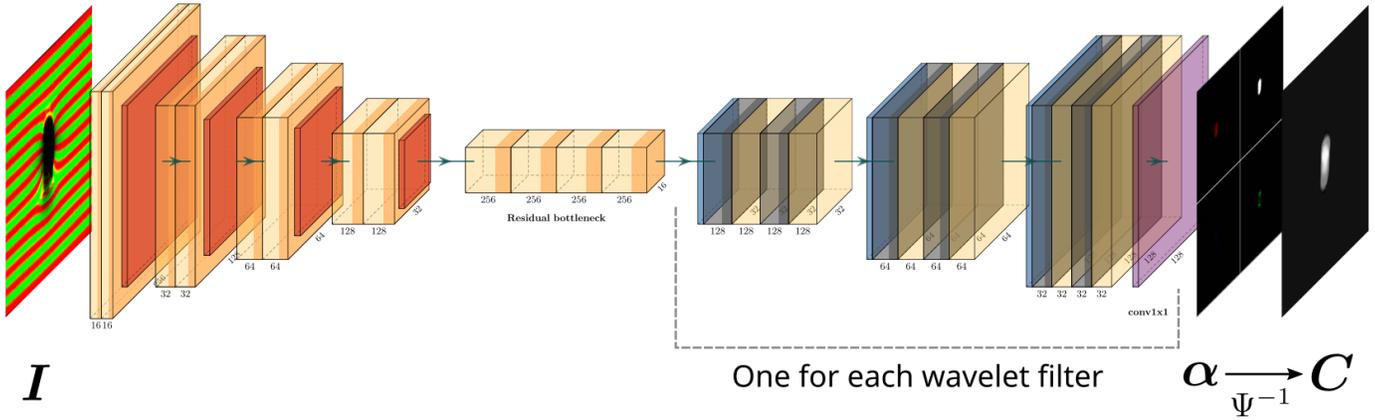


Fig. 3: Proposed Architecture of DeepSPIO. Each decoder upsamples its input using the shared features from the bottleneck. Then it performs a transposed convolution followed by convolutions to densify and specialize the reconstruction of each wavelet filter. The final convolution represents the corresponding wavelet filter.

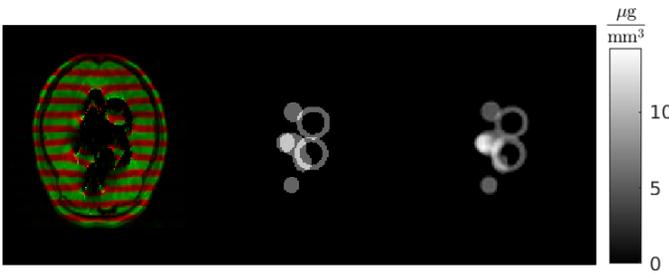


Fig. 4: Simulated SPIO distortions on IXI PD image, ground truth, and SPIO prediction. This image has an NMAE of 0.23. The group average is 0.17.

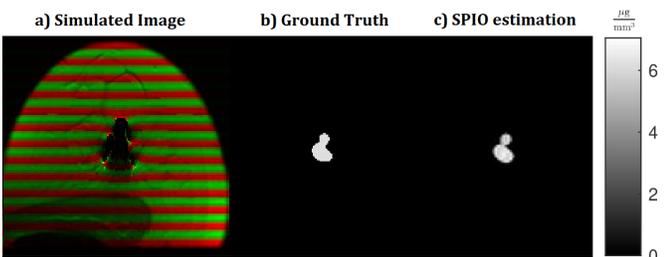


Fig. 5: Simulated SPIO distortions on a breast image, ground truth, and SPIO prediction. NMAE: 0.16.

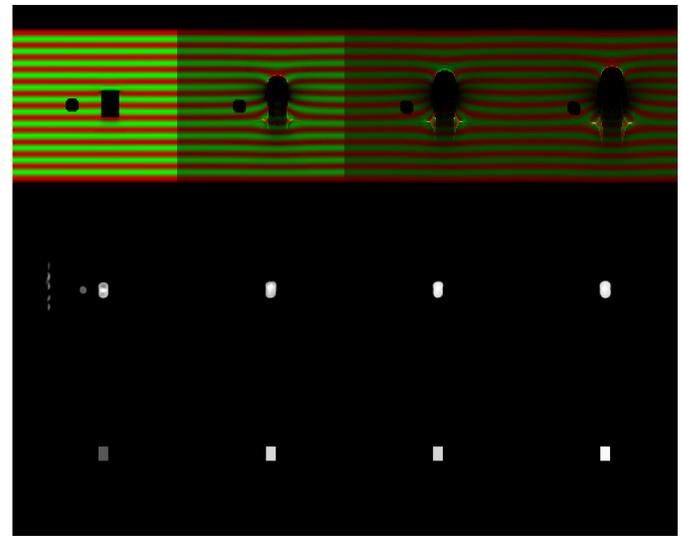


Fig. 6: Examples of the air experiment. The top row has the simulated images. The middle row has DeepSPIO predictions. And the bottom row has the corresponding ground truths. Cylinder concentrations increase from left to right ($0.22, 6.7, 13, 20 \mu\text{g}/\text{mm}^3$). The NMAE indices are: 1.96, 0.24, 0.30, and 0.21.

batch of samples with the following concentrations to validate our results: $C = \{13.9, 7, 3.5, 1.7\} \mu\text{g}/\text{mm}^3$.

The true concentration and the estimated concentration (mean of the center region) are shown in Table 1. The images were acquired with the View Line sequence in a 7T Bruker scanner. The MRI parameters of the sequences are as follows: Field of View (FOV): 50 mm squared; stripes width: 1 mm; resolution: 256×256 ; TE/TR: 8.7/800 ms; view lines: 25; turbo spin-echo factor: 3; total scan time: 30 minutes.

The acquired images and the output of our model for the SPIO concentration are shown in Fig. 8.

5.1 Noise sensitivity

To appreciate the effect of noise, we simulated a cylinder inside an agar container, similar to our real sample. The

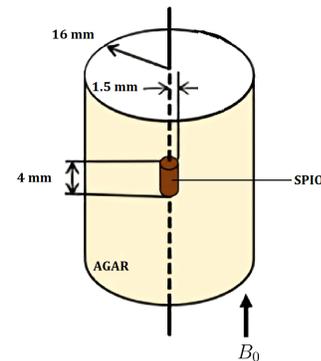


Fig. 7: Agar-filled phantom with SPIO particles. Various concentrations were used.

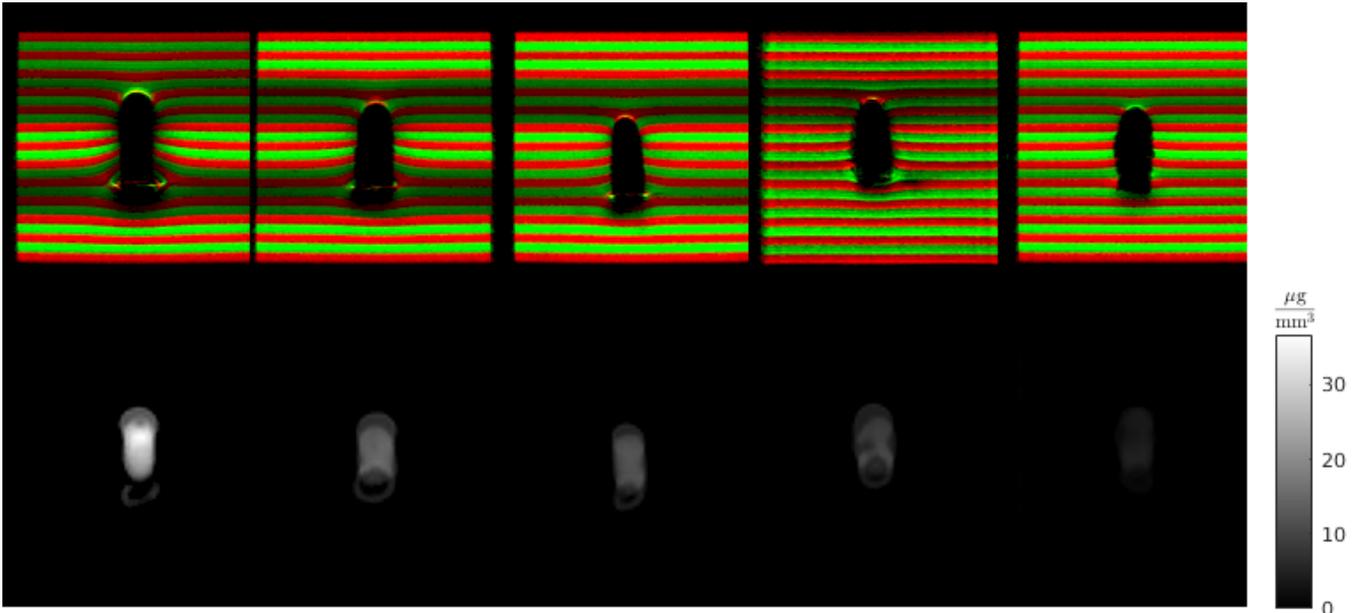


Fig. 8: Results on the first batch of real samples. With decreasing concentration from left to right.

TABLE 1: Comparison between average estimated concentration and average ground truth for both batches.

| Batch | Concentration ($\mu\text{g}/\text{mm}^3$) | | | |
|-------|---|-----------|-------|---------|
| | True | Estimated | Error | % Error |
| 1 | 27.90 | 27.35 | 0.55 | 1.97% |
| 1 | 13.95 | 13.17 | 0.78 | 5.59% |
| 2 | 13.95 | 11.78 | 2.17 | 15.5% |
| 1 | 9.30 | 10.92 | 1.62 | 17.4% |
| 1 | 6.98 | 7.88 | 0.90 | 12.89% |
| 2 | 6.98 | 6.42 | 0.56 | 8.02% |
| 1 | 3.49 | 3.53 | 0.04 | 1.14% |
| 2 | 3.49 | 3.10 | 0.39 | 11.17% |
| 2 | 1.75 | 1.41 | 0.34 | 19.42% |

concentration of SPIO was changed between 0.22 and 27.9 $\mu\text{g}/\text{mm}^3$. We tested the following noise levels for each concentration: 5 dB, 11 dB, 19 dB and ∞ . We simulated each concentration ten times, adding instances of Gaussian noise to the real and imaginary components of each pixel. Fig. 9 shows the mean and standard deviation of the NMAE index for each concentration. Fig. 10 shows the results in more detail, this time with 100 repetitions, for a concentration of 20.5 $\mu\text{g}/\text{mm}^3$. We can see how the mean error and its standard deviation are reduced as the SNR increases.

5.2 Analysis

We used the GRAD-CAM technique to visualize how the network is using the information in the input image and to make sure it is not using spurious artifacts [38]. GRAD-CAM produces visual representations in the form of attention maps.

The GRAD-CAM algorithm uses the gradient information flowing in the last convolutional layer of the neural network. GRAD-CAM produces a map showing each neuron’s importance in the decisions made by the network. This map is calculated by computing a score that weights the derivative of the output y of one of the decoders with

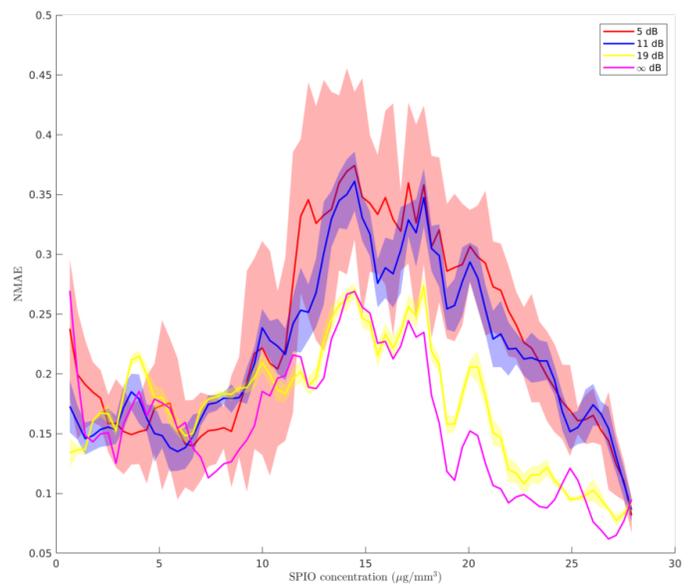


Fig. 9: Effect of noise. Model performance (measured by the NMAE) as a function of the concentration. For each SNR level, the mean and standard deviation are shown.

respect to the k -th feature map A^k of some given layer. Mathematically this is $\partial y / \partial A^k$, which corresponds to the gradient backpropagated. The following weight is defined

$$\alpha_k = \frac{1}{Z} \sum_i \sum_j \frac{\partial y}{\partial A_{ij}^k},$$

where i and j corresponds to the feature map’s dimensions, Z is the number of elements in the feature map, and α_k corresponds to the global average pooling of the k -th feature map of the selected layer. This expression can be interpreted as a partial linearization of the neural network downstream from the selected layer, and the importance of the k -th fea-

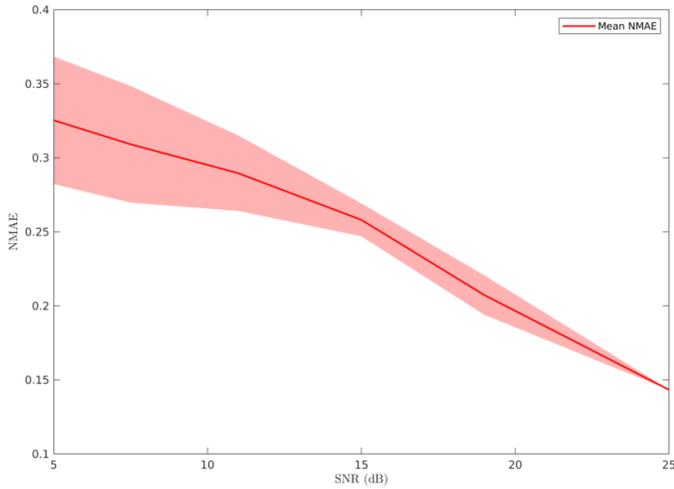


Fig. 10: Effect of noise: Model performance (measured by the NMAE) on a sample with a concentration of $20.5 \mu\text{g}/\text{mm}^3$ as a function of the SNR. The mean and standard deviation of the NMAE are shown.

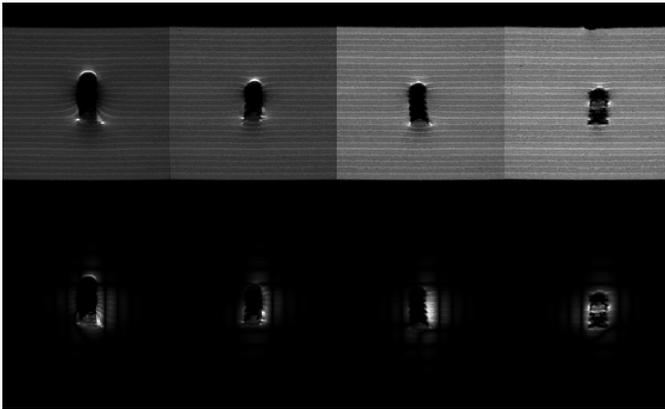


Fig. 11: Attention maps obtained using GRAD-CAM over the second batch of real samples. Using the last layer of the bottleneck as feature derivative. In the top row we show the images of the second batch. In the bottom row their corresponding attention maps.

ture map in making the estimation y . Finally, the attention map $L_{\text{GRAD-CAM}}$ is

$$L_{\text{GRAD-CAM}} = \text{ReLU}\left(\sum_k \alpha_k A_k\right).$$

This map is multiplied with the input image to visualize what parts of the image are used to make the actual estimation. In Figure 11 we show the attention maps of the network for the MRI images of experiment 4. We can see that it consistently uses the deformation and shape of the slice profiles to make the estimation.

We also use this technique to see what is encoding each layer. We can appreciate in the maps of the first layers that the network starts by obtaining the center lines of the slice selection. This result makes sense because the ΔB_z is encoded in the lines, as a deviation from the central frequency. In subsequent layers, before the bottleneck, it uses those lines to estimate concentration, as needed by the

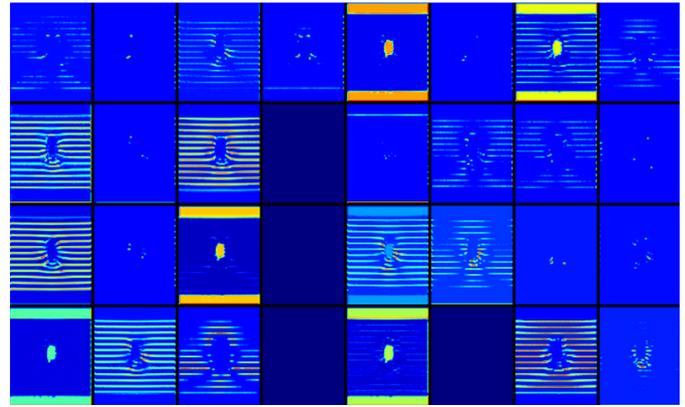


Fig. 12: Attention maps of each feature of the fifth convolutional layer.

theoretical model of the MRI signal (Eq. 1). In Figure 12, we can see the attention maps for the fifth convolutional layer, and how the field deviation is estimated. See the Appendix for all the attention maps.

To understand the shape of the error curves in the previous section, we performed a saliency analysis [39]. The saliency analysis's basic idea is to back-propagate the gradient of the output and project it into the input image. These maps are generally used in image processing tasks because they show how important are the features for the model. We used the gradient from our loss function.

In Figure 13, as expected, we can see that the obtained features of the noisy samples are of much less quality than their noise-free versions. This result is in concordance with the higher error rates for the noisy images.

In the same figure, it is also possible to appreciate how the features' importance changes with the concentration. They are less pronounced for concentrations in the middle of the range (less bright saliency maps in 13).

This behavior is due to the nature of the problem, which has two essential parts: finding the shape of the distribution of SPIO and finding its concentration. We can see that in low concentrations, it is straightforward for the network to predict the shape of the distribution. But it is difficult to estimate the concentration since there is no significant geometric distortion in the stripes. This effect is most notorious at concentrations below $3 \mu\text{g}/\text{mm}^3$, starting to become indistinguishable at around $1 \mu\text{g}/\text{mm}^3$. As the concentration increases (around 7 to $16 \mu\text{g}/\text{mm}^3$), information about the shape decreases, but information about the geometric distortions of the stripes increases. Finally, for higher concentrations, the geometric distortions are high enough to compensate for the lack of shape information, and the predictions are accurate again.

We also computed the saliency maps for the experiment with the simulated air bubble (Figure 14). As in the previous figure, it is possible to appreciate how for low concentrations, the main features are the shapes themselves. For higher concentrations, the relevant features are the geometric distortions of the stripes. For the lowest concentration ($1 \mu\text{g}/\text{mm}^3$), our method may confuse the field distortions coming from the SPIO with the ones coming from the air. This is not the case for higher concentrations, as it is

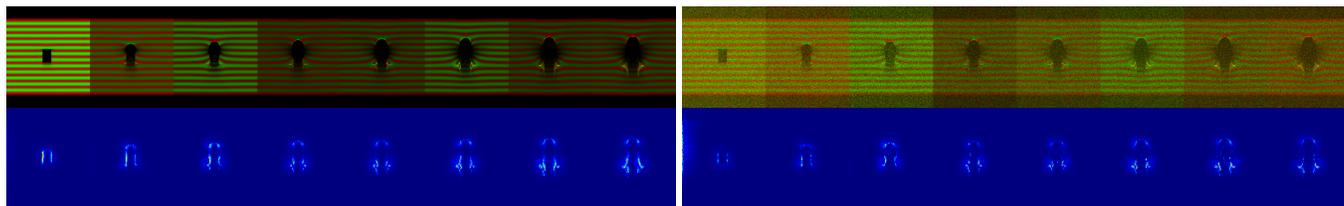


Fig. 13: Saliency maps obtained using guided back-propagation in simulated concentrations. First for a noiseless situation and then for an SNR level of 5 dB. In the top row we show the simulated images for the following concentrations: 0.22, 3.23, 6.91, 10.58, 14.26, 17.94, 21.62, and 25.31 $\mu\text{g}/\text{mm}^3$. In the bottom row, their corresponding saliency maps.

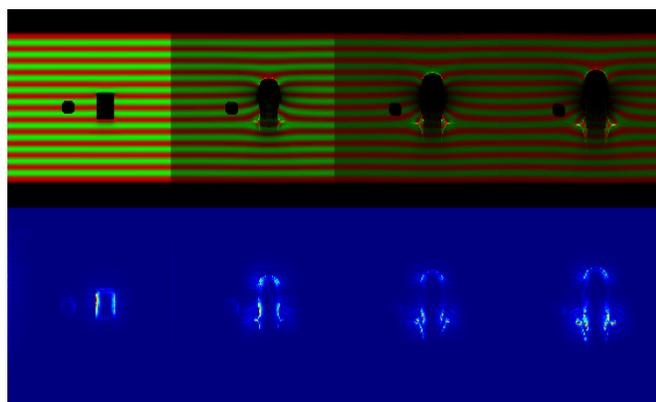


Fig. 14: Saliency maps from the simulations with an air sphere and a cylinder with SPIO. The concentrations, from left to right are 0.22, 6.91, 14.26 and 21.62 $\mu\text{g}/\text{mm}^3$. The top row has the simulated images and the bottom row has the corresponding saliency maps.

confirmed in the saliency maps, where the network did not consider features from the air bubble.

Although the proposed experiment is a simplified model of the different air cavities that can be found in the body, it helps to isolate and visualize the effects. It shows that the effect is negligible outside the sphere itself, independently of the shape.

6 DISCUSSION

In this study, we proposed a solution to the challenging problem of accurately quantifying high concentrations of SPIO particles using a fully convolutional neural network: DeepSPIO. DeepSPIO estimates the concentration distribution accurately in all the cases we tested. For training, we used images simulated from the View Line acquisition with simple shapes of SPIO distributions inserted in brain images of the IXI dataset.

Roughly speaking, one can appreciate two distinct image features in the View Line images that can be used to estimate the SPIO concentration. One feature, produced by the Off-resonance, is the bending of the slice selection. The only spins excited are those that have a resonant frequency in the bandwidth of the excitation pulse. Since the SPIO concentration distorts the field, that frequency changes considerably, and the excitation is displaced and warped, producing curved lines. The second feature, also produced by the Off-resonance, is a noticeable signal drop off near the SPIO

location. This signal loss is due to the significant frequency deviation such that the slice thickness gets reduced and tends to disappear. Additionally, the through-plane phase dispersion cancels the signal.

It is interesting to observe how the network processes these features, which could be appreciated in the maps obtained with GRAD-CAM and the saliency analysis. We can see that the attention of the first layer is concentrated in the center of each line of the slice selection. In other words, it is looking at the curvature of the lines. This feature is what an analytical solution would require. Subsequent layers gradually reduce the attention on the center lines to add other details. This behavior, together with the fact that the network was trained with synthetic data, confirms that the model learned to solve the problem and is not a result of overfitting to the training data.

The saliency maps also indicate that it is more challenging to estimate concentrations in the middle range of concentrations. This is due to the less explicit shape of the space occupied by the particles, and the lack of features farther apart to compensate for the lost shape. This effect explains the increase of the error for concentrations in the range from 7 $\mu\text{g}/\text{mm}^3$ to 16 $\mu\text{g}/\text{mm}^3$. As expected, noise in the images also reduces useful features and therefore increased the variability of the estimation.

To address the lack of accuracy in the middle ranged concentrations, in future work, we propose to optimize the acquisition parameters. Candidate variables to study are the bandwidth or the view line width since a reduced slice thickness would improve the detection of field distortions, at the cost of longer acquisitions. Timing parameters, particularly TE, could be relevant for improving the precision in this middle range, as well.

Although we only used synthetic data to train on a particular dataset, the model obtains accurate estimations of the SPIO distribution on different types of images. These results, plus the attention maps, show that the network is learning the underlying problem. The proposed approach also has a convenient and practical consequence. It allows to use forward models of different levels of sophistication and to perform transfer learning to measure the distribution in different parts of the body more accurately. In our experiments, although we estimated the field distortions in three dimensions, we only used two-dimensional images. Using 3D images could have a positive effect on the quality of the estimated model, although at a higher computational cost.

We used a wavelet transform of only four channels, but the proposed architecture can be easily extended to any

number of channels. Adding more wavelet filters may have a beneficial effect, as shown in other applications [31], [32].

An essential benefit of using Deep Learning to solve the SPIO quantification problem is that calculating the distribution map of a given sample is done within seconds. Obviously, the training takes a long time (40 hours in our computer for 4 million images), but this is done only once, and more importantly, it is not sensitive to the type of images used during training. The shape of the SPIO during training tends to be much more complicated than actual distributions. This network can already be applied to any anatomy.

7 CONCLUSION

We have described the bases of a new class of neural network model that solves the SPIO quantification problem. The quantification of SPIO concentration distribution is a difficult problem to solve with traditional analytical models, that typically use the phase of the signal, because the field variations are so significant that the signal disappears, and the phase becomes very unreliable. We proposed a network architecture that first encodes the image using strided convolutions. After the encoding, the data goes through a bottleneck with several residual blocks that allow for shortcuts. Finally, the data is decoded using one decoder per coefficient of a wavelet transform to exploit the statistical differences between wavelet channels. This architecture was successfully used to estimate the SPIO concentration in simulated data and real data obtained from an MRI scanner. It was able to estimate the concentration and geometric distribution accurately. Attention maps show that the network is looking at the slice selection bending, which coincides with the theoretical model and reassures that the network is learning the underlying problem and not overfitting to the training.

ACKNOWLEDGMENT

The authors gratefully acknowledge ANID for funding this research through PIA-Anillo ACT1416, ACT192064, Fondecyt 1191710, and the Millenium Science Initiative of the Ministry of Economy, Development, and Tourism, grant Nucleus for Cardiovascular Magnetic Resonance.

REFERENCES

- [1] J. Langley, W. Liu, E. K. Jordan, J. A. Frank, and Q. Zhao, "Quantification of SPIO Nanoparticles in vivo Using the Finite Perturber Method," *Magnetic resonance in medicine*, vol. 65, no. 5, pp. 1461–1469, May 2011. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3612521/>
- [2] Y.-X. J. Wang, "Superparamagnetic iron oxide based MRI contrast agents: Current status of clinical application," *Quantitative Imaging in Medicine and Surgery*, vol. 1, no. 1, p. 6, 2011. 1
- [3] K. Coenegrachts, H. Orlent, L. t. Beek, M. Haspeslagh, S. Bipat, J. Stoker, and H. Rigauts, "Improved focal liver lesion detection: Comparison of single-shot spin-echo echo-planar and superparamagnetic iron oxide (SPIO)-enhanced MRI," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 1, pp. 117–124, 2008. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.21247>
- [4] G. McAuley, M. Schrag, P. Sipos, S.-W. Sun, A. Obenaus, J. Neelavalli, E. M. Haacke, B. Holshouser, R. Madácsi, and W. Kirsch, "Quantification of punctate iron sources using magnetic resonance phase," *Magnetic Resonance in Medicine*, vol. 63, no. 1, pp. 106–115, Jan. 2010. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.22185>
- [5] K. G. B. Rasmussen, M. J. Kristensen, R. G. Blendal, L. R. Ostergaard, M. Plochanski, K. O'Brien, C. Langkammer, A. Janke, M. Barth, and S. Bollmann, "DeepQSM - Using Deep Learning to Solve the Dipole Inversion for MRI Susceptibility Mapping," *bioRxiv*, Mar. 2018. [Online]. Available: <http://biorxiv.org/lookup/doi/10.1101/278036>
- [6] J. Yoon, E. Gong, I. Chatnuntawech, B. Bilgic, J. Lee, W. Jung, J. Ko, H. Jung, K. Setsompop, G. Zaharchuk, E. Y. Kim, J. Pauly, and J. Lee, "Quantitative susceptibility mapping using deep neural network: QSMnet," *NeuroImage*, vol. 179, pp. 199–206, Oct. 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1053811918305378>
- [7] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *arXiv:1505.04597 [cs]*, May 2015, arXiv: 1505.04597. [Online]. Available: <http://arxiv.org/abs/1505.04597>
- [8] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017. 1, 2
- [9] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, "Deep learning for undersampled MRI reconstruction," *Physics in Medicine & Biology*, vol. 63, no. 13, p. 135007, Jun. 2018. [Online]. Available: <https://doi.org/10.1088%2F1361-6560%2F63%2F13>
- [10] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, "Noise2Noise: Learning image restoration without clean data," in *Proceedings of the 35th International Conference on Machine Learning*, ser. *Proceedings of Machine Learning Research*, J. Dy and A. Krause, Eds., vol. 80. Stockholmsmässan, Stockholm Sweden: PMLR, 10–15 Jul 2018, pp. 2965–2974. [Online]. Available: <http://proceedings.mlr.press/v80/lehtinen18a.html>
- [11] R. Salomir, B. D. d. Senneville, and C. T. Moonen, "A fast calculation method for magnetic field inhomogeneity due to an arbitrary distribution of bulk susceptibility," *Concepts in Magnetic Resonance: Magnetic Resonance Engineering*, vol. 19B, no. 1, pp. 26–34, 2003. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/cmr.b.10083>
- [12] K. Shmueli, J. A. de Zwart, P. van Gelderen, T.-Q. Li, S. J. Dodd, and J. H. Duyn, "Magnetic susceptibility mapping of brain tissue in vivo using MRI phase data," *Magnetic Resonance in Medicine*, vol. 62, no. 6, pp. 1510–1522, Dec. 2009. 2
- [13] T. Liu, I. Khalidov, L. de Rochefort, P. Spincemaille, J. Liu, A. J. Tsiouris, and Y. Wang, "A novel background field removal method for MRI using projection onto dipole fields (PDF)," *NMR in biomedicine*, vol. 24, no. 9, pp. 1129–1136, Nov. 2011. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3628923/>
- [14] C. Milovic, B. Bilgic, B. Zhao, J. Acosta-Cabrero, and C. Tejos, "Fast nonlinear susceptibility inversion with variational regularization," *Magnetic Resonance in Medicine*, vol. 80, no. 2, pp. 814–821, 2018. 2
- [15] Y. Kee, Z. Liu, L. Zhou, A. Dimov, J. Cho, L. de Rochefort, J. K. Seo, and Y. Wang, "Quantitative Susceptibility Mapping (QSM) Algorithms: Mathematical Rationale and Computational Implementations," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 11, pp. 2531–2545, Nov. 2017. 2
- [16] J. Liu, T. Liu, L. de Rochefort, J. Ledoux, I. Khalidov, W. Chen, A. J. Tsiouris, C. Wisnieff, P. Spincemaille, M. R. Prince, and Y. Wang, "Morphology Enabled Dipole Inversion for Quantitative Susceptibility Mapping Using Structural Consistency Between the Magnitude Image and the Susceptibility Map," *NeuroImage*, vol. 59, no. 3, pp. 2560–2568, Feb. 2012. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3254812/>
- [17] L. de Rochefort, T. Liu, B. Kressler, J. Liu, P. Spincemaille, V. Lebon, J. Wu, and Y. Wang, "Quantitative susceptibility map reconstruction from MR phase data using bayesian regularization: validation and application to brain imaging," *Magnetic Resonance in Medicine*, vol. 63, no. 1, pp. 194–206, Jan. 2010. 2

[18] W. Lu, K. B. Pauly, G. E. Gold, J. M. Pauly, and B. A. Hargreaves, "SEMAC: Slice encoding for metal artifact correction in MRI," *Magnetic Resonance in Medicine*, vol. 62, no. 1, pp. 66–76, Jul. 2009. [Online]. Available: <http://doi.wiley.com/10.1002/mrm.21967> 2

[19] K. M. Koch, J. E. Lorbiecki, R. S. Hinks, and K. F. King, "A multispectral three-dimensional acquisition technique for imaging near metal implants," *Magnetic Resonance in Medicine*, vol. 61, no. 2, pp. 381–390, Feb. 2009. 2

[20] Z. H. Cho, D. J. Kim, and Y. K. Kim, "Total inhomogeneity correction including chemical shifts and susceptibility by view angle tilting," *Medical Physics*, vol. 15, no. 1, pp. 7–11, Feb. 1988. 2

[21] S. Liu et al., "B0 Field Map Estimation with View Line Sequence." *ISMRM*, 2019. 2, 3

[22] J. C. Ye and W. K. Sung, "Understanding Geometry of Encoder-Decoder CNNs," [arXiv:1901.07647 \[cs, stat\]](https://arxiv.org/abs/1901.07647), Jan. 2019, [arXiv: 1901.07647](https://arxiv.org/abs/1901.07647). [Online]. Available: <http://arxiv.org/abs/1901.07647> 2

[23] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1361841517301135> 2

[24] T. Liu, P. Spincemaille, L. d. Rochefort, B. Kressler, and Y. Wang, "Calculation of susceptibility through multiple orientation sampling (COSMOS): A method for conditioning the inverse problem from measured magnetic field map to susceptibility source image in MRI," *Magnetic Resonance in Medicine*, vol. 61, no. 1, pp. 196–204, 2009. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.21828> 2

[25] J. F. Blinn, "A Generalization of Algebraic Surface Drawing," *ACM Transactions on Graphics*, vol. 1, no. 3, p. 22. 3.2

[26] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," [arXiv:1502.03167 \[cs\]](https://arxiv.org/abs/1502.03167), Feb. 2015, [arXiv: 1502.03167](https://arxiv.org/abs/1502.03167). [Online]. Available: <http://arxiv.org/abs/1502.03167> 4.1

[27] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," [arXiv:1512.03385 \[cs\]](https://arxiv.org/abs/1512.03385), Dec. 2015, [arXiv: 1512.03385](https://arxiv.org/abs/1512.03385). [Online]. Available: <http://arxiv.org/abs/1512.03385> 4.2

[28] M. Lin, Q. Chen, and S. Yan, "Network In Network," [arXiv:1312.4400 \[cs\]](https://arxiv.org/abs/1312.4400), Dec. 2013, [arXiv: 1312.4400](https://arxiv.org/abs/1312.4400). [Online]. Available: <http://arxiv.org/abs/1312.4400> 4.2

[29] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for Activation Functions," [arXiv:1710.05941 \[cs\]](https://arxiv.org/abs/1710.05941), Oct. 2017, [arXiv: 1710.05941](https://arxiv.org/abs/1710.05941). [Online]. Available: <http://arxiv.org/abs/1710.05941> 4.3

[30] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical Evaluation of Rectified Activations in Convolutional Network," [arXiv:1505.00853 \[cs, stat\]](https://arxiv.org/abs/1505.00853), May 2015, [arXiv: 1505.00853](https://arxiv.org/abs/1505.00853). [Online]. Available: <http://arxiv.org/abs/1505.00853> 4.3

[31] H. Huang, R. He, Z. Sun, and T. Tan, "Wavelet-SRNet: A Wavelet-Based CNN for Multi-scale Face Super Resolution," in *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct. 2017, pp. 1698–1706, iSSN: 2380-7504. 4.3, 6

[32] X. Fu, B. Liang, Y. Huang, X. Ding, and J. Paisley, "Lightweight Pyramid Networks for Image Deraining," [arXiv:1805.06173 \[cs\]](https://arxiv.org/abs/1805.06173), May 2018, [arXiv: 1805.06173](https://arxiv.org/abs/1805.06173). [Online]. Available: <http://arxiv.org/abs/1805.06173> 4.3, 6

[33] K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015, pp. 1026–1034. 4.4

[34] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," [arXiv:1412.6980 \[cs\]](https://arxiv.org/abs/1412.6980), Dec. 2014, [arXiv: 1412.6980](https://arxiv.org/abs/1412.6980). [Online]. Available: <http://arxiv.org/abs/1412.6980> 4.4

[35] F. Rousselle, C. Knaus, and M. Zwicker, "Adaptive sampling and reconstruction using greedy error minimization," *ACM Trans. Graph.*, vol. 30, no. 6, pp. 159:1–159:12, Dec. 2011. [Online]. Available: <http://doi.acm.org/10.1145/2070781.2024193> 4.4

[36] F. Chollet et al., "Keras," <https://keras.io>, 2015. 5

[37] M. Abadi, A. Agarwal, and P. B. et al., "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: <http://tensorflow.org/> 5

[38] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization," [arXiv:1610.02391 \[cs\]](https://arxiv.org/abs/1610.02391), Oct. 2016, [arXiv: 1610.02391](https://arxiv.org/abs/1610.02391). [Online]. Available: <http://arxiv.org/abs/1610.02391> 5.2

[39] T. N. Mundhenk, B. Y. Chen, and G. Friedland, "Efficient saliency maps for explainable ai," 2019. 5.2



Gabriel della Maggiora Computer Scientist (2016) and MSc. in Electrical Engineering (2019) from Pontificia Universidad Católica de Chile. His research interests are Deep Learning architectures for inverse problems, Generative Models and Reinforcement Learning.



Carlos Castillo-Passi Student for a PhD in Biological and Medical Engineering at the Institute of Biological and Medical Engineering in the Pontificia Universidad Católica de Chile. MSc in Electrical Engineering (2018), and Electrical Engineering (2018) both from Pontificia Universidad Católica de Chile. Interested in inverse problems, image reconstruction and formation, micro-structural imaging, and MRI hardware.



Wenqi Qiu graduated from Xiamen University in 2014, and finished a master's course in the department of electrical engineering and information system, the University of Tokyo in 2018. She works in the nuclear magnetic spectrum form 2014-2016 in Fujian Key Laboratory of Plasma and Magnetic Resonance, China. Then mainly focus on magnetic resonance imaging from 2016-2018 in Sekino Laboratory at the University of Tokyo. During this period, she also visited and studied as an exchange student in Pontificia Universidad Católica de Chile, Santiago, Chile.



Shuang Liu was born in Wuhan, China, in 1992. He received his B.Eng. degree in electronic engineering from Wuhan University of Technology, Wuhan, China, in 2015, and the M.Eng. degree in electronic engineering from the University of Tokyo, in 2019. He is currently a Ph.D. student with Sekino lab at the University of Tokyo at Tokyo, Japan, where his research interests mainly focus on the inverse problem of electromagnetics in biomedical engineering, in more specifically, including quantifying nanoparticles by Magnetic Resonance Imaging and Transcranial magnetic Stimulation coil design. From 2013 to 2014, he was an exchange student in Sahashi lab, Department of Electronic Engineering, School of Engineering, Tohoku University, Japan, he was responsible to simulate the Spin Transfer Torque device by FEM method. In 2019, He was involved with Pontificia Universidad Católica de Chile, Santiago, Chile, and Weill Cornell Medicine, New York, USA, as a visiting Graduate Student. He was the recipient of several awards including the Second Prize of the AP-RASC'19 Student Paper Competition, AP-RASC(2019), and the outstanding paper award of The Institute of Electrical Engineers of Japan (IEEJ) Dept. A (2019).



Carlos Milovic Doctor in Engineering Sciences (PhD), Pontificia Universidad Católica de Chile. Postdoctoral Research Fellow, Department of Medical Physics and Biomedical Engineering, University College London. Current work: inverse problems, image reconstruction, MRI.



Masaki Sekino Masaki Sekino PhD in Engineering from the University of Tokyo (2005). Associate professor at the Department of Electrical Engineering and Information Systems, Graduate School of Engineering, the University of Tokyo. Current work: biomedical applications of magnetic sensing, MRI, wearable electronic devices.



Cristian Tejos Cristian Tejos received a B.Sc. in Engineer (1997) and an M.Sc. in Electrical Engineering (2000) both from Pontificia Universidad Catolica de Chile. Afterwards, he received an M.Sc. in Engineering and Physical Science in Medicine from Imperial College London (2001) and Ph.D. from University of Cambridge (2007). He is currently Associate Professor at the Department of Electric Engineering Department and the Biomedical Imaging Center in the Pontificia Universidad Catolica de Chile.



Sergio Uribe Sergio Uribe is an associate professor in the Radiology Department and the Director of the Biomedical Imaging Center at Pontificia Universidad Católica de Chile (PUC). He is also the director of the Millennium Nucleus for Cardiovascular Magnetic Resonance. Sergio received his B.Sc. and M.Sc. in Electrical Engineering at PUC in 2005. Thereafter, he obtained a PhD in the Division of Imaging Sciences in 2009 in the School of Medicine at King's College London. His main research areas are biomedical

imaging and biomedical engineering.



Pablo Irrrazaval Electrical Engineer (1988) from Pontificia Universidad Catolica de Chile; MSc (1991) and PhD (1995) in Electrical Engineering from Stanford University. Professor at the Electric Engineering Department and the Institute for Biological and Medical Engineering in the Pontificia Universidad Catolica de Chile.