Medical Image Analysis 80 (2022) 102509

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

⊥-loss: A symmetric loss function for magnetic resonance imaging reconstruction and image registration with deep learning

Maarten L. Terpstra[a,b,*], Matteo Maspero[a,b], Alessandro Sbrizzi[a,b], Cornelis A.T. van den Berg[a,b]

[a] Department of Radiotherapy, University Medical Center Utrecht, Heidelberglaan 100, Utrecht, 3508GA, the Netherlands
[b] Computational Imaging Group for MR Diagnostics & Therapy, University Medical Center Utrecht, Heidelberglaan 100, Utrecht, 3508GA, the Netherlands

ARTICLE INFO

Article history:
Received 20 December 2021
Revised 8 April 2022
Accepted 1 June 2022
Available online 2 June 2022

Keywords:
Loss function
Deep learning
Artificial intelligence
Optimization
Symmetry
Complex number
MRI
Image reconstruction
Motion estimation

ABSTRACT

Convolutional neural networks (CNNs) are increasingly adopted in medical imaging, e.g., to reconstruct high-quality images from undersampled magnetic resonance imaging (MRI) acquisitions or estimate subject motion during an examination. MRI is naturally acquired in the complex domain $\mathbb{C}$, obtaining magnitude and phase information in k-space. However, CNNs in complex regression tasks are almost exclusively trained to minimize the L2 loss or maximizing the magnitude structural similarity (SSIM), which are possibly not optimal as they do not take full advantage of the magnitude *and* phase information present in the complex domain. This work identifies that minimizing the L2 loss in the complex field has an asymmetry in the magnitude/phase loss landscape and is biased, underestimating the reconstructed magnitude. To resolve this, we propose a new loss function for regression in the complex domain called ⊥-loss, which adds a novel phase term to established magnitude loss functions, e.g., L2 or SSIM. We show ⊥-loss is symmetric in the magnitude/phase domain and has favourable properties when applied to regression in the complex domain. Specifically, we evaluate the $\perp+\ell^2$-loss and ⊥+SSIM-loss for complex undersampled MR image reconstruction tasks and MR image registration tasks. We show that training a model to minimize the $\perp+\ell^2$-loss outperforms models trained to minimize the L2 loss and results in similar performance compared to models trained to maximize the magnitude SSIM while offering high-quality phase reconstruction. Moreover, ⊥-loss is defined in $\mathbb{R}^n$, and we apply the loss function to the $\mathbb{R}^2$ domain by learning 2D deformation vector fields for image registration. We show that a model trained to minimize the $\perp+\ell^2$-loss outperforms models trained to minimize the end-point error loss.

1. Introduction

Magnetic resonance imaging (MRI) is a noninvasive imaging technique to obtain an anatomical image with high resolution and excellent soft-tissue contrast. These properties have made MRI an indispensable diagnostic tool and is increasingly used for interventional guidance, such as high-intensity focused ultrasound (Wijlemans et al., 2012), catheter guidance during surgery (Rogers et al., 2014), and radiotherapy (Raaymakers et al., 2009; Mutic and Dempsey, 2014).

MRI must be acquired, reconstructed, and processed with high accuracy in real-time for interventional guidance applications. However, MRI acquisition can be time-consuming due to hardware limitations that ensure patient safety, e.g., low gradient slew rate to avoid peripheral nerve stimulation (Zhang et al., 2003) or limited radio-frequency power to prevent patient heating (Collins and Wang, 2011). However, the limited shot-encoding power of MR acquisitions is the most time-consuming part of image formation (Wright, 1997), resulting in repeated sampling of the frequency domain (k-space) to fulfil the Nyquist sampling criterion (Shannon, 1948). These constraints limit patient throughput and prohibit MRI applications that require high spatio-temporal resolution, e.g., cardiac imaging (Bustin et al., 2020), speech imaging (Lingala et al., 2016), or tracking abdominal motion (Keiper et al., 2020). One way to accelerate MRI is to acquire fewer data, i.e., undersampling the frequency domain as the number of k-space samples determines acquisition time. However, undersampling violates the Nyquist criterion, introducing image artifacts. Several methods have been proposed to remove these undersampling artifacts. For example, parallel imaging (Pruessmann

* Corresponding author at: Department of Radiotherapy, University Medical Center Utrecht, Heidelberglaan 100, Utrecht 3508GA, the Netherlands.
E-mail address: m.l.terpstra-5@umcutrecht.nl (M.L. Terpstra).

https://doi.org/10.1016/j.media.2022.102509

et al., 1999; Griswold et al., 2002) exploits information redundancy using multiple receiver elements. Also, iterative reconstruction algorithms as compressed sensing have been proposed (Lustig et al., 2007), which resolve image artifacts by casting MR reconstruction as a sparse denoising problem. However, the resulting acceleration factor of parallel imaging remains limited (Wiesinger et al., 2004), e.g., up to four-fold undersampling, and compressed sensing reconstructions cause a significant reconstruction latency, precluding real-time applications.

Recently, machine learning has been proposed as an alternative to traditional methods to solve inverse problems (Adler and Öktem, 2017). Specifically, several methods have been proposed to use convolutional neural networks (CNNs) to accelerate undersampled MRI reconstruction (Hammernik et al., 2018; Schlemper et al., 2018), perform organ segmentation (Fu et al., 2021), or estimate motion from undersampled MRI (Terpstra et al., 2021). These methods have attractive properties compared to traditional methods, such as the ability to obtain high-quality MRI reconstructions with high undersampling factors (eight-fold or higher), reducing acquisition time. Moreover, CNNs exhibit low inference times by taking advantage of parallel GPU architectures, even though training of CNNs can take hours or days. Using CNNs to accelerate the acquisition, reconstruction, and processing of MRI could enable new applications such as real-time interventional guidance using MRI (Jaubert et al., 2021) or real-time adaptive MRI-guided radiotherapy (Keall et al., 2019a).

In this work, we consider complex regression with CNNs for MRI applications. These CNNs are trained by finding parameters that minimize a loss function over a training set. The quality of the estimated solution largely depends on the loss function, as these loss functions determine the impact of residual artifacts and the importance of specific (image) features, such as texture, contrast, or the effects of noise propagation. It has been shown that loss functions with desirable properties, such as monotonicity, smoothness (Kanai et al., 2021), or symmetry (Patel and Sastry, 2021; Kunin et al., 2021), can lead to better generalization, noise robustness, and faster convergence (Janocha and Czarnecki, 2017), depending on the task at hand. As MRI is an inherently complex signal, loss functions for reconstructing or processing MRI are naturally defined in the complex domain. For example, for image reconstruction, popular choices of loss functions between a complex target image $\mathbf{Y} \in \mathbb{C}^{m \times n}$ and a complex estimated image $\hat{\mathbf{Y}} \in \mathbb{C}^{m \times n}$ include minimizing the complex difference using the $\ell^p_{\mathbb{C}}$-norm, e.g., $\ell^2_{\mathbb{C}}(\mathbf{Y}, \hat{\mathbf{Y}}) = ||\Re(\mathbf{Y} - \hat{\mathbf{Y}})||^2 + ||\Im(\mathbf{Y} - \hat{\mathbf{Y}})||^2$ or $\ell^1_{\mathbb{C}}(\mathbf{Y}, \hat{\mathbf{Y}}) = ||\Re(\mathbf{Y} - \hat{\mathbf{Y}})||_1 + ||\Im(\mathbf{Y} - \hat{\mathbf{Y}})||_1$, where $||\cdot||^2$ is the squared Frobenius norm, $||\cdot||_1$ is the ℓ_1 norm, $\Re(x)$ is the real part of x, and $\Im(x)$ is the imaginary part of x, or $\ell_{\text{SSIM}} = 1 - \text{SSIM}(|\mathbf{Y}|, |\hat{\mathbf{Y}}|)$, where SSIM is the structural similarity index measure (Wang et al., 2004). For image registration, the most common loss function between two deformation vector fields (DVFs) is the end-point error (EPE), which is equivalent to $\ell^2_{\mathbb{C}}$ in the $\mathbb{R}^2$ domain (Butler et al., 2012; Dosovitskiy et al., 2015). The geometric interpretation of minimizing the $\ell^2_{\mathbb{C}}$-norm is illustrated in Fig. 1.

These loss functions do not take full advantage of the phase structure of the data; for example, ℓ_{SSIM} discards phase information, while it has been shown that using this information could improve image reconstruction performance (Haji-Valizadeh et al., 2021). On the other hand, the $\ell^p_{\mathbb{C}}$ is separately defined on the real and imaginary components of the complex-valued signal, possibly precluding the $\ell^p_{\mathbb{C}}$ from taking full advantage of the magnitude and phase properties. Moreover, these loss functions do not necessarily produce a symmetrically-distributed error, while it has been demonstrated that using symmetric loss functions can improve task performance (Patel and Sastry, 2021). Specifically, a symmetrically-distributed error when estimating motion for image registration is a desirable property as the registration error is equally distributed in all directions.

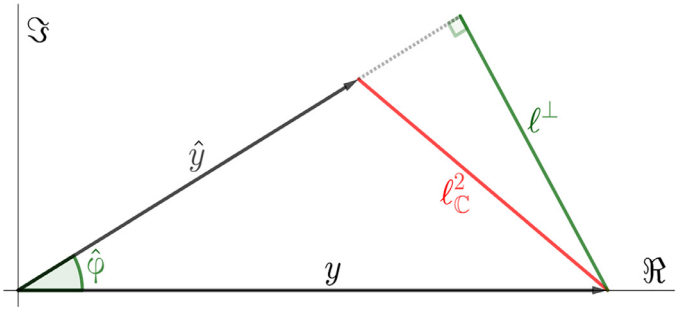

Fig. 1. Geometric illustration of regression in the $\mathbb{R}^2$ domain. An estimated vector $\hat{Y}$ compared to a target vector Y, with magnitude ratio $\lambda = |\hat{Y}|/|Y|$ and phase difference $\hat{\varphi}$. The $\ell^2_{\mathbb{C}}$ loss is the magnitude of the red vector. Our proposed loss function $\perp$-loss determines the phase error $\hat{\varphi}$ as the scalar rejection $\ell^{\perp}$, i.e., the length of the perpendicular line from $\hat{Y}$ to Y. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We hypothesize that a new loss function based on a complex signal's magnitude and phase components could lead to improved regression in the complex domain using deep learning models, thus improving image quality when reconstructing undersampled MRI and leading to better image registration. Therefore, we propose $\perp$-loss, a symmetric loss function defined in the magnitude and phase domain. In this work, we will:

1. Analyze the behavior of $\ell^p_{\mathbb{C}}$ in the complex plane and investigate the symmetry of the loss landscape.
2. Introduce the $\perp$-loss function, which operates on the polar representation of complex numbers and adds a novel phase term to magnitude loss functions. We examine the loss landscape produced by this loss function, comparing it to $\ell^p_{\mathbb{C}}$.
3. Study the performance of the $\ell^2_{\mathbb{C}}$ loss, ℓ_{SSIM}, $\perp$+ℓ^2-loss, and $\perp$+SSIM-loss functions for undersampled MRI reconstruction tasks using deep neural networks.
4. Explore the application of the $\perp$-loss to other domains, considering regression in the $\mathbb{R}^2$ domain for 2D image registration.

2. Theory

This section examines the topology of the $\ell^p_{\mathbb{C}}$ loss landscape and proposes a new loss function. We consider a complex image $\mathbf{Y} \in \mathbb{C}^{m \times n}$, representing the signal of each voxel as $Y_{jk} = a + bi$ where $a, b \in \mathbb{R}$ are the real and imaginary components (Cartesian representation), or $Y_{jk} = |Y_{jk}|e^{i\varphi_{jk}}$ where $|Y_{jk}|$ is the *magnitude* and φ_{jk} as the *phase* (polar representation).

2.1. Proof of asymmetry of $\ell^2_{\mathbb{C}}$

It has been shown that least-squares regression in the $\mathbb{R}^2$-space in the presence of noise results is biased. In particular, performing least-squares regression on noisy data results in parameters that underestimate the recovered magnitude (Fermüller et al., 2001; Fuller, 2009; Gleser, 1981; Nagel and Haag, 1998; Buczko and Willert, 2017). We make the case that this bias is also present for regression tasks in the complex domain.

Suppose we wish to estimate a complex value, e.g., estimating a voxel $\hat{Y}$ against the ground-truth voxel Y in the target image $\mathbf{Y}$. The reconstructed voxel $\hat{Y}$ is typically estimated by regression, minimizing a loss function such as the complex extension of the ℓ^2-norm, i.e., $\ell^2_{\mathbb{C}}(Y, \hat{Y}) = ||\Re(Y - \hat{Y})||^2 + ||\Im(Y - \hat{Y})||^2$. The reconstructed magnitude bias is expressed as $|\hat{Y}| = \lambda|Y|$ with $\lambda > 0$, and the remaining phase error is denoted as $\hat{\varphi}$.

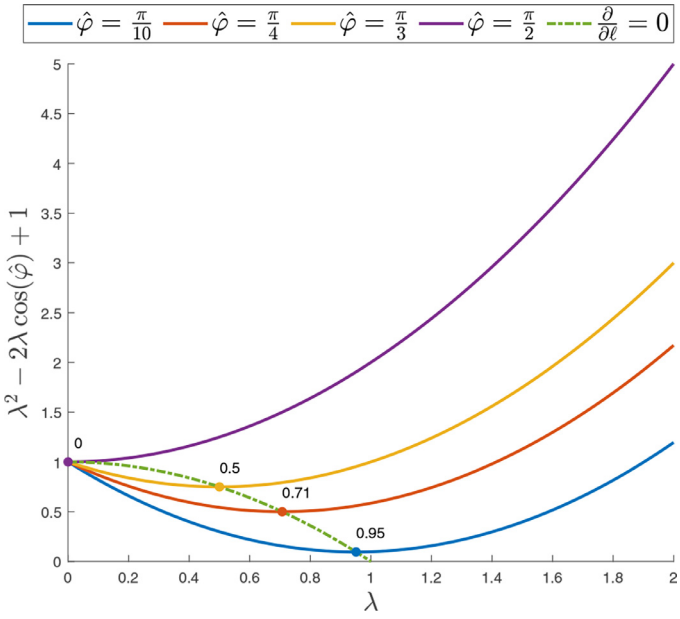

Fig. 2. The loss of $\ell^2_{\mathbb{C}}$. The parabola of the loss for a given value of $\hat{\varphi}$ is shown as a function of λ. The minimal value is indicated in every line plot, and the analytical curve of minimal loss is shown in green. It is clear that as $\hat{\varphi}$ increases, the minimum of the loss function goes to $\lambda \to 0$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In Fig. 1, this estimation is illustrated using vectors in the complex plane, where the $\ell^2_{\mathbb{C}}$ error, $\hat{\boldsymbol{Y}}$, and $\boldsymbol{Y}$ form a triangle. Therefore $\ell^2_{\mathbb{C}}$ can be expressed using the law of cosines as

$$\ell^2_{\mathbb{C}} = |\hat{\boldsymbol{Y}}|^2 + |\boldsymbol{Y}|^2 - 2|\hat{\boldsymbol{Y}}||\boldsymbol{Y}|\cos\hat{\varphi} \to \ell^2_{\mathbb{C}} \propto \lambda^2 - 2\lambda\cos\hat{\varphi} + 1.$$

This expression is zero if and only if $\boldsymbol{Y}$ and $\hat{\boldsymbol{Y}}$ are identical, i.e., $\lambda = 1$ and $\hat{\varphi} = 0$. We can observe that the $\ell^2_{\mathbb{C}}$ follows a parabola, depending on the value of λ and $\cos\hat{\varphi}$. For $\hat{\varphi} = 0$, this function is minimal when $\lambda = 1$. However, the minimum of the curve when $\hat{\varphi} \neq 0$ is at $\lambda < 1$ and tends toward $\lambda \to 0$ as $|\hat{\varphi}|$ increases, as shown in Fig. 2. This shift depending on $\hat{\varphi}$ implies that $\ell^2_{\mathbb{C}}$ underestimates the reconstructed magnitude, assigning a lower loss to estimates with $\lambda < 1$ than to estimates with $\lambda > 1$.

We can show that $\ell^2_{\mathbb{C}}$ is biased towards reconstructions with $\lambda \leq 1$ by examining the gradient of $\ell^2_{\mathbb{C}}$. The loss of $\ell^2_{\mathbb{C}}$ is minimal when $\frac{\ell^2_{\mathbb{C}}}{\partial\lambda} = 0$. Taking the derivative of $\ell^2_{\mathbb{C}}$ for λ yields

$$\frac{\partial \ell^2_{\mathbb{C}}}{\partial \lambda} = 0 \to 2\lambda - 2\cos\hat{\varphi} = 0$$
$$2\lambda = 2\cos\hat{\varphi}$$
$$\lambda = \cos\hat{\varphi}.$$

As $\cos\hat{\varphi} \leq 1$ for every value of $\hat{\varphi}$, the loss function is minimal when $\lambda \leq 1$, and $\lambda = 1$ if and only if $\hat{\varphi} = 0 \pmod{2\pi}$.

This analysis proves that using the $\ell^2_{\mathbb{C}}$ loss function to reconstruct $\hat{\mathbf{Y}}$ typically favors reconstructions where $|\hat{Y}| \leq |Y|$.

2.2. Proposed solution

We hypothesize that this bias occurs because the angle and magnitude errors are minimized simultaneously by manipulating the real and imaginary components. We introduce a possible solution by proposing a new loss function that separates the loss in angle and magnitude losses.

We propose representing the phase error between $\hat{\boldsymbol{Y}}$ and $\boldsymbol{Y}$ as $\ell^{\perp}$, the length of the line segment which is perpendicular to $\hat{\boldsymbol{Y}}$, ending at $\boldsymbol{Y}$, as visualized in Fig. 1 as the green line $\ell^{\perp}(\hat{\boldsymbol{Y}}, \boldsymbol{Y})$. More specifically, $\ell^{\perp}$ is the scalar rejection $\ell^{\perp} = ||\boldsymbol{Y} - \text{proj}_{\hat{\boldsymbol{Y}}}\boldsymbol{Y}||$. This segment is independent of the magnitude error λ and computed as

$$\ell^{\perp}(\boldsymbol{Y}, \hat{\boldsymbol{Y}}) = \frac{\boldsymbol{Y} \cdot \hat{\boldsymbol{Y}}^T}{|\hat{\boldsymbol{Y}}|}. \tag{1}$$

In the complex plane, this function is efficiently computed (Heckbert, 1994) as

$$\ell^{\perp}(Y, \hat{Y}) = \frac{|\Re(\hat{Y}) \cdot \Im(Y) - \Im(\hat{Y}) \cdot \Re(Y)|}{|\hat{Y}|}. \tag{2}$$

The magnitude error can then be independently represented from $\ell^{\perp}$ in the full loss function.

The definition of $\ell^{\perp}$ assumes $\hat{\varphi} \leq \pi/2$ and $\ell^{\perp}$ decreases when $\hat{\varphi}$ increases beyond $\pi/2$. To ensure $\ell^{\perp}$ smoothly increases when $\hat{\varphi} > \pi/2$, a smooth continuation of $\ell^{\perp}$ when $\hat{\varphi} \geq \pi/2$ has been defined in Eq. 4. Moreover, it was assumed that $\hat{Y}$ and Y have nonzero magnitude, which is not always the case. Therefore, we add $\epsilon = 10^{-8}$ to the denominator of Eq. 2 during implementation as defined in Eq. 3.

The complete loss function for complex images $\mathbf{Y}, \hat{\mathbf{Y}} \in \mathbb{C}^{m \times n}$ is then defined as

$$\ell^{\perp}(\hat{\mathbf{Y}}, \mathbf{Y}) = \frac{1}{mn}\sum_{j}^{m}\sum_{k}^{n} \frac{|\Re(\hat{Y}_{jk})\Im(Y_{jk}) - \Im(\hat{Y}_{jk})\Re(Y_{jk})|}{|\hat{Y}_{jk}| + \epsilon} \tag{3}$$

$$\mathcal{L}^{\perp}(\hat{\mathbf{Y}}, \mathbf{Y}) = \begin{cases} \ell^{\perp}(\hat{\mathbf{Y}}, \mathbf{Y}) & |\hat{\varphi}| < \frac{\pi}{2} \\ 2|\mathbf{Y}| - \ell^{\perp}(\hat{\mathbf{Y}}, \mathbf{Y}) & |\hat{\varphi}| \geq \frac{\pi}{2} \end{cases} \tag{4}$$

$$\perp(\hat{\mathbf{Y}}, \mathbf{Y}) = \mathcal{L}^{\perp}(\hat{\mathbf{Y}}, \mathbf{Y}) + f(|\mathbf{Y}|, |\hat{\mathbf{Y}}|) \tag{5}$$

In Eq. 5, f is the loss function operating on the magnitude part of $\hat{\mathbf{Y}}$ and $\mathbf{Y}$. For example, we can define $\perp$+ℓ^2-loss where f is the Frobenius norm of $|\hat{\mathbf{Y}}| - |\mathbf{Y}|$, or $\perp$+SSIM-loss where $f = 1 - \text{SSIM}(|\hat{\mathbf{Y}}|, |\mathbf{Y}|)$. As $\ell^{\perp}$ is independent of λ and the magnitude term is independent of $\hat{\varphi}$, $\perp$-loss is symmetric and assigns equal loss to vectors with magnitude bias λ and λ^{-1} for the same $\hat{\varphi}$.

$\perp$-loss as defined in Eq. 5 is currently only defined in the complex domain but can also be applied to different domains. In particular, as $\mathbb{C}$ is isomorphic to $\mathbb{R}^2$, $\perp$-loss may find application in many regression tasks, e.g., image registration (Buczko and Willert, 2017; Terpstra et al., 2020) or fluid flow analysis (Kim and Günther, 2019). Moreover, the concept of the scalar rejection that forms the basis of $\ell^{\perp}$ in Eq. 1 allows for direct extension to $\mathbb{R}^n$, enabling the application of $\perp$-loss to higher-dimensional problems.

3. Methods

First, we examine the symmetry of the loss landscapes of the $\ell^1_{\mathbb{C}}$, $\ell^2_{\mathbb{C}}$ and $\perp$+ℓ^2-loss functions. Subsequently, to study the image reconstruction quality and dataset or model dependence, we apply $\perp$+ℓ^2-loss and $\perp$+SSIM-loss to complex image reconstruction using two different networks and two different datasets. Finally, we study how $\perp$+ℓ^2-loss generalizes to problems defined in the $\mathbb{R}^2$ vector space, focusing on image registration.[1]

3.1. Loss landscape examination

To verify the symmetry of the loss functions, we have performed a simulation to visualize the loss landscape of $\ell^1_{\mathbb{C}}$, $\ell^2_{\mathbb{C}}$, and $\perp$+ℓ^2-loss loss functions. We generated a vector of $n = 5000$ complex numbers $x \in \mathbb{C}^n$, $x_j = a + bi$ with a and b independently drawn from a uniform distribution between -5 and

[1] Code will be made publicly available on https://gitlab.com/computational-imaging-lab/perp_loss.

5. Next, we perturbed every x_j with every $(\lambda, \hat{\varphi})$-pair generated by the Cartesian product between $\Lambda = \{0, 0.02, 0.04, \ldots, 2\}$ and $\Phi = \{0, \frac{\pi}{100}, \frac{2\pi}{100}, \frac{3\pi}{100}, \ldots, \pi\}$. That is, $\hat{x}_k = \lambda_k x_j e^{i\hat{\varphi}_k}$, $(\lambda_k, \hat{\varphi}_k) \in \Lambda \times \Phi$. This resulted in a vector $\hat{x}$ with 10,000 perturbed complex numbers for every x_j. Finally, we computed the $\ell^1_{\mathbb{C}}$, $\ell^2_{\mathbb{C}}$ and $\perp$+ℓ^2-loss loss between x_j and $\hat{x}$ to obtain a loss value per $(\lambda, \hat{\varphi})$-pair. Calculating this loss for every x_j and taking the mean resulted in the loss landscape of λ and $\hat{\varphi}$. These landscapes were compared for the $\ell^1_{\mathbb{C}}$, $\ell^2_{\mathbb{C}}$ and the $\perp$+ℓ^2-loss loss.

3.2. MRI reconstruction

We have trained two deep CNNs on two different datasets to examine the performance of $\perp$-loss when optimizing deep neural networks for undersampled MRI reconstruction. An "end-to-end variational network" (E2E-VarNet) (Sriram et al., 2020) was trained to reconstruct complex MRI from the fastMRI challenge (Zbontar et al., 2018) (Experiment A). Also, a recurrent inference machine (Putzky and Welling, 2017; Teuwen et al., 2020) (RIM) was trained to reconstruct complex MRI from the Calgary-Campinas MRI reconstruction challenge (Beauferris et al., 2020) (Experiment B).

3.2.1. Datasets

The fastMRI dataset (Zbontar et al., 2018) is a large, open dataset of knee MRI, providing 34,742 MRI slices of 973 volumes for training and 7135 slices of 199 volumes for validation. The MRI was acquired using 15 receiver channels at 1.5T and 3T systems, providing various contrasts at 0.5 mm^2 resolution. The MRI was acquired using an Cartesian 2D turbo spin echo sequence (TE = 27–34 ms, TR = 2200–3000 ms). Approximately half of the scans were acquired with fat suppression, while the other half did not employ fat suppression. The dataset provided unprocessed, fully-sampled, complex multi-coil k-space for every slice, along with fully-sampled ground-truth images. However, these ground-truth images only contained the magnitude component as they were computed using a root-sum-of-squares coil combination, which precluded using a complex-valued loss function. We computed coil-sensitivity maps (CSMs) and generated fully-sampled complex coil-weighted reconstructions as target images (Inati et al., 2014). Finally, the multi-coil k-space was retrospectively undersampled by multiplying it with an equispaced Cartesian undersampling mask ($R = 4$), preserving 8% of the center lines.

The Calgary-Campinas multi-channel MR dataset (Beauferris et al., 2020) is an open, 2D brain MRI dataset providing 12-channel k-space of 167 volunteers, acquired using a T_1-weighted gradient-recalled echo sequence (TE = 2.6–3.1 ms, TR = 6.3–7.4 ms, TI = 400–650 ms). The MRI was acquired using a 3T system at 1 mm^3 isotropic resolution. The dataset provided unprocessed, fully-sampled, complex multi-coil k-space for every slice, along with fully-sampled ground-truth images. As with the FastMRI dataset, the ground-truth images only contained the magnitude component as they were computed using a root-sum-of-squares coil combination, which precluded using a complex-valued loss function. Compuation of CSMs allowed generation of fully-sampled complex coil-weighted reconstructions as target images (Inati et al., 2014). Data were undersampled using provided random Poisson disc undersampling masks, yielding an acceleration factor $R = 5$. In total, the dataset provided 12,032 slices for training, 5120 slices for validation, and 7800 slices as the test set.

3.2.2. Architectures

The E2E-VarNet (Sriram et al., 2020) is an unrolled network for undersampled MRI reconstruction consisting of multiple cascades. Each cascade computes

$$k^{t+1} = k^t - \eta^t M(k^t - \hat{k}) + G(k^t)$$

where k^t is the k-space per coil from the previous cascade, M is the sampling mask, $\hat{k}$ is the k-space sampled by the MRI, η is a learnable parameter, and G is a convolutional neural network operating on the coil-combined, Fourier-transformed k^t. Its output was then again coil-weighted and Fourier-transformed. The CSMs were estimated from the central lines of k-space using a U-Net of depth 4 and 8-channel input. In our case, G is a U-Net of depth 4 with 18 input channels (Ronneberger et al., 2015). The entire model consisted of 8 cascades. The reconstructed image was obtained by transforming the coil-combined k-space of the final cascade to image-space. Sriram et al. showed that end-to-end estimation of the CSMs allows higher reconstruction quality than using than using traditional methods to estimate CSMs from the undersampled MRI, as the ground-truth CSMs are not available as these were derived from fully-sampled k-space.

The RIM is a recurrent neural network proposed for inverse problems, such as undersampled MRI reconstruction (Putzky and Welling, 2017; Teuwen et al., 2020). The network starts with an initial estimate of the of the image x_0, which is the Fourier transform of the undersampled, coil-combined k-space. Then, this estimate is updated such that

$$x_{t+1}, s_{t+1} = x_t + h_\theta\left(\nabla_{y|x_t}, x_t, s_t\right).$$

Here, $\nabla_{y|x_t}$ is the data fidelity term, h_θ is the recurrent neural network parameterized by θ, and s is the internal state of the recurrent neural network. The final reconstruction is given by x_T for some predefined number of steps T.

A RIM was used for complex image reconstruction with $T = 8$ and h_θ is a 2-layer convolutional gated recurrent unit (GRU) with 64 hidden features. Additionally, a U-Net of depth 4 was used to estimate CSMs from the undersampled k-space, as the ground-truth CSMs were not available during training and CSM estimation using deep CNNs showed improved performance compared to traditional methods to estimate CSMs.

3.2.3. Experiments

Two experiments were performed: The E2E-VarNet model was used to reconstruct complex undersampled MRI of the fastMRI dataset to examine the performance of $\perp$-loss compared to $\ell^2_{\mathbb{C}}$ and ℓ_{SSIM} (Experiment A), and the RIM was used to reconstruct complex undersampled MRI with added noise of the Calgary-Campinas dataset to examine the noise robustness of $\perp$-loss compared to $\ell^2_{\mathbb{C}}$ and ℓ_{SSIM} (Experiment B).

For every experiment, four separate models were trained:

1. A model trained to minimize the $\ell^2_{\mathbb{C}}$ loss function between the ground truth and the estimated image.
2. A model trained to minimize the $\perp$+ℓ^2-loss between the ground truth and the estimated image.
3. A model trained to minimize the function ℓ_{SSIM} loss function between the magnitude ground-truth and the magnitude estimated image, discarding phase information.
4. A model trained to minimize the function $\perp$+SSIM $= \ell^{\perp} \cdot \psi + (1 - \psi) \cdot \ell_{\text{SSIM}}$ between the ground-truth and the estimated image. Here, ψ is a learnable parameter between 0 and 1.

Each model was trained on an NVIDIA V100 GPU with 32GB VRAM. The models were trained deterministically with a fixed random seed and identical hyperparameters for a fair comparison. The models from experiment A were trained for 50 epochs using a batch size of 1. At each epoch, 25% of the volumes were randomly sampled from all training data to manage training time, using the Adam optimizer (Kingma and Ba, 2017) with a learning rate of 10^{-3}. After 40 epochs, the learning rate was reduced to 10^{-4}.

The models in experiment B were trained for 200,000 steps with a batch size of 4 using the Adam optimizer. The base learning rate was 10^{-4} and was halved after every 50,000 steps. During training of the models for Experiment B, Gaussian noise $\epsilon \sim \mathcal{N}(0, \xi)$ was added to the sampled k-space points, where $\xi \geq 0$ was drawn from a uniform distribution $\xi \sim \mathcal{U}(0, 0.2 \cdot |k_0|)$, with $|k_0|$ as the magnitude of the central point in k-space.

3.2.4. Evaluation

The reconstruction quality of the models was evaluated over the entire field of view using the mean and standard deviation of the peak signal-to-noise ratio (PSNR), SSIM, and visual information fidelity (VIF) (Sheikh and Bovik, 2005) metrics of the magnitude estimate to the magnitude target. VIF is a multi-resolution image quality metric based on the mutual information between two images and has been shown to strongly correlate with MRI quality as assessed by radiologists (Mason et al., 2020). The VIF between a target image I_{target} and estimated image I_{est} is computed as

$$\text{VIF}(I_{\text{target}}, I_{\text{est}}) = \sum_j \frac{\log_{10}\left(C_j(I_{\text{target}}) \cdot g_j(I_{\text{target}}, I_{\text{est}})^2\right)}{\log_{10}\left(C_j(I_{\text{target}})\right)}$$

Here, $j = 0, 1 \ldots, J$ is the resolution level, $C_j(I) = 1 + \frac{(I_j)^2 - (I^2)_j}{\sigma_N^2}$ is the information in an image I at resolution level j, and $g_j(I_{\text{target}}, I_{\text{est}}) = \frac{C_j(I_{\text{est}})}{C_j(I_{\text{target}})}$ is the mutual information between I_{target} and I_{est}. The jth sub-band of an image I is approximated by blurring I using a zero-meaned Gaussian kernel with $\sigma^2 \propto 2^{J-j}$ and downsampling the image by factor 2^j. σ_N^2 is a parameter of the vision model and was chosen as 0.4 for MRI images (Beauferris et al., 2020). The VIF is bounded by 0, but can reach values greater than one if the reconstructed image shows less noise or improved contrast compared to the target image.

Besides magnitude quality, we evaluated the mean and standard deviation of the mean-squared error between the estimated phase map and the target phase map. Wilcoxon signed-rank tests with $\alpha < 0.01$ were performed when comparing the results. During the evaluation of experiment B, Gaussian noise was added to the input data using $\xi \sim \mathcal{U}(0, 0.5 \cdot |k_0|)$, allowing comparison of the loss functions based on image quality depending on the noise level.

3.3. Image registration

To investigate whether the proposed $\perp$-loss generalizes beyond MR image reconstruction, we trained a model to learn deformation vector fields (DVFs) from a pair of 2D MR images. A DVF is a vector field $V \in \mathbb{R}^{2\times m\times n}$, where the first dimension is the displacement in the x-direction, while the second dimension is the displacement in the y-direction. A transformation $f : \mathbb{R}^{2\times m\times n} \to \mathbb{C}^{m\times n}$ mapped x-displacements of the DVF to the real part of the complex field and y-displacements of the DVF to the imaginary part of the complex field.

3.3.1. Data acquisition and processing

We have used magnitude-only sagittal cine-MRI of 135 patients with abdominal cancer undergoing radiotherapy simulation at our department (Terpstra et al., 2020) to train the image registration model on simulated ground-truth deformations. The data were acquired using a two-dimensional Cartesian balanced steady-state free precession (bSSFP) sequence on a 1.5 T MRI scanner (Ingenia MR-RT, Philips, Best, the Netherlands) using 28 receiver channels, TE/TR=1.3/2.8 ms, a flip angle of 50 degrees, a resolution of 1.4 mm^2 and a field of view of 320 mm^2, yielding an acquisition matrix size of 224×224 pixels. In total, 31,750 magnitude-only dynamics were collected. The signal intensity over all dynamics was linearly rescaled to [0, 1], clipping the top 99th intensity percentile of the dynamics in a cine-MRI. Images were augmented using random affine transformations (rotation between $[-20, 20]$ degrees, translations between $[-10, 10]$% of the image size, scaling between $[75, 125]$%, and shearing between $[-10, 10]$%), random horizontal and vertical flips, and cropping to a random region of 224×224 pixels.

The ground-truth deformation was generated using Gryds (Eppenhof and Pluim, 2019), which generates random, smooth DVFs $\mathcal{D} \in \mathbb{R}^{2\times m\times n}$ using a B-spline basis from $\mathbb{R}^{2\times 3\times 3} \sim \mathcal{N}(0, \mathcal{U}(0.001, 0.025))$. The motion parameters were selected such that the determinant of the Jacobian was higher than 0 everywhere to avoid folding by the DVF, which would have resulted in implausible motion (Reinhardt et al., 2008). Warping randomly-chosen cine frames using the ground-truth DVF $\mathcal{D}$ yielded a magnitude reference image, magnitude warped image. The ground-truth DVF was mapped to the complex domain by applying the transformation $f : \mathbb{R}^{2\times m\times n} \to \mathbb{C}^{m\times n}$ to $\mathcal{D}$.

3.3.2. Model architecture, training, and evaluation

We trained a residual U-Net (Ma et al., 2021) of depth 5 with two input channels and two output channels, with two residual units per level. Every residual unit consisted of a 3×3 convolution, a two-dimensional instance norm (Ulyanov et al., 2017), and PReLU non-linear activation (He et al., 2015). The first convolution used stride 2, while the latter convolution used stride 1. Two variants of this image registration model were trained:

1. One model was trained to minimize the end-point error (EPE), which is equivalent to the $\ell^2_{\mathbb{C}}$ loss function in the $\mathbb{R}^2$-domain. This is the de-facto loss function for training image registration models (Butler et al., 2012; Dosovitskiy et al., 2015).
2. One model was trained to minimize the $\perp$+ℓ^2-loss.

The models used the reference and warped image as input to reconstruct the DVF. The models were trained using the AdamW optimizer (Loshchilov and Hutter, 2019) for 150 epochs with a base learning rate of 1×10^{-4} and a weight decay of 5×10^{-4} using a batch size of 16 on an NVIDIA V100 GPU with 32 GB VRAM. After this training, the models were fine-tuned for 50 epochs using a learning rate of 10^{-6}.

The models were evaluated on the residual EPE and SSIM after image registration with the estimated DVF. Moreover, the models were evaluated on λ, i.e., the ratio between magnitude overestimation and magnitude underestimation. Wilcoxon signed-rank tests with $\alpha < 0.01$ were performed when comparing the results.

4. Results

4.1. Loss landscape

The loss landscapes of the $\ell^1_{\mathbb{C}}$, $\ell^2_{\mathbb{C}}$ and $\perp$+ℓ^2-loss functions are shown in Fig. 3. The $\ell^1_{\mathbb{C}}$ and $\ell^2_{\mathbb{C}}$ loss functions both display an asymmetry, assigning a higher loss to vectors with $\lambda > 1$ and a lower loss to vectors with $\lambda < 1$ with the same phase error. For $\perp$+ℓ^2-loss, this loss landscape is symmetric with steep gradients far from $\lambda = 1$ and a large region where the loss is low, as can be observed from the size of the area within the first isocontour around $\lambda = 1$ and $\hat{\varphi} = 0$.

4.2. Image reconstruction

4.2.1. Experiment A: E2E-VarNet

Each E2E-VarNet model was trained in approximately 48 hours. As shown in Table 1, the proposed $\perp$+ℓ^2-loss yields reconstructions with significantly higher SSIM and VIF than $\ell^2_{\mathbb{C}}$ (Wilcoxon, $p \ll 0.01$) – with an SSIM of 0.90 ± 0.07 versus 0.86 ± 0.10 – while

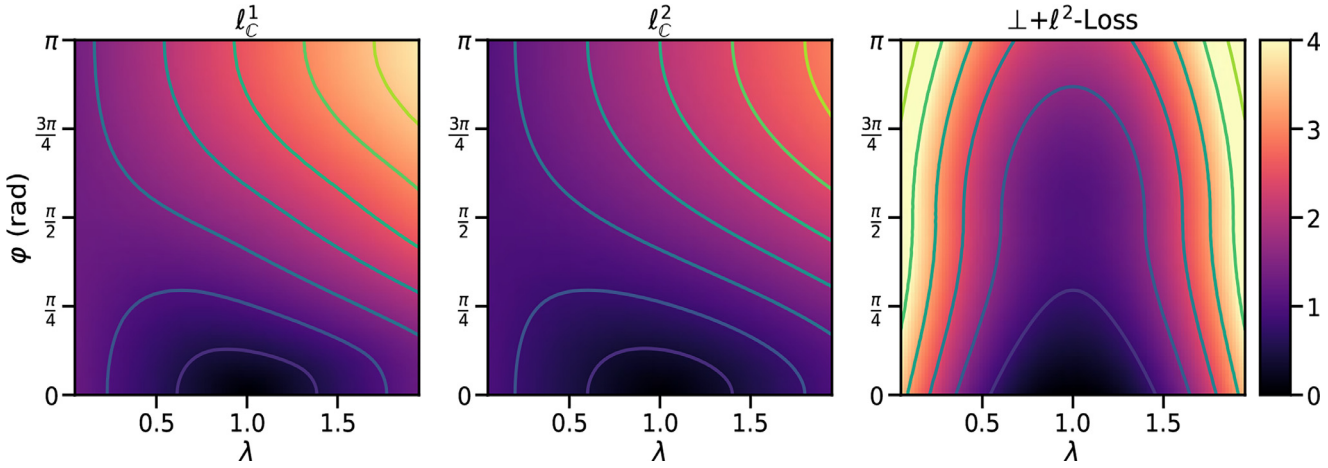

Fig. 3. The loss landscapes for the $\ell_{\mathbb{C}}^1$ loss function (left), $\ell_{\mathbb{C}}^2$ loss function (center) and $\perp+\ell^2$-loss (right). Lines through the loss landscape are isolines of the losses. The $\ell_{\mathbb{C}}^1$ and $\ell_{\mathbb{C}}^2$ loss functions show an asymmetry, resulting in a higher loss value for $\lambda > 1$, while $\perp+\ell^2$-loss is symmetric.

Table 1
Evaluation of E2E-VarNet and RIM models trained to minimize $\ell_{\mathbb{C}}^2$, $\perp+\ell^2$-loss, SSIM or $\perp$+SSIM-loss functions using the structural similarity, peak-signal-to-noise ratio, phase mean-squared error, and VIF metrics. Best results per model and metric are marked in boldface.

Loss \ Model	E2E-VarNet/FastMRI				RIM/Calgary-Campinas			
	SSIM (↑)	PSNR (↑)	Phase MSE (↓)	VIF (↑)	SSIM (↑)	PSNR (↑)	Phase MSE (↓)	VIF (↑)
$\ell_{\mathbb{C}}^2$	0.86 ± 0.10	31.2 ± 3.1	$\mathbf{0.04 \pm 0.03}$	0.80 ± 0.17	0.89 ± 0.04	31.2 ± 2.5	$\mathbf{0.03 \pm 0.01}$	0.98 ± 0.09
$\perp+\ell^2$-loss	0.90 ± 0.07	33.0 ± 3.0	$\mathbf{0.04 \pm 0.03}$	0.83 ± 0.16	$\mathbf{0.93 \pm 0.03}$	$\mathbf{31.4 \pm 2.6}$	$\mathbf{0.03 \pm 0.01}$	0.99 ± 0.09
SSIM	$\mathbf{0.91 \pm 0.07}$	$\mathbf{33.9 \pm 3.5}$	0.12 ± 0.03	$\mathbf{0.86 \pm 0.15}$	$\mathbf{0.93 \pm 0.03}$	31.0 ± 2.6	0.15 ± 0.01	0.99 ± 0.11
$\perp$+SSIM-loss	$\mathbf{0.91 \pm 0.07}$	33.1 ± 3.0	$\mathbf{0.04 \pm 0.03}$	$\mathbf{0.86 \pm 0.16}$	$\mathbf{0.93 \pm 0.03}$	30.9 ± 2.5	$\mathbf{0.03 \pm 0.01}$	$\mathbf{1.0 \pm 0.11}$

Values marked in bold are statistically significant.

the phase error is equally low. Example reconstructions are shown in Fig. 4, where it can be seen that the $\ell_{\mathbb{C}}^2$ reconstruction shows an apparent magnitude underestimation in the reconstructed knee tissue and a more significant error in the background region. The reconstruction from the $\perp+\ell^2$-loss model shows no clear preference for magnitude overestimation or underestimation. However, less signal is present in the background in this specific instance, indicating better denoising properties.

The model trained to maximize the magnitude SSIM results in even higher magnitude quality but does not reconstruct usable phase information, as the network received no loss for the phase term. Combining the SSIM loss with $\perp$-loss using the $\perp$+SSIM-loss results in reconstructions of equal quality as the SSIM model but has a similar phase reconstruction performance as the $\ell_{\mathbb{C}}^2$ loss and $\perp+\ell^2$-loss. However, there seems to be an increased residual magnitude error compared to the SSIM reconstruction, indicating a trade-off between denoised and dealiased magnitude reconstructions and high-quality phase maps.

4.2.2. Experiment B: RIM

Each RIM was trained in approximately 8 hours. Similar to the results of the E2E-VarNet model, using $\perp+\ell^2$-loss yields significantly higher reconstruction quality in the noise-free case than $\ell_{\mathbb{C}}^2$ (Wilcoxon, $p \ll 0.01$) – with SSIM values of 0.89 ± 0.04 versus 0.93 ± 0.03. Additional quantitative results are presented in Table 1. The model trained to maximize the magnitude SSIM outperforms both the $\perp+\ell^2$-loss and $\ell_{\mathbb{C}}^2$ models, but these reconstructions have high phase error. Based on the VIF metric, the $\perp$+SSIM-loss model outperforms the SSIM model (with a mean VIF of 1.0 ± 0.011 versus 0.99 ± 0.11) while having low phase error.

An example reconstruction without added noise is shown in Fig. 5. It can be observed that both SSIM and $\perp$+SSIM-loss models yield high-quality magnitude reconstructions. However, the SSIM model reconstructs poor phase images, while the phase images of the $\perp$+SSIM-loss model are similar to the ground-truth phase maps.

When evaluated with added noise, i.e., $\xi > 0$, the quality of the magnitude reconstructions degrade for all models. In Fig. 6, the performance of the four RIM models with added noise is evaluated based on the magnitude PSNR metric. Here, in the noise-free case, $\perp+\ell^2$-loss achieves the highest median PSNR of 32.8. However, the models were trained up to a noise level of 0.2. Up to this noise level, $\perp+\ell^2$-loss achieves the highest median PSNR. Beyond this noise level, the image quality of $\perp+\ell^2$-loss degrades faster than the competing models, yielding a median PSNR of 22.8 at noise level 0.5 versus 23.1 for the $\ell_{\mathbb{C}}^2$ model. The $\perp$+SSIM-loss model performs similarly to the SSIM model, which both yield an median PSNR of 32.1 in the noise-free case and 23.3 at noise-level 0.5.

The model trained to maximize the magnitude SSIM shows high phase error for every noise level and produced unusable phase maps. For the $\ell_{\mathbb{C}}^2$ models, the phase error slowly increased as the added noise increases. For the $\perp+\ell^2$-loss and $\perp+\ell^2$-loss, the phase error remained near-constant as the noise level was increased, showing superior phase reconstruction for all noise levels.

4.3. Image registration

Each image registration model was trained in approximately 6 hours. An example comparing the DVFs of the $\perp+\ell^2$-loss and EPE models to the ground-truth is shown in Fig. 7, where the model trained to minimize $\perp+\ell^2$-loss produces a lower residual error and a lower registration error, with a registration SSIM of 0.931 for the $\perp+\ell^2$-loss model versus 0.768 for the EPE model. Moreover, it can be seen from the difference DVF that the EPE model yields a relative magnitude underestimation of up to 25% compared to the target DVF in regions with significant deformation. Quantitative evaluation of both models demonstrates that minimizing the $\perp+\ell^2$-

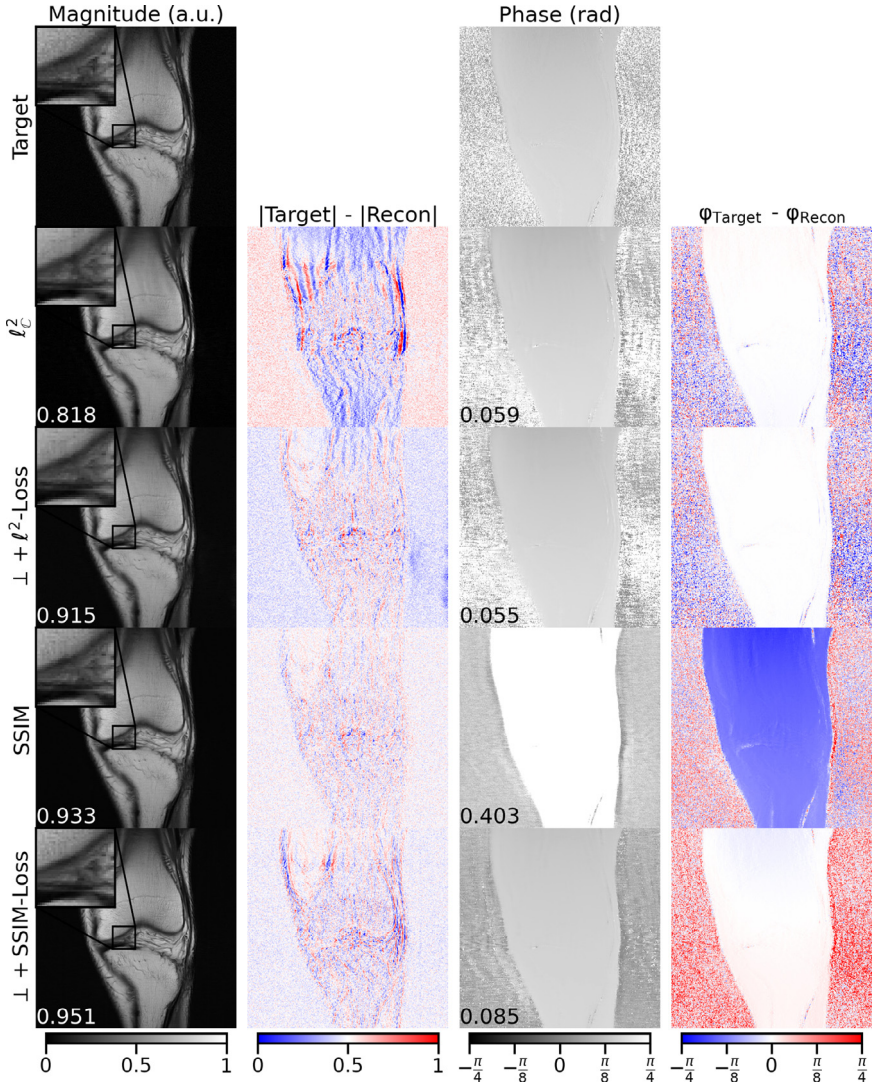

Fig. 4. Example reconstructions using an E2E-VarNet ($R = 4$). Examples of reconstructions of the same slice with normalized magnitude using four networks, trained to minimize the $\ell^2_{\mathbb{C}}$, $\perp$+ℓ^2-loss, SSIM-Loss, and $\perp$+SSIM-loss loss, respectively. The bottom-left number in magnitude reconstructions shows the foreground VIF (magnitude) while showing the foreground mean squared error for phase images. Using the $\perp$+ℓ^2-loss instead of the $\ell^2_{\mathbb{C}}$ loss significantly improves image quality. In the zoomed region of the $\ell^2_{\mathbb{C}}$ model, a hallucinated lesion is visible and is much less severe using the $\perp$+ℓ^2-loss model. Using the SSIM as a loss function yields even higher magnitude image quality and higher phase error than the $\ell^2_{\mathbb{C}}$ and $\perp$+ℓ^2-loss models. Using the $\perp$+SSIM-loss function obtains the highest magnitude image quality with low phase error. In the zoomed region, higher image quality and more contrast for the $\perp$+SSIM-loss model can be observed compared to the SSIM model.

Table 2
Evaluation of image registration models trained to minimize $\perp$+ℓ^2-loss or the EPE loss functions using the registered SSIM, remaining EPE, and λ, the ratio between magnitude overestimation and magnitude underestimation. Best results per metric are marked in boldface.

	SSIM (↑)	EPE (mm, ↓)	λ (≈ 1)
$\perp$+ℓ^2-loss	**0.927 ± 0.053**	**0.898 ± 0.598**	**0.960 ± 0.304**
EPE	0.895 ± 0.071	1.391 ± 0.704	0.720 ± 0.632

Values marked in bold are statistically significant.

loss produces significantly better DVFs, reducing the mean EPE from 1.39 mm to 0.89 mm (Wilcoxon, $p \ll 0.01$), as presented in Table 2 Moreover, the mean registered SSIM increased by approximately 0.03 (Wilcoxon, $p \ll 0.01$), and the mean value λ is significantly closer to 1. This indicates that the magnitude error is more symmetrically-distributed, i.e., magnitude underestimation is about as likely as magnitude overestimation for the $\perp$+ℓ^2-loss model. The error is significantly more biased towards magnitude underestimations for the EPE model than the $\perp$+ℓ^2-loss model (Wilcoxon, $p < 0.01$).

5. Discussion

In this work, we have identified that the $\ell^2_{\mathbb{C}}$ loss function exhibits a magnitude bias when applied to regression in the $\mathbb{C}$ and $\mathbb{R}^2$ domains, which may impact the performance of deep neural networks. Specifically, for the reconstruction and processing of complex MRI, we have shown this bias is detrimental to the performance of CNNs for complex MRI reconstruction and MRI registration. To address this issue, we have proposed a new loss function called $\perp$-loss, which produces a symmetric loss space when applied to regression in the $\mathbb{C}$ and $\mathbb{R}^2$ domains.

We have shown that $\perp$+ℓ^2-loss reconstructs undersampled MRI with higher quality than models trained to minimize the $\ell^2_{\mathbb{C}}$ loss. A hybrid $\perp$+SSIM-loss function allowed image reconstruction with similar image quality as networks trained to maximize the magnitude SSIM while generating high-quality phase maps. The E2E-VarNet model trained to minimize $\perp$+ℓ^2-loss as loss function increased the magnitude SSIM by approximately 0.04 and the PSNR by approximately 1.8 dB compared to the $\ell^2_{\mathbb{C}}$ loss. Using $\perp$+ℓ^2-loss as the loss function for the RIM model increased the magnitude

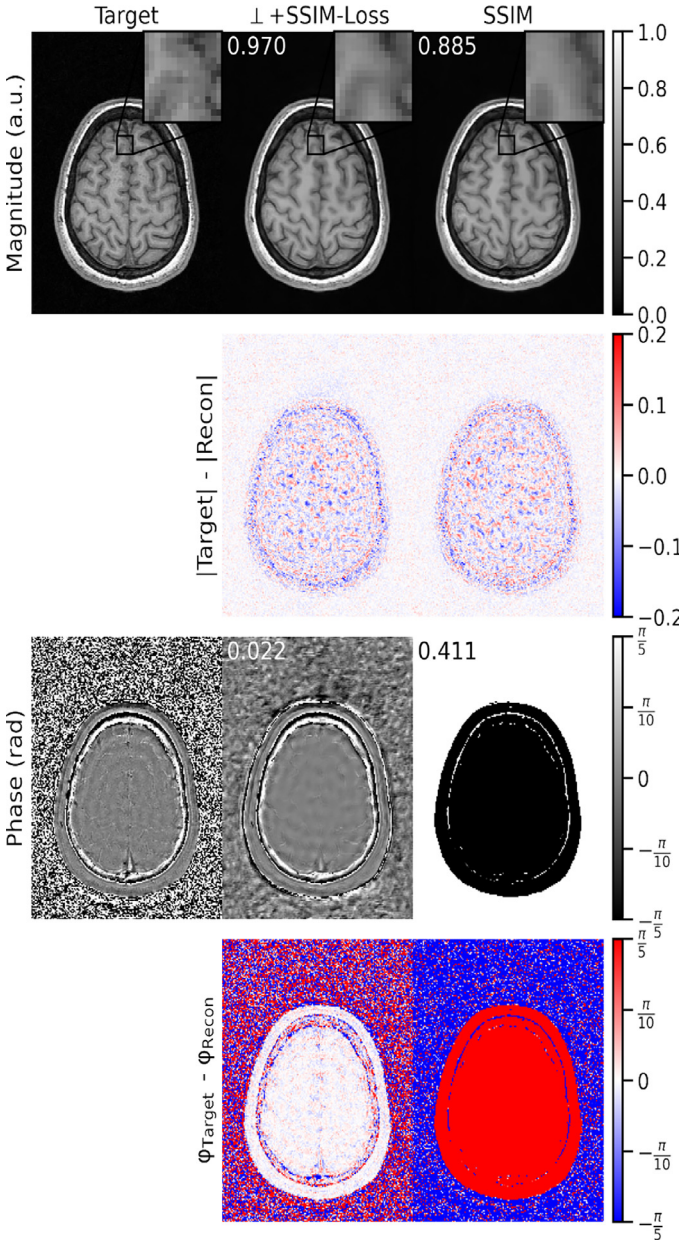

Fig. 5. Example reconstructions of the RIM ($R = 5$). Examples of reconstructions with the normalized magnitude of the same slice, using a RIM trained to minimize the ⊥+SSIM-loss and SSIM-Loss, respectively. Using ⊥+SSIM-loss as a loss function produces significantly higher-quality phase information (Indicated by the phase MSE, shown in the top-left corner of the phase images) and similar magnitude reconstructions (Indicated by the VIF on the magnitude images). The zoomed region indicates an artifact in the SSIM reconstruction.

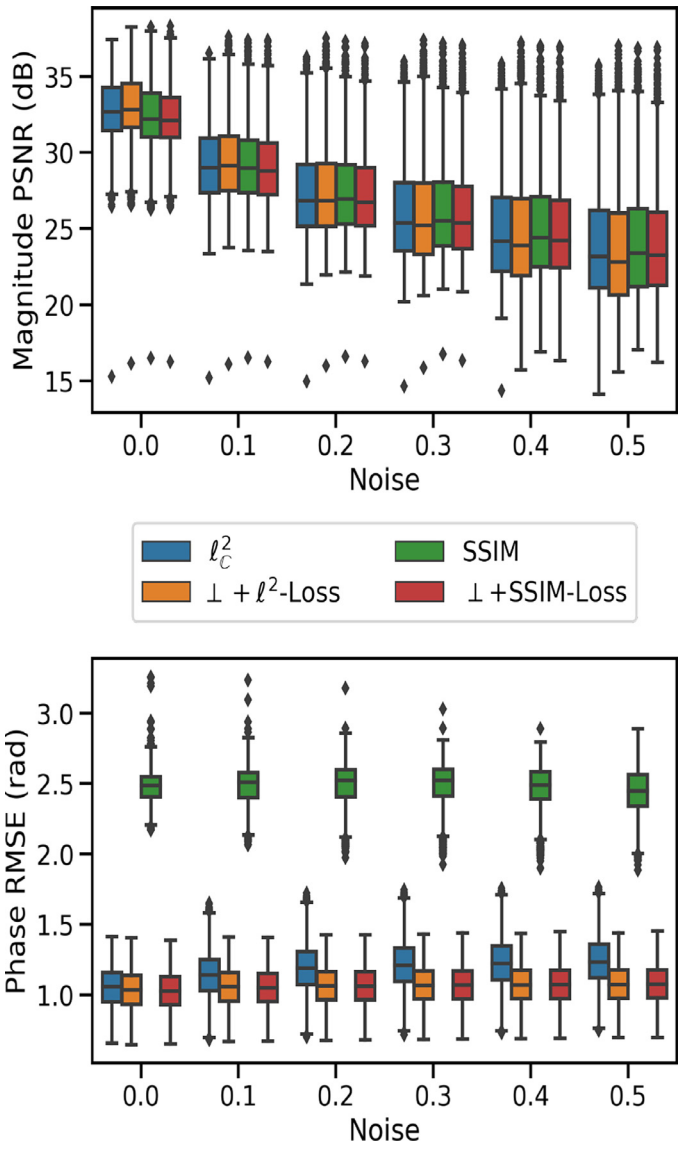

Fig. 6. Quantitative RIM results. The PSNR of the magnitude reconstructions (top) and the root-mean square error of the reconstructed phase (bottom) were evaluated for the four RIM models trained with different loss functions. On the horizontal axis is the magnitude of the added noise. In the low-noise regime, ⊥+ℓ^2-loss achieves a median PSNR of 32.8 versus 32.6 for $\ell^2_{\mathbb{C}}$. The PSNR of ⊥+ℓ^2-loss decreases faster at higher noise levels, while the ⊥+SSIM-loss model achieves a similar PSNR as the SSIM RIM. The SSIM model show very high phase error, while the ⊥-loss models show superior phase reconstruction compared to the $\ell^2_{\mathbb{C}}$ and SSIM models.

SSIM by approximately 0.04 and the VIF by approximately 0.01 compared to the $\ell^2_{\mathbb{C}}$ loss.

These models trained to optimize for ⊥+ℓ^2-loss display competing performance to state-of-the-art image reconstruction models while achieving high-quality phase reconstruction. However, some literature currently reports higher scores than the values we reported here. For example, an SSIM of 0.930 and PSNR of 40 when reconstructing the fastMRI "challenge" dataset using an E2E-VarNet model has been reported (Sriram et al., 2020). On the validation dataset, they reported an SSIM of 0.923 whereas we found an SSIM of 0.91 ± 0.07. This difference in performance could be caused by multiple reasons, such as a better training regime with larger batch sizes, longer training times, larger models, or our training with complex target images instead of standard magnitude-only training.

These results indicate that deep learning models can use phase information from the complex signal to improve magnitude reconstructions. This observation aligns with other works, which found that including phase information or designing a network that uses complex-valued model weights – thus taking full advantage of the complex information present in the data – improves both magnitude and phase reconstruction performance (Pezzotti et al., 2020; Cole et al., 2021). Moreover, improved phase reconstruction might enable deep learning applications, such as quantitative susceptibility mapping (QSM) (Wang and Liu, 2015) or four-dimensional flow MRI (Markl et al., 2012). We speculate that reconstructing phase information makes the inversion problem less ill-posed as $\ell^{\perp}$ acts as a phase regularizer.

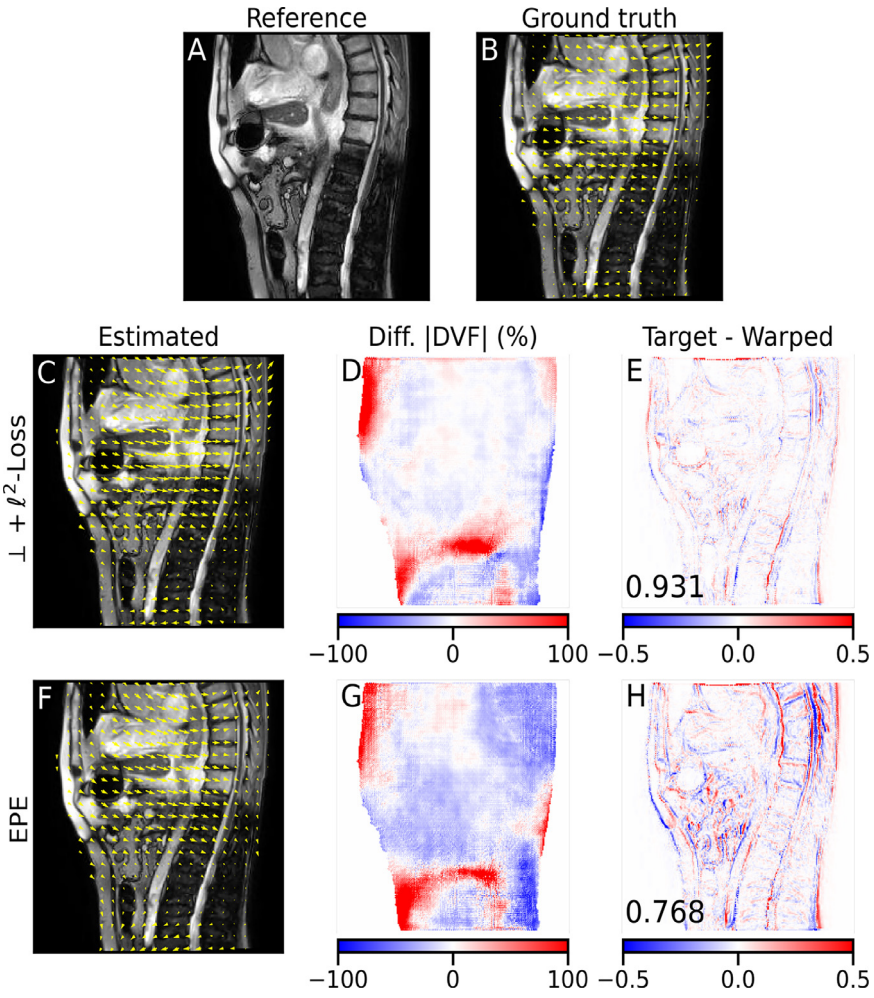

Fig. 7. Image registration results. The reference image (A), warped image (B) were input to the models. The ground-truth DVF is also shown in (B), but was not presented to the model. Figures C and F show the estimates produced by the $\perp$+ℓ^2-loss and EPE models, respectively. Figures D and G show the relative magnitude error between the produced DVF and target DVF for the $\perp$+ℓ^2-loss and EPE models, respectively. Finally, figures E and H show the normalized intensity difference (a.u.) between the images registered by the $\perp$+ℓ^2-loss and EPE DVF and the ground-truth image, respectively. In the bottom-left of C and H, the SSIM between the estimated registration and the ground-truth registration is shown. The EPE model shows a larger registration error (H), as visible by larger intensities in the difference image and the lower registration SSIM. Also, the estimated DVF by the EPE model shows a larger negative error (G), which indicates a magnitude underestimation, while the $\perp$+ℓ^2-loss shows errors closer to zero.

Finally, we demonstrated that $\perp$-loss is not limited to complex image reconstruction but can be applied to problems in other domains. In particular, we have shown that using $\perp$+ℓ^2-loss as the loss function outperforms the EPE when applied to an image registration task. The DVFs are estimated with lower residual error, reducing from approximately 1.4 mm to 0.9 mm, better image registration, and a more symmetrically distributed error.

In past literature (Buczko and Willert, 2017; Fermüller et al., 2001), it has been established that applying least squares regression to the $\mathbb{R}^2$ domain in the presence of noise yields a magnitude underestimation. Our findings align with these previous results, and we have shown that these issues also apply to the complex domain in general and occur during non-linear regression.

Currently, we have only used real-valued networks for comparison. Recently, networks with complex-valued weights have been proposed, and it has been shown that these models achieve superior performance (Cole et al., 2021). Future work might investigate whether these complex-valued networks could enable even better performance with $\perp$-loss, as every part from input to output is defined in the complex plane.

We have identified that it is beneficial to have a smooth and symmetric loss function for MR image reconstruction and image registration. However, it is crucial to consider the domain of the problem, and these qualities do not necessarily transfer to other tasks. For example, it has been identified that non-convex loss functions achieve better performance for image classification tasks (Zhao et al., 2010) while convex loss functions are generally preferred, and for anomaly detection, where it is beneficial to use an asymmetric loss function that penalizes outliers harder than inliers (Fu and Wang, 2021).

A symmetrically-distributed error in image registration is an attractive property for many applications. For example, in radiotherapy, accurate tracking of tumor motion enables real-time adaptive treatments, where the position and shape of the radiation beam are adapted to the current anatomy (Keall et al., 2019b). A biased image registration algorithm might underdose a tumor while risking increased toxicity to increased dose delivery to healthy tissue (Cai et al., 2007). With a symmetric loss function like $\perp$-loss, a more conformal dose delivery could be obtained.

For future work, it would be helpful to further study the properties of $\perp$-loss. For example, it would be interesting to analyze further the noise robustness of $\perp$-loss to conventional loss functions. Currently, $\perp$+ℓ^2-loss does not outperform $\ell^2_{\mathbb{C}}$ in the magnitude domain when subjected to more noise than seen during

training. In the phase domain, however, the performance of $\perp$+ℓ^2-loss and $\perp$+SSIM-loss is minimally perturbed by the added noise. The cause of this phenomenon is unknown, but we speculate this might be caused by a low gradient in the phase direction in high-noise environments when $\lambda \approx 1$ (Fig. 3). Moreover, the loss function consists of weighing between the magnitude term and phase term. This parameter is currently optimized as part of the network, while more manual tuning could change the emphasis of the loss between the domains. Further experiments could investigate the performance of $\perp$-loss in a low SNR regime or at higher undersampling factors.

For image registration, it would be interesting to extend the definition of $\perp$-loss to three dimensions. It is simple to express the scalar rejection as defined in Eq. 1 in $\mathbb{R}^n$, thus facilitating extension to $\mathbb{R}^3$ and beyond. Further analysis could establish whether the attractive properties of the $\perp$-loss function in the complex plane, e.g., symmetric output or higher image registration performance, translates to higher dimensions.

Finally, it would be interesting to apply $\perp$-loss to non-deep learning image reconstruction. For example, $\perp$-loss could be used as a cost function in an iterative compressed sense algorithm. Future research could investigate whether using $\perp$+ℓ^2-loss or $\perp$+SSIM-loss could improve image reconstruction or result in faster iteration convergence.

As using $\perp$+SSIM-loss allows for better image reconstruction model training, further undersampling of k-space with equal image quality could be achieved by such models. This would increase MRI efficiency, increasing patient throughput. Moreover, reconstructing images with higher quality could lead to better diagnosis by radiologists.

6. Conclusion

We have identified that the conventional $\ell^2_{\mathbb{C}}$ loss function gives rise to an asymmetric loss landscape in the complex field, resulting in an underestimation bias of the reconstructed magnitude.

To resolve this problem, we have presented $\perp$-loss, a loss function defined in a complex vector space for MRI reconstruction and image registration with a symmetric magnitude/phase loss landscape. We have applied $\perp$+ℓ^2-loss to undersampled complex MR image reconstruction, obtaining higher-quality reconstructions than when minimizing the $\ell^2_{\mathbb{C}}$ loss. Compared to state-of-the-art models that maximize the SSIM, we achieved competitive performance in the magnitude domain and superior performance for phase reconstruction using $\perp$+SSIM-loss.

Finally, we showed that $\perp$+ℓ^2-loss generalized beyond complex image reconstruction and could be applied for image registration. We achieved higher performance than models minimizing the end-point error, with lower residual error, better image registration performance, and a symmetric output error.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is part of the research program HTSM with project number 15354, which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO) and Philips Healthcare. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Quadro RTX 5000 GPU used for prototyping this research.

References

Adler, J., Öktem, O., 2017. Solving ill-posed inverse problems using iterative deep neural networks. Inverse Probl. 33 (12), 124007. doi:10.1088/1361-6420/aa9581.
Beauferris, Y., Teuwen, J., Karkalousos, D., Moriakov, N., Caan, M., Rodrigues, L., Lopes, A., Pedrini, H., Rittner, L., Dannecker, M., Studenyak, V., Grger, F., Vyas, D., Faghih-Roohi, S., Jethi, A. K., Raju, J. C., Sivaprakasam, M., Loos, W., Frayne, R., Souza, R., 2020. Multi-channel MR Reconstruction (MC-MRRec) Challenge – Comparing Accelerated MR Reconstruction Models and Assessing Their Generalizability to Datasets Collected with Different Coils. arXiv:2011.07952
Buczko, M., Willert, V., 2017. Monocular outlier detection for visual odometry. In: 2017 IEEE Intelligent Vehicles Symposium (IV), pp. 739–745.
Bustin, A., Fuin, N., Botnar, R.M., Prieto, C., 2020. From compressed-sensing to artificial intelligence-based cardiac MRI reconstruction. Front. Cardiovasc. Med. 7, 17.
Butler, D.J., Wulff, J., Stanley, G.B., Black, M.J., 2012. A naturalistic open source movie for optical flow evaluation. In: Fitzgibbon, A. (Ed.), European Conf. on Computer Vision (ECCV). Springer-Verlag, pp. 611–625.
Cai, J., Read, P.W., Baisden, J.M., Larner, J.M., Benedict, S.H., Sheng, K., 2007. Estimation of error in maximal intensity projection-based internal target volume of lung tumors: A Simulation and comparison study using dynamic magnetic resonance imaging. Int. J. Radiat. Oncol.*Biol.*Phys. 69 (3), 895–902.
Cole, E., Cheng, J., Pauly, J., Vasanawala, S., 2021. Analysis of deep complex-valued convolutional neural networks for MRI reconstruction and phase-focused applications. Magn. Reson. Med. 86 (2), 1093–1109.
Collins, C.M., Wang, Z., 2011. Calculation of radiofrequency electromagnetic fields and their effects in MRI of human subjects. Magn. Reson. Med. 65 (5), 1470–1482.
Dosovitskiy, A., Fischer, P., Ilg, E., Hausser, P., Hazirbas, C., Golkov, V., Van Der Smagt, P., Cremers, D., Brox, T., 2015. Flownet: learning optical flow with convolutional networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2758–2766.
Eppenhof, K.A.J., Pluim, J.P.W., 2019. Pulmonary CT registration through supervised learning with convolutional neural networks. IEEE Trans. Med. Imaging 38 (5), 1097–1105.
Fermüller, C., Shulman, D., Aloimonos, Y., 2001. The statistics of optical flow. Comput. Vis. Image Underst. 82 (1), 1–32.
Fu, L., Wang, Y.-G., 2021. Robust regression with asymmetric loss functions. Stat. Methods Med. Res. 30 (8), 1800–1815.
Fu, Y., Lei, Y., Wang, T., Curran, W.J., Liu, T., Yang, X., 2021. A review of deep learning based methods for medical image multi-organ segmentation. Phys. Med. 85, 107–122. doi:10.1016/j.ejmp.2021.05.003.
Fuller, W.A., 2009. Measurement Error Models, 305. John Wiley & Sons.
Gleser, L.J., 1981. Estimation in a multivariate "errors in variables"regression model: large sample results. Ann. Stat. 9 (1), 24–44.
Griswold, M.A., Jakob, P.M., Heidemann, R.M., Nittka, M., Jellus, V., Wang, J., Kiefer, B., Haase, A., 2002. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn. Reson. Med. 47 (6), 1202–1210.
Haji-Valizadeh, H., Guo, R., Kucukseymen, S., Tuyen, Y., Rodriguez, J., Paskavitz, A., Pierce, P., Goddu, B., Ngo, L.H., Nezafat, R., 2021. Comparison of complex k-space data and magnitude-only for training of deep learning based artifact suppression for real-time cine MRI. Front. Phys. 9, 475.
Hammernik, K., Klatzer, T., Kobler, E., Recht, M.P., Sodickson, D.K., Pock, T., Knoll, F., 2018. Learning a variational network for reconstruction of accelerated MRI data. Magn. Reson. Med. 79 (6), 3055–3071.
He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. arXiv:1502.01852
Heckbert, P., 1994. Graphics Gems IV (IBM version). Elsevier.
Inati, S., Hansen, M., Kellman, P., 2014. A fast optimal method for coil sensitivity estimation and adaptive coil combination for complex images. In: Proceedings of the 22nd Annual Meeting of ISMRM, Milan, p. 4407.
Janocha, K., Czarnecki, W. M., 2017. On loss functions for deep neural networks in classification. arXiv preprint arXiv:1702.05659
Jaubert, O., Montalt-Tordera, J., Knight, D., Coghlan, G.J., Arridge, S., Steeden, J.A., Muthurangu, V., 2021. Real-time deep artifact suppression using recurrent U-Nets for low-latency cardiac MRI. Magn. Reson. Med. 86 (4), 1904–1916.
Kanai, S., Yamada, M., Takahashi, H., Yamanaka, Y., Ida, Y., 2021. Smoothness analysis of loss functions of adversarial training. arXiv preprint arXiv:2103.01400
Keall, P., Poulsen, P., Booth, J.T., 2019. See, think, and act: real-time adaptive radiotherapy. Semin. Radiat. Oncol. 29 (3), 228–235. Adaptive Radiotherapy and Automation
Keall, P., Poulsen, P., Booth, J.T., 2019. See, think, and act: real-time adaptive radiotherapy. In: Seminars in Radiation Oncology, vol. 29. Elsevier, pp. 228–235.
Keiper, T.D., Tai, A., Chen, X., Paulson, E., Lathuilire, F., Briault, S., Hbert, F., Cooper, D.T., Lachaine, M., Li, X.A., 2020. Feasibility of real-time motion tracking using cine MRI during MR-guided radiation therapy for abdominal targets. Med. Phys. 47 (8), 3554–3566.
Kim, B., Günther, T., 2019. Robust reference frame extraction from unsteady 2D vector fields with convolutional neural networks. Comput. Graph. Forum (Proc. EuroVis) 38 (3), 285–295.
Kingma, D. P., Ba, J., 2017. Adam: a method for stochastic optimization. arXiv:1412.6980
Kunin, D., Sagastuy-Brena, J., Ganguli, S., Yamins, D.L., Tanaka, H., 2021. Neural mechanics: symmetry and broken conservation laws in deep learning dynamics. In: International Conference on Learning Representations.

Lingala, S.G., Sutton, B.P., Miquel, M.E., Nayak, K.S., 2016. Recommendations for real-time speech MRI. J. Magn. Reson. Imaging 43 (1), 28–44.
Loshchilov, I., Hutter, F., 2019. Decoupled Weight Decay Regularization. arXiv:1711.05101
Lustig, M., Donoho, D., Pauly, J.M., 2007. Sparse MRI: the application of compressed sensing for rapid MR imaging. Magn. Reson. Med. 58 (6), 1182–1195.
Ma, N., Wenqi Li, Brown, R., Yiheng Wang, Behrooz, Gorman, B., Johnson, H., Yang, I., Kerfoot, E., Charliebudd, Yiwen Li, Adil, M., Yuan-Ting Hsieh, Arpit Aggarwal, Trentz, C., Aji, A., Masadcv, Graham, M., Murray, B., Gagan Daroach, Petru-Daniel Tudosiu„ Myron, McCormick, M.„ Ambros, Balamurali, Baker, C., Sellner, J., Fidon, L., Cgrain, 2021. Project-MONAI/MONAI: 0.5.3. https://zenodo.org/record/4323058.
Markl, M., Frydrychowicz, A., Kozerke, S., Hope, M., Wieben, O., 2012. 4D flow MRI. J. Magn. Reson. Imaging 36 (5), 1015–1036.
Mason, A., Rioux, J., Clarke, S.E., Costa, A., Schmidt, M., Keough, V., Huynh, T., Beyea, S., 2020. Comparison of objective image quality metrics to expert radiologists scoring of diagnostic quality of MR images. IEEE Trans. Med. Imaging 39 (4), 1064–1072. doi:10.1109/TMI.2019.2930338.
Mutic, S., Dempsey, J.F., 2014. The viewray system: magnetic resonance guided and controlled radiotherapy. Semin. Radiat. Oncol. 24 (3), 196–199. Magnetic Resonance Imaging in Radiation Oncology
Nagel, H.-H., Haag, M., 1998. Bias-corrected optical flow estimation for road vehicle tracking. In: Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), pp. 1006–1011. doi:10.1109/ICCV.1998.710839.
Patel, D., Sastry, P.S., 2021. Memorization in deep neural networks: does the loss function matter? In: Karlapalem, K., Cheng, H., Ramakrishnan, N., Agrawal, R.K., Reddy, P.K., Srivastava, J., Chakraborty, T. (Eds.) Advances in Knowledge Discovery and Data Mining. Springer International Publishing, Cham, pp. 131–142.
Pezzotti, N., Yousefi, S., Elmahdy, M. S., van Gemert, J., Schlke, C., Doneva, M., Nielsen, T., Kastryulin, S., Lelieveldt, B. P. F., van Osch, M. J. P., de Weerdt, E., Staring, M., 2020. An Adaptive Intelligence Algorithm for Undersampled Knee MRI Reconstruction. arXiv:2004.07339
Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P., 1999. SENSE: sensitivity encoding for fast MRI. Magn. Reson. Med. 42 (5), 952–962.
Putzky, P., Welling, M., 2017. Recurrent inference machines for solving inverse problems. arXiv preprint arXiv:1706.04008
Raaymakers, B.W., Lagendijk, J.J.W., Overweg, J., Kok, J.G.M., Raaijmakers, A.J.E., Kerkhof, E.M., van der Put, R.W., Meijsing, I., Crijns, S.P.M., Benedosso, F., van Vulpen, M., de Graaff, C.H.W., Allen, J., Brown, K.J., 2009. Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept. Phys. Med. Biol. 54 (12), N229–N237.
Reinhardt, J.M., Ding, K., Cao, K., Christensen, G.E., Hoffman, E.A., Bodas, S.V., 2008. Registration-based estimates of local lung tissue expansion compared to xenon CT measures of specific ventilation. Med. Image Anal. 12 (6), 752–763.
Rogers, T., Ratnayaka, K., Lederman, R.J., 2014. MRI catheterization in cardiopulmonary disease. Chest 145 (1), 30–36.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, Cham, pp. 234–241.
Schlemper, J., Caballero, J., Hajnal, J.V., Price, A.N., Rueckert, D., 2018. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. IEEE Trans. Med. Imaging 37 (2), 491–503.
Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423.
Sheikh, H.R., Bovik, A.C., 2005. A visual information fidelity approach to video quality assessment. The First International Workshop on Video Processing and Quality Metrics for Consumer Electronics, vol. 7. sn
Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C. L., Yakubova, N., Knoll, F., Johnson, P., 2020. End-to-End Variational Networks for Accelerated MRI Reconstruction. arXiv:2004.06688
Terpstra, M.L., Maspero, M., Bruijnen, T., Verhoeff, J.J.C., Lagendijk, J.J.W., van den Berg, C.A.T., 2021. Real-time 3D motion estimation from undersampled MRIusing multi-resolution neural networks. Med. Phys. 48 (11), 6597–6613.
Terpstra, M.L., Maspero, M., d'Agata, F., Stemkens, B., Intven, M.P.W., Lagendijk, J.J.W., van den Berg, C.A.T., Tijssen, R.H.N., 2020. Deep learning-based image reconstruction and motion estimation from undersampled radial k-space for real–time MRI-guided radiotherapy. Phys. Med. Biol. 65 (15), 155015.
Teuwen, J., Moriakov, N., Karkalousos, D., Caan, M., Yiasemis, G., 2020. Direct. https://github.com/directgroup/direct.
Ulyanov, D., Vedaldi, A., Lempitsky, V., 2017. Instance normalization: the missing ingredient for fast stylization. arXiv:1607.08022
Wang, Y., Liu, T., 2015. Quantitative susceptibility mapping (QSM): decoding MRI data for a tissue magnetic biomarker. Magn. Reson. Med. 73 (1), 82–101.
Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E., 2004. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13 (4), 600–612.
Wiesinger, F., Boesiger, P., Pruessmann, K.P., 2004. Electrodynamics and ultimate SNR in parallel MR imaging. Magn. Reson. Med. 52 (2), 376–390.
Wijlemans, J., Bartels, L., Deckers, R., Ries, M., Mali, W.T.M., Moonen, C., Van Den Bosch, M., 2012. Magnetic resonance-guided high-intensity focused ultrasound (MR-HIFU) ablation of liver tumours. Cancer Imaging 12 (2), 387.
Wright, G., 1997. Magnetic resonance imaging. IEEE Signal Process. Mag. 14 (1), 56–66.
Zbontar, J., Knoll, F., Sriram, A., Muckley, M. J., Bruno, M., Defazio, A., Parente, M., Geras, K. J., Katsnelson, J., Chandarana, H., Zhang, Z., Drozdzal, M., Romero, A., Rabbat, M. G., Vincent, P., Pinkerton, J., Wang, D., Yakubova, N., Owens, E., Zitnick, C. L., Recht, M. P., Sodickson, D. K., Lui, Y. W., 2018. fastMRI: an open dataset and benchmarks for accelerated MRI. CoRR abs/1811.08839
Zhang, B., Yen, Y.-F., Chronik, B.A., McKinnon, G.C., Schaefer, D.J., Rutt, B.K., 2003. Peripheral nerve stimulation properties of head and body gradient coils of various sizes. Magn. Reson. Med. 50 (1), 50–58.
Zhao, L., Mammadov, M., Yearwood, J., 2010. From convex to nonconvex: a loss function analysis for binary classification. In: 2010 IEEE International Conference on Data Mining Workshops, pp. 1281–1288.