

Deep Learning for Low-Field to High-Field MR: Image Quality Transfer with Probabilistic Decimation Simulator

Hongxiang Lin¹, Matteo Figini¹, Ryutaro Tanno^{1,2}, Stefano B. Blumberg¹,
Enrico Kaden¹, Godwin Ogbale³, Biobele J. Brown⁴, Felice D’Arco⁵,
David W. Carmichael^{6,7}, Ikeoluwa Lagunju⁴, Helen J. Cross^{5,6},
Delmiro Fernandez-Reyes^{1,4}, and Daniel C. Alexander¹

¹ Centre for Medical Image Computing and Department of Computer Science,
University College London, UK

² Machine Intelligence and Perception Group, Microsoft Research Cambridge, UK

³ Department of Radiology, College of Medicine, University of Ibadan, Nigeria

⁴ Department of Paediatrics, College of Medicine, University of Ibadan, Nigeria

⁵ Great Ormond Street Hospital for Children, London, UK

⁶ UCL Great Ormond Street Institute of Child Health, UK

⁷ Department of Biomedical Engineering, King’s College London, UK

harry.lin@ucl.ac.uk

Abstract. MR images scanned at low magnetic field ($< 1\text{T}$) have lower resolution in the slice direction and lower contrast, due to a relatively small signal-to-noise ratio (SNR) than those from high field (typically 1.5T and 3T). We adapt the recent idea of Image Quality Transfer (IQT) to enhance very low-field structural images aiming to estimate the resolution, spatial coverage, and contrast of high-field images. Analogous to many learning-based image enhancement techniques, IQT generates training data from high-field scans alone by simulating low-field images through a pre-defined decimation model. However, the ground truth decimation model is not well-known in practice, and lack of its specification can bias the trained model, aggravating performance on the real low-field scans. In this paper we propose a probabilistic decimation simulator to improve robustness of model training. It is used to generate and augment various low-field images whose parameters are random variables and sampled from an empirical distribution related to tissue-specific SNR on a 0.36T scanner. The probabilistic decimation simulator is model-agnostic, that is, it can be used with any super-resolution networks. Furthermore we propose a variant of U-Net architecture to improve its learning performance. We show promising qualitative results from clinical low-field images confirming the strong efficacy of IQT in an important new application area: epilepsy diagnosis in sub-Saharan Africa where only low-field scanners are normally available.

1 Introduction

Magnetic Resonance Imaging (MRI) is now ubiquitous in neurology with a strong trend towards the use of high-field scanners, with 1.5T and 3T being the cur-

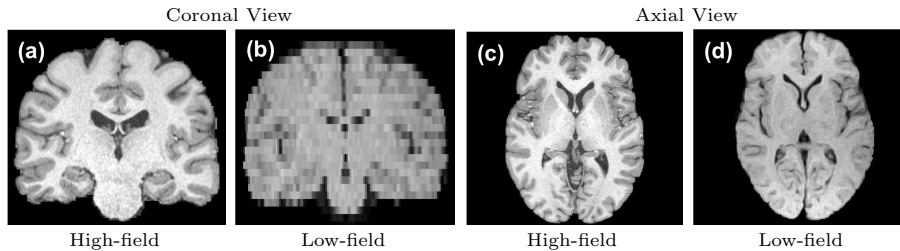


Fig. 1. High-field vs low-field MR scans: (a-b) Resolution change on coronal plane; (c-d) Contrast change on axial plane. Data sources: (a, c) 3T MRI from Human Connectome Project [2]; (b, d) 0.36T MRI acquired from University College Hospital, Ibadan.

rent clinical standard. However, low-field MRI scanners, less than 1T, are still common in low and middle income countries (LMICs), due to limited funds and frequent power outages. Low-field scanners suffer from lower signal-to-noise ratio (SNR) than high field at equivalent spatial resolution. To counteract the SNR reduction, practitioners commonly acquire images with non-adjacent thick slices to reduce the acquisition time and cross-talk artifacts in brain MRI scenario [1]. This leads to resolution reduction in the slice direction compared with the in-plane resolution and a loss of information due to gaps between slices; see Fig. 1(a-b). Moreover, the contrast between grey matter (GM) and white matter (WM) may be worse than in high field even at equivalent SNR and spatial resolution as illustrated in Fig. 1(c-d).

In this study, we aim to learn an image-translation mapping from low field to high field to perform super-resolution and contrast enhancement. In the literature, mathematical models have been proposed to describe the variation of MRI signal with the magnetic field [3,4], but such models are simplistic and do not include all effects on the final images, such as variability in the acquisition process. Furthermore, the reconstruction of missing information between the acquired slices is severely ill-posed, which hinders the practical capability of producing high-field like images. Several approaches in the literature aim to solve related problems. Bahrami *et al* [5] proposed a multi-level Canonical Correlation Analysis for estimating 7T from 3T images using paired training data. Wolterink *et al* [6] used the idea of cycle consistency to leverage the abundance of unpaired training sets and learn to synthesise CT from MRI. This approach is, however, known to be susceptible to hallucinations and may introduce spurious features in the output images [7].

Image Quality Transfer (IQT) is a machine learning framework used to enhance low-quality clinical data to the abundant neurological information in high-quality images. Most implementations of IQT simulate low quality data from high quality providing matched-paired for training. In [8,9,10,11] for instance, the corresponding low-field data are synthesised by downsampling and matching voxel-wise intensities coming from prior or empirical knowledge about actual low-field data. However, the trained model strongly depends on the accuracy of

low-field synthesis. To improve model generalisability, the prediction of a trained model should be built on unseen test data with less dependency of simulation.

In this paper, we build on the IQT framework to construct a mapping that estimates high-field images from the matched low-field inputs. The paired data, particularly in large numbers, are hard to acquire in one area due to the rare availability of high-field scanners in LMICs and low-field scanners in high income countries (HICs). Our key technical contribution is to propose a probabilistic decimation (downsampling) model to improve robustness of IQT training and to enhance images from low-field scanners. More specifically, low-field data generation comes from a probabilistic model which comprises random tissue-specific intensity statistics (e.g. SNR) and probabilistic semantic segmentation. We assume that an *a priori* distribution related to the tissue-specific SNR is available. The segmentation mask estimated by Statistical Parametric Mapping [12] is also probabilistic in terms of the tissue type. Therefore for one high-field subject, we can simultaneously generate the corresponding multiple low-field data and form the paired training data, a novel way of performing data augmentation. We then learn the low-field-to-high-field transformation by adapting the U-Net architecture [13] with a super-resolution module, a ‘‘bottleneck block’’, extending its depth to enable it to capture more global features of image contrast.

2 Methods

2.1 Formulation

Let a 3D low-field input patch x of size $w \times h \times d$ be corrupted by smoothing, low contrast, and random noise. It is randomly cropped from the original low-field MR volume denoted by X . Our aim is to reconstruct the sub-voxel information in the slice thickness direction and to attain the high SNR and contrast transferring to the corresponding high-field output patch y of size $w \times h \times kd$, where k is an up-sampling rate. Then we assemble all output patches into a high-field MR volume denoted by Y . The relationship between x and y is modelled by a degradation process of image quality, described by a function S such that

$$x = S(y, \alpha) + \epsilon, \quad (1)$$

where α denotes a vector of SNR components corresponding to prior knowledge of WM and GM in the low-field input volume, i.e. $\alpha = (SNR_X^{WM}, SNR_X^{GM})$. It is randomly sampled from the Gaussian distribution $\mathcal{N}(\mu, \Sigma)$ where μ is a mean vector and Σ is a covariance matrix. ϵ denoting background noise has a Gaussian distribution $\mathcal{N}(0, \sigma_{BG}^2)$. Section 2.2 will specify the formulation and algorithm for modelling S . We then employ deep learning, specifically a convolutional neural network, to estimate the inverse mapping S^\dagger .

We use a given M -paired training set $\mathcal{T}_M = \{(x_i, y_i)\}_{i=1}^M$ with a fixed α to train our convolutional neural networks over all sampled patches from all MR volumes. We optimise the network parameters θ by minimising the average of

the pixel-wise mean squared error (MSE) denoted by $\|\cdot\|_2^2$ over all training sets:

$$\theta^* = \arg \min_{\theta} \frac{1}{M} \sum_{i=1}^M \|S_{\theta}^{\dagger}(x_i) - y_i\|_2^2. \quad (2)$$

2.2 Probabilistic Decimation Simulator

Equation (1) enables us to produce additional training data by randomly sampling the coefficient α from an *a priori* distribution, forming the so-called probabilistic decimation simulator. It translates the voxel-wise low-field SNRs, related to the sampled α and the tissue category, to the high-field image and down-samples with a factor of k . We use this simulator to generate N low-field patches for each high-field patch y_i and form a new training set $\mathcal{T}_{M,N} = \{(x_{ij}, y_i) | i = 1, \dots, M, j = 1, \dots, N\}$. Henceforth, the new model is trained on the augmented set $\mathcal{T}_{M,N}$ with the following expression:

$$\theta^* = \arg \min_{\theta} \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \|S_{\theta}^{\dagger}(x_{ij}) - y_i\|_2^2. \quad (3)$$

We develop Alg. 1 for implementing the probabilistic decimation simulator for neural images. We transform high-field images $Y(\mathbf{v})$ to synthetic low-field images denoted by $\hat{X}(\mathbf{v})$ for any voxel coordinate \mathbf{v} by adapting the SNR in WM and GM to the values obtained in our reference low-field dataset. We assume that SNRs of WM and GM have a 2D Gaussian distribution and the background noise in the low-field or the high-field images has a 1D Gaussian distribution with a zero mean and a standard deviation of σ_X or σ_Y , respectively. We also assume $\sigma_X \gg \sigma_Y$ since the random noise in high field is negligible. The simulation procedure starts with the skull-stripped $Y(\mathbf{v})$ with isotropic voxels of length e_z . We then down-sample along the slice thickness direction (vertical, or z -direction). A 1D Gaussian filter $h_{\sigma}(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/(2\sigma^2)}$ is applied to the high-field images along the z -direction, where the σ is linked to a full-width at half maximum (FWHM): $\text{FWHM} = 2\sqrt{2\ln 2}\sigma$. The FWHM of the Gaussian filter is set to the slice thickness, or in terms of σ : $\sigma = ke_z/\sqrt{8\ln 2}$. Then the distance between slices is set to be larger than this slice thickness, emulating the gap between slices. The slices of the original image falling in the gaps have virtually no effect on the signal in the simulated image, similar to what happens in real acquisitions. The high-field images $Y(\mathbf{v})$ are first segmented into tissue categories $j = WM, GM, others$ (denoted by $M^j(\mathbf{v})$) using the unified segmentation algorithm in Statistical Parametric Mapping [12]. In this algorithm, the mask $M^j(\mathbf{v})$ corresponds to the probability that each voxel \mathbf{v} belongs to the tissue category j . SNR of the high-field image with respect to the tissue category j is defined as:

$$SNR_Y^j = \frac{\sum_{\mathbf{v}} M^j(\mathbf{v})Y(\mathbf{v})}{\sigma_Y \sum_{\mathbf{v}} M^j(\mathbf{v})}. \quad (4)$$

Algorithm 1 Probabilistic Decimation Simulator for low-field Image

Input: high-field Images $Y(\mathbf{v})$, masks $M^j(\mathbf{v})$ for $j = WM, GM, others$, downsampling scale $k \in \mathbb{N}$, background noise levels σ_X and σ_Y , low-field SNR distribution $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$.

- 1: $Y_{\downarrow k}(\mathbf{v}) = Y_{\downarrow k}(\tilde{v}, v') = \sum_{v''} Y(\tilde{v}, kv' - v'')h_{\sigma}(v'')$; \triangleright Downsample on v'' component.
- 2: $Y_{\downarrow k}^j(\mathbf{v}) = M^j(\mathbf{v})Y_{\downarrow k}(\mathbf{v})$; \triangleright Apply masks.
- 3: $SNR_Y^j = \sum_{\mathbf{v}} Y_{\downarrow k}^j(\mathbf{v}) / (\sigma_Y \sum_{\mathbf{v}} M^j(\mathbf{v}))$; \triangleright Compute SNRs for high field.
- 4: $(SNR_X^{WM}, SNR_X^{GM}) \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$; \triangleright Sample SNRs for low field.
- 5: $l^j = \begin{cases} SNR_X^j / SNR_Y^j, & j = WM, GM, \\ 1, & others; \end{cases}$ \triangleright Evaluate ratio of image intensity.
- 6: $\hat{X}(\mathbf{v}) = \sum_{j \in \{WM, GM, others\}} l^j Y_{\downarrow k}^j(\mathbf{v})$; \triangleright Transfer contrast.
- 7: $\hat{X}_{\epsilon}(\mathbf{v}) = \hat{X}(\mathbf{v}) + \epsilon(\mathbf{v})$ where $\epsilon(\mathbf{v}) \sim \mathcal{N}(0, \sigma_X^2)$. \triangleright Add noise.

Output: Noisy synthetic low-field image $\hat{X}_{\sigma}(\mathbf{v})$.

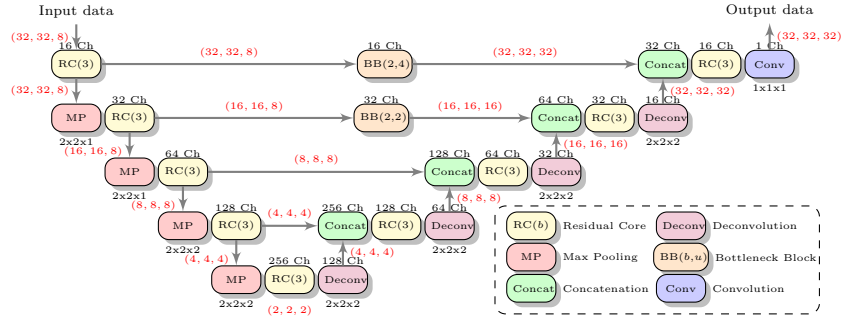
This allows us to evaluate ratios of low-field-to-high-field image intensity for both WM and GM; see Step 5. We then re-scale the high-field images with the ratios of image intensity according to tissue category, which results in the synthetic low-field images $\hat{X}(\mathbf{v})$. We finally add Gaussian white noise to $\hat{X}(\mathbf{v})$, with a standard deviation of σ_X .

2.3 Deep Learning Framework

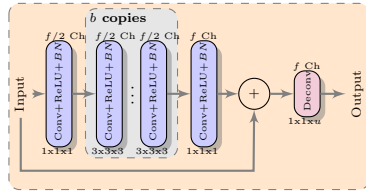
The classical 3D isotropic U-Net [14] maps two identical-size cubes serving as input and output through the encoder-decoder framework. Each level, defined as a collection of operations in between two shape deformations, for a typical U-Net consists of several convolutional layers together with a pooling layer. The activation from each level in the encoder is concatenated to the input features to the same level in the decoder, enabling the network to integrate both local and global image features. U-Net uses the “same” zero-padding technique so that feature sizes keep invariant during convolution.

In this work, we extend the U-Net architecture into mapping input and output patches differing with up-scaling factor k in the slice direction. Considering the case of $k = 4$ illustrated in Fig. 2, this anisotropic U-Net first partially down-samples the first two dimensions until the down-scaling features become isotropic and thereafter conducts isotropic down- and up-sampling. To achieve this, we define the following two operations:

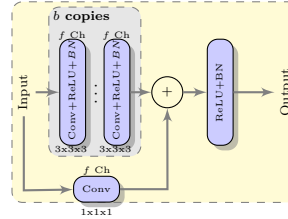
Bottleneck Block. To incorporate a super-resolution transformation into U-Net, we propose a bottleneck block used to connect corresponding levels of the contracting and expanding paths, as shown in Fig. 2(b). The design is inspired by bottleneck block in ResNet [15] and FSRCNN [16]. The bottleneck block $BB(b, u)$ has three hyperparameters: the input filter f , the number of shrinking layers b and the up-sampling scaling factor u . It shrinks half of the filters on consecutive $3 \times 3 \times 3$ convolutional layers between two endpoint convolutions with a kernel size of $1 \times 1 \times 1$. All convolution layers are activated by Rectified Linear Unit (ReLU) with Batch Normalization (BN). The skip connection enables the



(a) Anisotropic U-Net Architecture.



(b) Bottleneck Block $BB(b, u)$



(c) Residual Core $RC(b)$

Fig. 2. (a) The diagram of anisotropic U-Net (example for the up-scaling factor of $k = 4$). The operations, (b) Bottleneck Block $BB(b, u)$ with f filters and (c) Residual Core $RC(b)$ with f filters, are detailed. The round boxes correspond to the different operations illustrated in the bottom right of (a). The number of output channels, abbreviated as “Ch”, and the kernel size are denoted on top and bottom of the boxes. The arrows represent transfer of data with its corresponding shape highlighted.

training of deeper networks [15]. Resolution change is efficiently carried out by a transpose convolution, or deconvolution, with the same kernel and stride of $(1, 1, u)$.

Residual Core. To have more convolutional layers on each level, the residual core that is a revision of residual element in [17] is introduced in Fig. 2(c). This is a combination of several sequential $3 \times 3 \times 3$ convolutional layers, followed by ReLU and BN layers, skip connected with an $1 \times 1 \times 1$ fully convolutional layer. Then the output is attained before ReLU and BN again. Utilizing the consecutive convolutional layers enlarges each receptive field on each level. Moreover, the appended skip connection is able to avoid the vanishing gradient problem in neural networks with gradient-based learning methods.

3 Experiments

3.1 Implementation Details

Datasets. High-resolution axial T1-weighted images were obtained from the publicly available Human Connectome Project (HCP) dataset [2], acquired on a

3T Siemens Connectome scanner with an isotropic voxel size of $0.7 \times 0.7 \times 0.7$ mm³. To investigate sensitivity of the proposed U-Net, we trained it on two training sets with two up-scaling factors of $k = 4$ or 8. Specifically, the slice thickness/gap is 2.1 mm/0.7 mm for $k = 4$, and 4.2 mm/1.4 mm for $k = 8$. As a reference for low field, T1-weighted images were acquired on a 0.36T MagSense 360 MRI System scanner with a non-isotropic voxel size of $0.9 \times 0.9 \times 7.2$ mm³ including 6.0 mm slice thickness and 1.2 mm gaps. The distribution of white matter and grey matter SNRs in the low field was acquired from 28 image data from children with epilepsy in University College Hospital, Ibadan, whose ages are within a range from 2 to 15 years.

IQT Pipeline. In the training stage, we randomly selected 30 subjects with skull-stripping from HCP dataset and employed them to synthesise the low-field images using Alg. 1 based on *a priori* variable SNRs. Regarding patch extraction, we cropped the low-field patches with the step size of 8, 16, and $16/k$ along x -, y -, and z -directions, respectively. We also cropped the high-field patches with the same volume and position as the corresponding low-field patches. The low-field and high-field patch sizes were $32 \times 32 \times (32/k)$ and $32 \times 32 \times 32$, respectively. Then the patches capturing 80% background voxels were excluded from a patch library.

We examined if overfitting occurred with a validation set and judged the performance of the trained neural network with an evaluation set. We split all 30 subjects into 12, 3, and 15 for training, validation, and evaluation sets. Moreover, we investigated the image quality by calculating the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [18]. We employed a two-tailed Wilcoxon signed-rank test to determine the statistical significance of the performance difference between two comparing methods.

Neural Networks. We conducted an ablation study on the proposed U-Net, denoted by ANISO U-Net(b), in the case of $b = 2$ or 3 for shrinking layers in the bottleneck block. We evaluated our networks against the 3D cubic B-spline interpolation and several existing U-Net baselines equivalently switching off the corresponding blocks, i.e. bottleneck block and residual core, in ANISO U-Net. One is an isotropic 3D U-Net (ISO U-Net) [14] implemented with 5 levels and 3 convolutional layers per level. The input of ISO U-Net is isotropically interpolated using cubic B-splines. The other one is 3D-SRU-Net [13] that up-samples each level output on the contraction path before concatenation. It contained 3 levels for the down-sampling scale $k = 4$ and 4 levels for $k = 8$. We unified hyper-parameters of the three U-Nets as follows. Number of filters on the first level was 16 with the number of filters doubling at each subsequent level. All U-Nets were implemented in Python using Keras library [19] with Tensorflow backend. They were calculated on a Nvidia GTX 1080 Ti GPU. Training used ADAM [20] as the optimizer with a starting learning rate of 10^{-3} and a decay of 10^{-6} . We initialized the parameters with Glorot normal initializer [21]. The batchsize was 32 and the loss function is the pixel-wise mean squared error (MSE). All the experiments started converging after about 30 epochs and we employed early stopping after 5 epochs of no improvement on the validation set.

Table 1. The performance of the proposed model on up-scaling factors $k = 4$ or 8 . The mean and standard deviation of PSNR and the mean SSIM (MSSIM) are calculated over 15 evaluation subjects. For each case, we show the best performance over an ensemble of 5 trained models. Bold font denotes the best mean or standard deviation. The asterisk * denotes p -value < 0.01 compared with the rest methods.

Method	$k = 4$		$k = 8$	
	PSNR (dB)	MSSIM	PSNR (dB)	MSSIM
Cubic B-spline	$20.689 \pm 2.540^*$	$0.692 \pm 0.0384^*$	$18.974 \pm 2.535^*$	$0.567 \pm 0.0471^*$
ISO U-Net	30.798 ± 2.573	$0.916 \pm 0.0227^*$	27.073 ± 2.469	0.846 ± 0.0278
3D-SRU-Net	30.764 ± 2.638	0.922 ± 0.0191	27.275 ± 2.542	0.847 ± 0.0290
ANISO U-Net(2)	$31.045 \pm 2.654^*$	0.923 ± 0.0197	27.346 ± 2.517	$0.852 \pm 0.0280^*$
ANISO U-Net(3)	30.918 ± 2.639	$0.921 \pm 0.0199^*$	27.054 ± 2.544	0.847 ± 0.0282

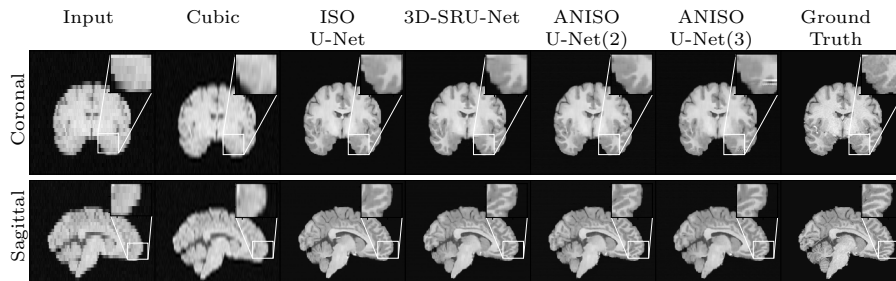


Fig. 3. Visualization of U-Net reconstructions with the up-scaling factor $k = 8$.

3.2 Evaluation on Fixed SNR Data Sets

We evaluated the ability of the proposed U-Net in an ideal case where the SNR-related coefficient $\alpha = (SNR_X^{WM}, SNR_X^{GM})$ in Eq. (1) is deterministic. We fixed the SNR_X^{WM} and SNR_X^{GM} as 61 and 53, respectively, in the IQT pipeline by reconstructing images in the evaluation set at Step 4 in Algorithm 1. Table 1 shows that our model, ANISO U-Net(2), achieved the best performance in terms of the average PSNR and SSIM, and especially, significantly outperformed the others in terms of PSNR at $k = 4$ and the mean SSIM (MSSIM) at $k = 8$. The reconstruction degraded as the up-scaling factor increased. Figure 3 shows the U-Net reconstructions on coronal and sagittal planes. Qualitatively we observed clear recovery of high resolution information and enhancement of contrast. The reconstructed images from all networks nicely highlighted features visible in the ground truth images that were obscured in the low quality input. The quantitative results in Table 1 show little difference among the U-Net outputs but they might not be able to reflect subtle qualitative differences. The zoomed patches in Fig. 3 highlight differences more clearly and we believe ANISO U-Net(2) approximates the ground truth most closely and with the least artefacts as shown in the ANISO U-Net(3) result of Fig. 3. Delicately selecting hyper-parameters can avoid overfitting, and hence can mitigate the artifacts.

3.3 Evaluation on Variable SNR Data Sets

We evaluated the performance of several deep learning architectures including the proposed anisotropic U-Net with variable-SNR low-field data. SNR_X^{WM} and SNR_X^{GM} are now sampled from a two-dimensional Gaussian distribution $\mathcal{N}(\mu, \Sigma)$ where the coefficients are:

$$\mu = (64.50, 54.14) \quad \Sigma = \begin{pmatrix} 78.47, & 71.50 \\ 71.50, & 73.91 \end{pmatrix}.$$

The simulator shown in Alg. 1 randomly generated N low-field input images with different SNR for the chosen 15 training subjects in the HCP data set. We trained the deep learning models on the dataset with the augmenting factor $N = 1, 2, 4$ and 8. We randomly selected 12.5% patches for training in each overlap patch library. For each neural network, an ensemble of 5 models were trained in terms of different augmented dataset.

Table 2 shows the mean and standard deviation of PSNR and MSSIM over 15 test subjects in terms of the augmented datasets and deep learning architectures. As a result, probabilistic decimation model was generally able to produce more stable reconstruction than the deterministic model if the unseen test data were also generated from the variable SNR. Both accuracy and robustness corresponding to mean and standard deviation of MSSIM improved in various degree as the number of generated low-field image samples increased, and in addition, the performances for the two methods were statistically significant in terms of $N = 8$ at $k = 8$. Regarding PSNR, the performance upgraded after augmentation but the robustness reflected by the standard deviation did not improve correspondingly. In addition, we observed that PSNR and MSSIM at $k = 4$ only slightly improve when the augmenting factor N became larger, which means the improvement of performance arising from augmentation gradually reached an upper bound.

3.4 Test on Patient Data

We tested our IQT approach on the data from a 10-year-old epilepsy patient who has two cortical-subcortical cystic lesions with surrounding edema on low-field T1-weighted images at the GM-WM junction of the parietal lobes. In this case, we used IQT with ANISO U-Net(2) trained on the HCP dataset with the augmenting factor $N = 1$ and the up-scaling factor of $k = 4$. Figure 4 shows the axial and coronal results enhanced from the low-field T1-weighted image of the patient. The IQT approach improved the GM-WM contrast globally, and significantly enhanced the resolution in coronal and sagittal planes. The enhanced image strongly highlights the two lesions in this patient which are very subtle on the input T1-weighted image. In this particular patient, the lesions were clearly visible on the original T2-weighted image, which validates that IQT highlights the lesions in the correct locations, as Fig. 4(c) shows. However, in general not all lesions are clearly visible on any MRI sequence, especially at low field, and Fig. 4 highlights the potential of our algorithms to reveal subtle

Table 2. The performance of probabilistic decimation simulation for augmentation with a factor of N . The mean and standard deviation of PSNR and MSSIM are calculated over 15 evaluation subjects. We show the best performance over an ensemble of 5 trained models. The “const” at N samples/subject column means that the models were trained on the fixed SNR data sets as described in Section 3.2. Bold font denotes the best mean or standard deviation. The asterisk * denotes p -value < 0.01 compared with the other augmentation factors.

Method	N samples /subject	$k = 4$		$k = 8$	
		PSNR (dB)	MSSIM	PSNR (dB)	MSSIM
3D-SR U-Net	“const”	27.214 ± 3.030	0.871 ± 0.0332	24.687 ± 2.464	0.758 ± 0.0371
	1	27.988 ± 2.445	0.861 ± 0.0286	23.777 ± 2.693	0.757 ± 0.0401
	2	29.453 ± 2.585	0.901 ± 0.0240*	25.513 ± 2.711*	0.799 ± 0.0366*
	4	30.257 ± 2.647	0.918 ± 0.0201	26.025 ± 2.587	0.816 ± 0.0339*
	8	29.958 ± 2.541	0.911 ± 0.0203	26.391 ± 2.621	0.832 ± 0.0316*
ANISO U-Net(2)	“const”	27.311 ± 3.522	0.870 ± 0.0338	24.754 ± 2.367	0.769 ± 0.0341
	1	28.664 ± 2.552	0.890 ± 0.0240	23.418 ± 2.322	0.757 ± 0.0345*
	2	29.216 ± 2.308	0.893 ± 0.0248	25.862 ± 2.617*	0.803 ± 0.0343
	4	30.248 ± 2.565	0.916 ± 0.0195	26.231 ± 2.613	0.807 ± 0.0381
	8	30.344 ± 2.421	0.914 ± 0.0188	27.053 ± 2.398	0.843 ± 0.0330*

lesions enhancing diagnosis and potentially enabling effective treatment via clear localisation.

4 Discussion and Conclusion

In this work, we present an IQT approach to enhance low-field MRIs aiming to match resolution as well as contrast of high-field images. We introduce the anisotropic U-Net characterised by a deeper hierarchy and super resolving connections between input and output layers. We propose the probabilistic decimation simulator by synthesising multiple low-field images with respect to distinct grey-white matter SNR sampled from an *a priori* distribution. We demonstrate that the proposed method improves the robustness on the unseen test data of variable SNR at the evaluation stage. We validate our proposed U-Net on the evaluation dataset and the results potentially show generalisability to the actual clinical low-field images.

This work offers several avenues for future improvement and application. Here the metrics (MSSIM and PSNR) used for quantitative assessment reflect the performance on only synthetic images. This demonstrates efficacy, but evaluation on a sizeable data set of clinical images and clinical significance from radiologists are essential for further translation. Therefore, additional qualitative evaluation by radiologist ratings and, ultimately, demonstration of improved decision making is essential to confirm impact of the approach. Nevertheless, we believe our methods have great potential to identify subtle lesions in epilepsy and other neurological conditions and thus to improve patient outcomes in LMICs in the future.

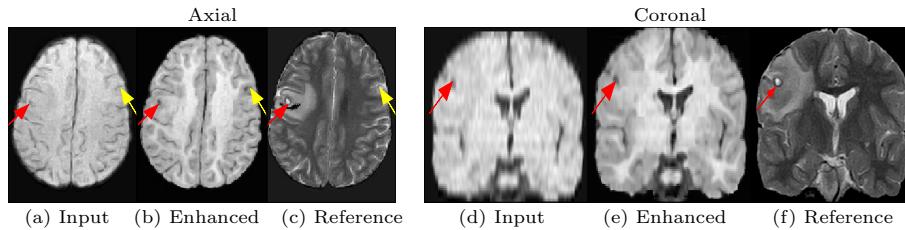


Fig. 4. The IQT prediction on the low-field epileptic patient data for (a-c) axial plane and (d-f) coronal plane. (a) and (d): Low-field T1-weighted input with cubic B-spline interpolation; (b) and (e): IQT-enhanced T1-weighted output using ANISO U-Net(2); (c) and (f): low-field T2-weighted image as a reference of ground truth. Two sub-centimeter parenchymal cystic lesions at the GM-WM junction of the parietal lobes are pointed out by the red and the yellow arrows. They are barely visible in (a) and (d) but greatly enhanced in (b) and (e). (c) and (f), not involved in the IQT experiment, verified their location in an independent acquisition.

Acknowledgements

This work was supported by EPSRC grants (EP/R014019/1, EP/R006032/1 and EP/M020533/1) and the NIHR UCLH Biomedical Research Centre. Data were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by NIH and Washington University. The 0.36T MRI data were acquired at the University College Hospital, Ibadan, Nigeria.

References

1. Wadghiri, Y.Z., Johnson, G., Turnbull, D.H.: Sensitivity and performance time in MRI dephasing artifact reduction methods. *Magnetic Resonance in Medicine*. **45**(3), 470–476 (2001)
2. Sotiropoulos, S.N., et al.: Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *NeuroImage*. **80**, 125–143 (2013)
3. Marques, J.P., Simonis, F.F.J., Webb, A.G.: Lowfield MRI: An MR physics perspective. *Journal of Magnetic Resonance Imaging*. **49**(6), 1528–1542 (2019)
4. Brown, R.W., Cheng, Y.-C.N., Haacke, E.M., Thompson, M.R., Venkatesan, R.: *Magnetic resonance imaging: physical principles and sequence design*. 2nd edn. John Wiley & Sons, Inc., Hoboken, New Jersey (2014)
5. Bahrami, K., Shi, F., Zong, X., Shin, H.W., An, H., Shen, D.: Reconstruction of 7T-Like Images from 3T MRI. *IEEE Transactions on Medical Imaging*. **35**(9), 2085–2097 (2016)
6. Wolterink, J.M., Dinkla, A.M., Savenije, M.H.F., Seevinck, P.R., van den Berg, C.A.T., Išgum, I.: Deep MR to CT Synthesis Using Unpaired Data. In: *SASHIMI 2017*. LNCS 10557, pp 14–23. Springer, Cham (2017).
7. Cohen, J.P., Luck, M., Honari, S.: Distribution Matching Losses Can Hallucinate Features in Medical Image Translation. In: *MICCAI 2018*. LNCS 11070, pp 529–536. Springer, Cham (2018)

8. Alexander, D.C., Zikic, D., Ghosh, A., Tanno, R., Wottschel, V., Zhang, J., Kaden, E., Dyrby, T.B., Sotiropoulos, S.N., Zhang, H., Criminisi, A.: Image quality transfer and applications in diffusion MRI. *NeuroImage*. **152**, 283–298 (2017)
9. Tanno, R., Ghosh, A., Grussu, F., Kaden, E., Criminisi, A., Alexander, D.C.: Bayesian Image Quality Transfer. In: *MICCAI 2016*. LNCS 9901, pp 265–273. Springer, Cham (2016)
10. Tanno, R., Worrall, D.E., Ghosh, A., Kaden, E., Sotiropoulos, S.N., Criminisi, A., Alexander, D.C.: Bayesian image quality transfer with CNNs: Exploring uncertainty in dMRI super-resolution. In: *MICCAI 2017*, LNCS 10433, pp. 611–619. Springer, Cham (2017)
11. Blumberg, S.B., Tanno, R., Kokkinos, I., Alexander, D.C.: Deeper Image Quality Transfer: Training Low-Memory Neural Networks for 3D Images. In: *MICCAI 2018*. LNCS 11070, pp. 118–125. Springer, Cham (2018)
12. Ashburner, J., Friston, K.J.: Unified segmentation. *NeuroImage*. **26**, 839-851 (2005)
13. Heinrich, L., Bogovic, J.A., Saalfeld, S.: Deep Learning for Isotropic Super-Resolution from Non-isotropic 3D Electron Microscopy. In: *MICCAI 2017*, LNCS 10434, pp. 135–143. Springer, Cham (2017)
14. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In: *MICCAI 2016*, LNCS 9901, pp. 424–432 (2016)
15. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. In *CVPR 2016*, pp. 770–778 (2016)
16. Dong, C., Loy, C.C., Tang, X.: Accelerating the Super-Resolution Convolutional Neural Network. In: *ECCV 2016*, LNCS 9906, pp.391-407. Springer, Cham (2016)
17. Guerrero, R., Qin, C., Oktay, O., Bowles, C., Chen, L., Joules, R., Wolz, R., Valdes-Hernandez, M.C., Dickie, D.A., Wardlaw, J., Rueckert, D.: White matter hyperintensity and stroke lesion segmentation and differentiation using convolutional neural networks. *NeuroImage: Clinical*. **17**, 918-934 (2018)
18. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*. **13**(4), 600-612 (2004)
19. Chollet, F. et al.: Keras. <https://keras.io> (2015)
20. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In *arXiv preprint arXiv:1412.6980* (2014)
21. Glorot, X. and Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. In: *AISTATS 2010*, PMLR **9**, pp. 249-256 (2010)