

SELF-SUPERVISION MEETS BOOTSTRAP ESTIMATION: NEW PARADIGM FOR UNSUPERVISED RECONSTRUCTION WITH UNCERTAINTY QUANTIFICATION

Anonymous authors

Paper under double-blind review

ABSTRACT

Deep learning-based self-supervised reconstruction (SSR) plays a vital role in diverse domains, including unsupervisedly reconstructing magnetic resonance imaging (MRI). Current powerful methodologies for self-supervised MRI reconstruction usually rely on capturing the relationships between different views or transformations of the same data such as serving as inputs and labels respectively, which show notable influence from analogous approaches in computer vision. Although yielding somewhat promising results, their designs are often heuristic without deep insights into reconstructed object characteristics, and the analytical and mathematical principles of such methods are not expressive. This paper addresses these issues with a novel SSR paradigm, *BootRec*, that not only provides an explanation for self-supervised reconstruction but also facilitates the development of downstream algorithms. Self-supervised MRI reconstruction is modeled as error-oriented parameter estimation - Bootstrap estimation for SSR (*BootRec*). In *BootRec*, we demonstrate the mathematical equivalence between bootstrapping in a sample set and the commonly used re-undersampling operation for SSR. This insight is further incorporated into designing models to estimate errors of MRI SSR results without accessing labeled data. The estimation can further serve as the loss function for unsupervisedly training the models. Experiments show that our new paradigm *BootRec* enables advanced MRI reconstruction performance against other zero-shot methods. The code is available at <https://github.com/user19781945/rep10825984>.

1 INTRODUCTION

Magnetic resonance imaging (MRI) reconstruction receives continuous attention for its significance in medical imaging and challenges in often unsupervised settings due to costly labeling and obtaining ground truth. MRI reconstruction inherently requires a lengthy step of repeatedly collecting measurements in the frequency domain to fill the k -space before recovering the spatial signals using inverse Fourier transform (IFT). The advancements in techniques such as parallel imaging for acquiring signals and compressed sensing (CS) for reconstruction have provided approaches to reduce imaging time. Specifically, CS makes it possible to acquire fewer measurements than the Nyquist rate while reducing the aliasing artifacts (Donoho, 2006; Lustig et al., 2008).

The introduction of deep neural networks for deep learning (DL) to CS-MRI has also led to breakthroughs in a higher acceleration ratio and better reconstruction quality in MRI reconstruction (Chen et al., 2022; Wang et al., 2021; Lin & Heckel, 2022; Fabian et al., 2021). However, these DL methods, though powerful, have several challenges in further applications. The first problem is that supervised DL training demands numerous labeled training data. In the situation of MRI reconstruction, it means that enough fully sampled images must be provided, which is impossible in many situations. Another important shortcut is the black-box nature of DL models, making the reconstruction lack explanation and uncertainty estimation. Hence, it is hard to evaluate the risk in real-world medical practice when doctors need to make critical decisions according to the images (Edupuganti et al., 2020).

We propose a new paradigm of Bootstrap estimation for self-supervised reconstruction (BootRec) of MRI. BootRec models MRI SSR as a parameter estimation problem, and applies Bootstrap estimation to quantify the errors. The learning target is then shifted to minimize the estimated mean squared error (MSE) between the reconstructed fully sampled images and the unknown ground truth. Summary of different pipelines and the insights of our modeling are in Figure 1.

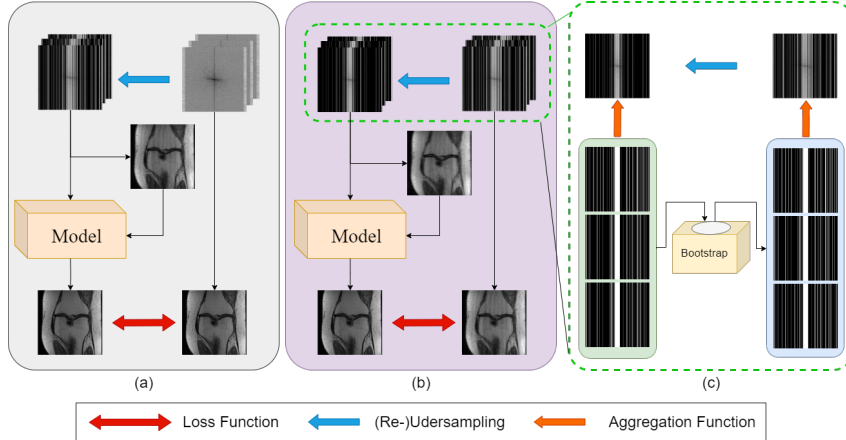


Figure 1: Demonstration of different pipelines of DL-based reconstruction models. All self-supervised methods incorporate some kind of **re-undersampling**. (a) Supervised training with paired fully sampled images as labels. (b) self-supervision via data undersampling pipeline. (c) Insights of modeling re-undersampling as Bootstrap. Only masks of virtual sample sets are plotted for simplification.

The main contributions of our BootRec paradigm are summarized as follows: (1) We construct a new framework that models MRI SSR as a parameter estimation problem. (2) We demonstrate the equivalence between Bootstrap sampling and re-undersampling in certain conditions. (3) We propose using Bootstrap MSE estimation as uncertainty quantification for SSR. (4) We propose new algorithms to train self-supervised models and achieve advanced results.

The notations used in the paper are summarized in Appendix A for reference.

2 BACKGROUND & RELATED WORK

2.1 DEEP-LEARNING-BASED RECONSTRUCTION FOR MRI

The imaging process of parallel CS-MRI in one coil can be formulated as Equation 1, where \mathbf{y} represents the acquired k-space data, \mathbf{x} is the spatial anatomy data, \mathbf{r} is the noise, \mathcal{F} is the Fourier transform, \mathbf{U} is the 0-1 valued matrix indicating the sampling points in k-space (which is called measurement matrix in CS), and \mathbf{C} is the coil sensitivity. Note that \mathbf{x} , \mathbf{y} and \mathbf{r} should be multi-dimensional values. The notions in the paper are summarized in Appendix A. For simplification and to be consistent with other references, we represent them as flattened vectors.

$$\mathbf{y} = \mathbf{U}\mathcal{F}\mathbf{C}\mathbf{x} + \mathbf{r} \quad (1)$$

For simplification in the later analysis and without influencing the conclusion, we will skip the combination of multiple coils and ignore the noise \mathbf{r} as it is usually modeled as Gaussian noise with mean of zero. We also ignore the coil sensitivity or merge them into \mathbf{x} , thus obtaining a simplified equation of CS-MRI:

$$\mathbf{y} = \mathbf{U}\mathcal{F}\mathbf{x} \quad (2)$$

Given the acquired \mathbf{y} , the reconstruction is built as a reverse problem to recover \mathbf{x} using some reconstruction model. Traditionally, the reconstructor is iterative based on CS theory, while in deep-learning-based methods, the model can be a neural network parameterized by θ (Chen et al., 2022; Yang et al., 2016). We represent any reconstruction model as Equation 3.

$$\hat{\mathbf{x}} = f(\mathbf{y}, \mathbf{U}) \quad (3)$$

2.2 SELF-SUPERVISED TRAINING OF RECONSTRUCTION MODELS

Early trials of unsupervised training of reconstruction models implement dictionary learning and other classical algorithms in CS (Majumdar, 2018; Singhal & Majumdar, 2020). Other methods include leveraging unpaired fully-sampled data (Oh et al., 2020; Chung et al., 2021; Korkmaz et al., 2022) and Deep Image Prior (DIP) (Ulyanov et al., 2018). Benefiting from the success of self-supervised methods in computer vision like contrastive learning (Chen et al., 2020) and masked Autoencoder (He et al., 2022), self-supervised training in MRI reconstruction has made progress in recent years and surpassed other methods (Zhou et al., 2022; Yaman et al., 2020; Zou et al., 2022; Wang et al., 2022b).

The basic pipeline of self-supervised model is in Figure 1. As shown in Equation 4t, reconstruction is conducted on the re-undersampled measurements y , and the basic form loss of is $\mathcal{L}(\hat{x}^R, \hat{x})$. The key points are the design of re-undersampling masks U^R and loss functions. In different models (Wang et al., 2022b; Yaman, 2022), different kinds of sampling methods (uniform, Gaussian, etc.) and ratios of re-undersampling are proposed and evaluated. The loss function of the self-supervision mainly comes from the undersampled k-space not being selected in re-undersampling, which can be defined in the frequency or spatial domain (Jafari et al., 2021; Senouf et al., 2019), with a wide range of choices from imaging processing.

$$\hat{x}^R = f(U^R y, U) \quad (4)$$

Generally speaking, the explorations of effective self-supervised algorithms for CS-MRI reconstruction are heuristic. Instead, Bootrec will try to provide a methodology and explanation for this field.

2.3 UNCERTAINTY QUANTIFICATION OF MRI RECONSTRUCTION

DL models show impressive advantages in many fields with a major concern about their result’s reliability, such as the hallucination of large language models (OpenAI, 2023). In MRI reconstruction, a concern is that DL models may “imagine” the anatomies and mislead the diagnosis. Uncertainty Quantification (UQ) can ameliorate the problem by providing “confidence level” of the results, enabling decision-makers aware of the risk of unauthentic imaging (Gawlikowski et al., 2021), and doctors can choose to conduct further examinations for results of high uncertainty.

Derived from its origin, uncertainty of reconstruction can be divided into two categories (Kendall & Gal, 2017), aleatoric uncertainty stemming from the ill-posedness of the problem and epistemic uncertainty from the uncertainty of model parameters. The notion of uncertainty is also to be clarified. In the field of image tasks, the variance of the result is widely used, and other choices include quantiles and entropies (Angelopoulos et al., 2022). In MRI reconstruction and other image regression tasks, the residual error of the prediction also made notable progresses (Wang et al., 2022a).

Uncertainty quantification has been considered in the community of computational imaging. In the field of MRI reconstruction, Edupuganti et al. (2020) leverages the variational Autoencoder (VAE) to convert the deterministic result to be probabilistic. Schlemper et al. (2018b) and Ekmekci & Cetin (2022) builds a Bayesian neural network (BNN) and models the inherent uncertainty with a Gaussian distribution. A main limitation of existing methods is that supervised training is needed for the quantification so they cannot be applied to unsupervised models.

We find Bootstrap estimated MSE can be viewed as UQ to some extent, which models the aleatoric uncertainty from (re)-undersampling well. Further experiments are conducted to assess the quantification.

3 MODELING RECONSTRUCTION AS PARAMETER ESTIMATION

The BootRec framework consists of the following modules: (1) *aggregation function* for preprocessing and wrap reconstruction model as parameter estimator; (2) *a virtual sample set* as a mathematical tool to map a single observation to a sample set; (3) *pseudo resampling trick* to map Bootstrap sampling of a virtual sample set to re-undersampling of measurement; and (4) *training algorithms* for the specific loss function based on bootstrapping. These are detailed below.

3.1 DISTRIBUTION OF MRI ACQUISITION OBSERVATION

Firstly, we assume the sampling mask \mathbf{U} obeys a multivariate Bernoulli distribution where each variable is independent, as Equation 5. We also make a constraint that all positions keep the possibility to be sampled, that is, $\mathbf{P}_{\mathbf{U}_i} > 0$ for any given position i .

$$\mathbf{U} \sim \mathcal{B}(\mathbf{U}; 1, \mathbf{P}_{\mathbf{U}}) \quad (5)$$

BootRec initially operates by training a separated model for each data point (zero-shot reconstruction(Yaman, 2022)) and we'll discuss more general situation in Section 5.3. In the zero-shot case, the target of the i_{th} reconstruction is fixed as $x^{(i)}$, so we directly use \mathbf{x} as $x^{(i)}$ in the following derivation.

With former Equation 2, the randomness from the mask is introduced so the sampled k-space data can also be viewed as random variables. Usually, the sampling mask and the acquired k-space data are provided and processed simultaneously in CS, so we define the observation as $\mathbf{s} = (\mathbf{y}, \mathbf{U})$ for convenience. The distribution of \mathbf{s} can be fully parameterized by \mathbf{x} and $\mathbf{P}_{\mathbf{U}}$, written as $p(\mathbf{s}; \mathbf{x}, \mathbf{P}_{\mathbf{U}})$. The reconstruction task in Equation 3 can then be viewed as estimating the parameter of $p(\mathbf{s})$ given observations of \mathbf{s} .

3.2 ESTIMATOR CONSTRUCTION WITH AGGREGATION FUNCTION

To formulate estimators from the reconstruction models, the main distinction is that observations are processed individually and independently without forming a set, as in Equation 3, so we propose *aggregation function* as an adapter.

An aggregation function is a special mapping from observation sets $\{\mathbf{s}^{(1)}, \mathbf{s}^{(2)}, \dots, \mathbf{s}^{(n)}\}$ to a single observation $\mathbf{s}^* = (\mathbf{y}^*, \mathbf{U}^*)$ and serves as a component of the estimator. The output can then be directly passed to any existing reconstruction model.

In the situation of n (positive integer) independent observations $\{\mathbf{s}^{(1)}, \mathbf{s}^{(2)}, \dots, \mathbf{s}^{(n)}\}$, the aggregation function is defined as Equation 6. Intuitively it takes the average observed value in positions being selected at least once and keeps the other positions zero-valued. A case of $n = 3$ is demonstrated in Figure 2.

$$h(\mathbf{s}^{(1)}, \mathbf{s}^{(2)}, \dots, \mathbf{s}^{(n)})_i = (\mathbf{U}^* \mathbf{y}, \mathbf{U}^*)_i = (\mathbf{y}^*, \mathbf{U}^*)_i \quad (6)$$

$$\mathbf{U}_i^* = \begin{cases} 1 & (\mathbf{U}^{(1)} + \mathbf{U}^{(2)} + \dots + \mathbf{U}^{(n)})_i \neq 0 \\ 0 & (\mathbf{U}^{(1)} + \mathbf{U}^{(2)} + \dots + \mathbf{U}^{(n)})_i = 0 \end{cases} \quad (7)$$

After the aggregation function, it is obvious that \mathbf{y}^* will still be k-space data in which all the selected measurements and gradients in \mathbf{U}^* remain the same as that of the corresponding position in \mathbf{y} , as a normal masking operation. Also, if only one observation is acquired, the aggregation function will be transparent and will not be adjusted.

Aggregation function can be composed by any reconstruction method to form a new reconstruction method $f_{AF} = f \circ h$. The new function can take the sample set from the distribution of $P(\mathbf{s})$ and serve as an estimator of the parameter \mathbf{x} without modifying the reconstruction process defined by f .

$$\hat{\mathbf{x}} = f_{AF}(\mathbf{s}^{(1)}, \mathbf{s}^{(2)}, \dots, \mathbf{s}^{(n)}) \quad (8)$$

3.3 EQUIVALENT SAMPLE DISTRIBUTION

Under the definition of aggregation function, \mathbf{U}^* means a "selected at least once" matrix and \mathbf{U}^* obeys a multivariate Bernoulli distribution whose distribution parameter can be computed as Equation 10. The parameterized distribution of output observations is then $p(\mathbf{s}^*; \mathbf{x}, \mathbf{P}_{\mathbf{U}^*})$ correspondingly.

$$\mathbf{U}^* \sim \mathcal{B}(\mathbf{U}^*; 1, \mathbf{P}_{\mathbf{U}^*}) \quad (9)$$

$$\text{diag}(\mathbf{P}_{\mathbf{U}^*}) = \mathbf{I} - (\mathbf{I} - \text{diag}(\mathbf{P}_{\mathbf{U}}))^n \quad (10)$$

Observing the process of aggregation function, we notice that multiple sample sets may be mapped to the same observation. Given a sample set containing n identically distributed observations

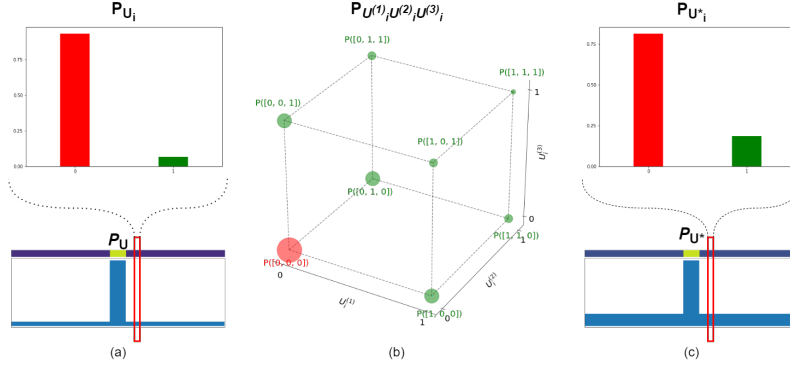


Figure 2: Demonstration of Aggregation Function with 3 observations (only the distribution of masks are presented for simplification). (a)/(c): The probability of being selected in different positions (P_U) and the distribution of a specific position being selected or not (P_{U_i}) before/after aggregation function. (b) The distribution of selections in the set of when size $n = 3$, and the color means the corresponding value after the aggregation function.

$\{\mathbf{s}^{(j)} | \mathbf{s}^{(j)} \sim p(\mathbf{s}), j = 0, 1, \dots, n\}$, if the result of adding it to the aggregation function satisfy that $h(\mathbf{s}^{(1)}, \mathbf{s}^{(2)}, \dots, \mathbf{s}^{(n)})_i = (\mathbf{y}^*, \mathbf{U}^*)_i$, we define $p(\mathbf{s})$ the n -cardinality *Equivalent Sample Distribution* of $p(\mathbf{s}^*)$, whose parameter P_U satisfies Equation 10. The two distributions are connected by the aggregation function. In Figure 2, if (c) visualizes the distribution of actual observations, then (a) shows an equivalent sample distribution when $n = 3$. Solving Equation 10 by setting P_U as unknowns, we get Equation 11, which computes the parameters of equivalent sample distribution.

$$\text{diag}(P_{U_i}) = \mathbf{I} - (\mathbf{I} - \text{diag}(P_{U^*}))^{1/n} \quad (11)$$

4 CONNECTING BOOTSTRAP WITH RE-UNDERSAMPLING

4.1 BOOTSTRAP ESTIMATION OF MULTIPLE OBSERVATIONS

For parameter estimation, the inference of the population is performed with the collected sample set of a certain size n . However, without reference to the population, the quality of the estimation cannot be computed. Bootstrap method solves the problem by sampling a new sample set of the same size n in the original sample set (with replacement) m times, and using the resampled sets (called Bootstrap sample sets) to form m Bootstrap estimations. The quality of the estimation with the original sample set can then be inferred by assessing the Bootstrap estimation with respect to the original estimation, which is accessible. With much more that can be studied in applying bootstrapping, we here only focus on the non-parameterized Bootstrap method and the Bootstrap estimation of MSE.

In the scale of our modeled reconstruction problem, we can represent the sample set of size n with $\{\mathbf{s}^{(1)}, \mathbf{s}^{(2)}, \dots, \mathbf{s}^{(n)}\}$ and the original estimation as $\hat{\mathbf{x}}$. The k_{th} Bootstrap sample set can be represented by $\{\mathbf{s}^{B_k(1)}, \mathbf{s}^{B_k(2)}, \dots, \mathbf{s}^{B_k(n)}\}$, and the corresponding estimation as Equation 12.

$$\hat{\mathbf{x}}^{B_k} = f_{AF}(\mathbf{s}^{B_k(1)}, \mathbf{s}^{B_k(2)}, \dots, \mathbf{s}^{B_k(n)}) \quad (12)$$

The MSE of the estimation $\hat{\mathbf{x}}$ can be estimated by bootstrapping using Equation 13. For MRI reconstruction, we can see that without reference to the fully sampled image \mathbf{x} , it is still possible to estimate the MSE of the reconstruction result.

$$\hat{mse}(\hat{\mathbf{x}}) = \frac{1}{m} \sum_{k=1}^{k=m} (\hat{\mathbf{x}}^{B_k} - \hat{\mathbf{x}})^2 \quad (13)$$

4.2 VIRTUAL SAMPLE SET AND PSEUDO RESAMPLING TRICK

In the previous section, we show that we can measure the quality of MRI reconstruction using Bootstrap method. However, in real-life applications, it is unrealistic to assume that there will be multiple observations to form a sample set of enough size to perform bootstrapping. In fact, often only one observation may be available in a specific scan. An intuitive method is to randomly generate a sample set with Equation 6 as a constraint or to assign the points to observations in the virtual sample simply uniformly. These methods will lead to high variance in computation with no prior knowledge leveraged. Instead, we propose to get the distribution of the observations by mapping the observation to a virtual sample set derived from the equivalent sample distribution defined in Section 3.3, which generates estimations equally distributed as a direct reconstruction given the single observation.

If the single observation follows distribution $p(s^*)$, the equivalent sample distribution is then $p(s)$ correspondingly, which forms the prior distribution of the observations in the virtual sample set of corresponding size n . However, given a specific observation $s^* = (\mathbf{y}^*, \mathbf{U}^*)$, the operation of “sampled at least once” is constrained so we should instead model the observations in the virtual sample set as a conditioned distribution, as formalized in Equation 14.

$$P(\mathbf{U}_i^V | \mathbf{U}_i^*) = \mathbf{1}_{\mathbf{U}_i^* = 1} Pr(\mathbf{U}_i^V = 1 | \mathbf{U}_i^* = 1) \quad (14)$$

We can calculate the probabilities with Bayes’ theorem and then use approximate values to help implementation, the result is in Equation 15 and details of derivation can be found in Appendix C. The distribution of virtual sample set elements is parameterized as $p(s^V; \mathbf{x}, \mathbf{P}_{U^V})$.

$$\mathbf{P}_{U_i^V} = Pr(\mathbf{U}_i^V = 1) = \begin{cases} 1 & \mathbf{P}_{U_i} = 1 \\ 1/n & \mathbf{U}_i^* \neq 0 \& \mathbf{P}_{U_i} \neq 1 \\ 0 & \mathbf{U}_i^* = 0 \end{cases} \quad (15)$$

With the distribution of observations in the virtual sample set, the distribution of the output of the aggregation function, marked as $p(s^{B*})$ can be computed with the same methods as Equation 10. As a result, we can skip sampling the virtual sample set and directly draw instances from $p(s^{B*})$ to get the results of the aggregation function, which is similar to the kernel trick in kernel methods, so we name it Pseudo Resampling Trick.

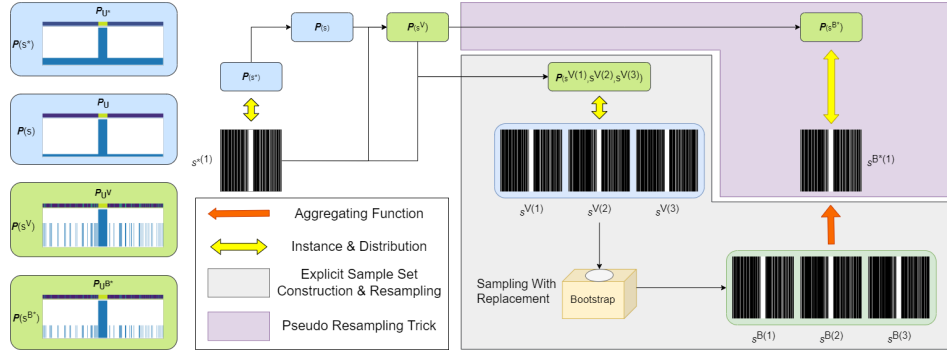


Figure 3: Demonstration of virtual sample set and pseudo resampling trick. The virtual sample set of an observation and its corresponding distribution are visualized. The gray area is the explicit construction of the virtual sample set and conducting Bootstrap sampling, while the purple area corresponds to the pseudo resampling trick.

4.3 SMARTER RE-UNDERSAMPLING AND TRAINING WITH BOOTREC

It’s easy to find that $\{i | \mathbf{U}_i^{B*} = 1\} \subseteq \{i | \mathbf{U}_i^* = 1\}$, so the process results in re-undersampling in k -space. On the contrary, for any given re-undersampling mask \mathbf{U}^R applied to the sample, the process can be described as the pseudo resampling virtual sample set, as long as the distribution of $\mathbf{U}^{R*} = \mathbf{U}^R \odot \mathbf{U}^*$ is the same as \mathbf{U}^{B*} . Based on this insight, new algorithms can be developed to

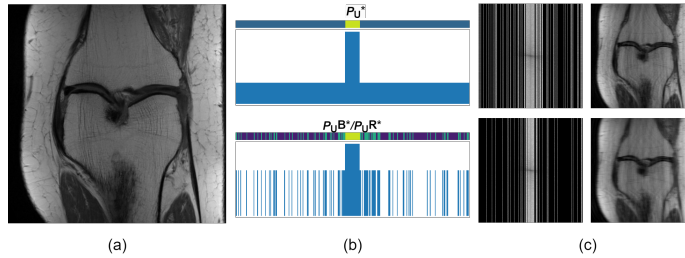


Figure 4: Example of dataset used. (a): A fully sampled image. (b): The upper is the distribution of P_{U^*} and the lower is $P_{U^{B^*}}$ or $P_{U^{R^*}}$. (c): The (re-)undersampled k-space and corresponding images.

implement the Bootstrap computation like Equation 13. The pseudo-code is provided in Appendix B.

A further step is to use estimated MSE as a proxy for the loss function in training learning-based reconstruction models. Algorithm 2 can derive the loss function as Equation 16 and Appendix D provide an example pipeline of training.

$$\mathcal{L}_{bootrec}(\hat{\mathbf{x}}, \mathbf{y}^*) = mse = \frac{1}{m} \sum_{k=1}^{k=m} (\hat{\mathbf{x}}^{B_k} - \hat{\mathbf{x}})^2 \quad (16)$$

The key attributes of our methodology include:

1. The re-undersampling pattern is derived from the distribution of sampling and is dynamic¹.
2. Spatial loss is used and the final target of $\hat{\mathbf{x}}$ is involved in the training process.
3. The self-supervision loss can be interpreted as errors estimated.

5 EXPERIMENTS

5.1 IMPLEMENTATION METHODS AND BASELINES

The experiment is conducted in fastMRI dataset (Zbontar et al., 2018), applying the setting of Wang et al. (2022b), where 232 volumes are split into 2D slices and divided with around 8:1:1 for training, validation, and testing, with a sampling ratio of 33% and a fixed mask for all data points. Coil sensitivity maps are built with ESPiRiT (Uecker et al., 2014). We use a large hyper-parameter $n = 1000$, 100 epochs for training models, and 100 iterations in each zero-shot epoch. More details are in Appendix F.

5.2 EVALUATING ESTIMATED MSE AS UNCERTAINTY QUANTIFICATION

To test the effectiveness of Algorithm 2 in an independent reconstructor, a DC-CNN model (Schlemper et al., 2018a) is trained with supervised MSE loss. The visualization of the results is in Figure 5. The data are collected with the test set so the model doesn't meet them in training or validation. We evaluate the correlation between the estimated MSE and the ground truth computed from the label and the prediction and the influence of different m values. The correlations of the estimated and ground truths mean that it's possible to identify hard sample with the estimation.

5.3 ESTIMATED MSE AS LOSS FUNCTION FOR SELF-SUPERVISED TRAINING

We test the capability of optimizing reconstruction model according to Bootstrap estimated MSE in the zero-shot reconstruction scenery (Yaman, 2022), where an untrained neural network is optimized

¹In fact, some intuition of selecting such masks can also be derived, such as the re-undersampling ratio should not be lower than $\lim_{n \rightarrow +\infty} 1 - (1 - 1/n)^n = 1 - \frac{1}{e}$, otherwise it would result in a negative cardinality of the virtual sample set. Another situation is the corresponding n is not integer.

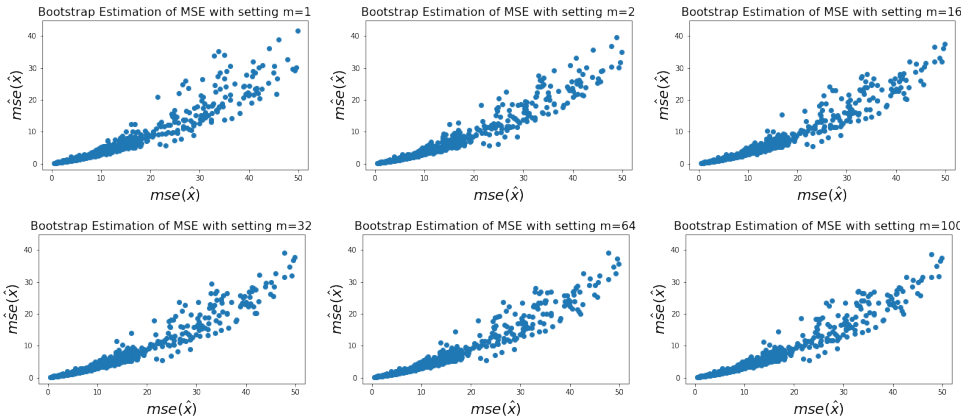


Figure 5: Estimated MSE with Algorithm 2 v.s. MSE computed with the ground truth. All figures show clear linear correlations between the estimations and ground truths, and the different values of m seem to have little influence on the correlation.

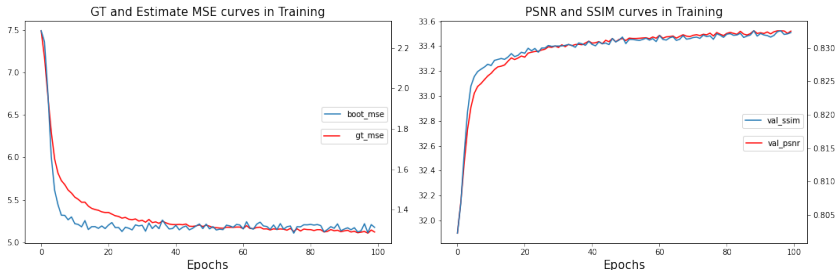


Figure 6: Training curves in zero-shot training. Estimated and GT MSE show consistent tendency, proving the effectiveness of optimizing Estimated MSE.

according to the acquired data with some loss function. Note the optimization of the loss function cannot be conducted directly with gradient descent, since the model will collapse. Accordingly, we make some special designs including stopping the gradient of the original estimation(Grill et al., 2020; Chen & He, 2021) and enforcing consistency on positions not sampled. The details of the implementation can be found in Appendix D. We set $m = 1$ in the experiment.

We use data from the fastMRI validation set. Our method is compared with other zero-shot methods like deep image prior (Ulyanov et al., 2018; Jafari et al., 2021; Senouf et al., 2019) and self-supervision via data undersampling (SSDU) (Yaman, 2022), where independent zero-shot models trained separately per image. Details of the models and results can be found in Appendix E. We also show the performances of SENSE reconstruction (Pruessmann et al., 1999), supervised DC-CNN methods, and the state-of-the-art (SOTA) unsupervised model from Wang et al. (2022b)². The quantitative results can be found in Figure 7 and visual examples are displayed in Figure 8.

Zero-shot model performances In Figure 7 we show that pure k-space loss functions failed to perform much better than the simple zero-filled SENSE method in our implementation, while our model shows clear advantages. We also probe the metrics during training in Figure 6 and found that the model continuously performs better as the optimization goes. Another positive finding is the MSE of the reconstruction shows a synchronous tendency as the Bootstrap estimated values, proving the accuracy of the estimation.

Handling multiple observations Another intriguing prospect is generalizing the loss to the training over multiple samples in a dataset. For now, our theory doesn't directly cover the multi-sample situation. If multiple samples are trained, the distribution is compositional and the to-be-estimated

²This model binds another more powerful backbone than DC-CNN in training so may have extra advantages.

parameter is transferred to be the parameter of the distribution of $\mathbf{x}^{(i)}$. Since the pseudo resampling and other tools are defined to be applied only for \mathbf{x} (here it only means a particular image), part of our theory needs to be re-explained and we leave it for future study. However, we demonstrate the effectiveness of directly applying Algorithm 2 in multiple samples. The model trained with multiple samples has better performances than the zero-shot models and even has a better structural similarity index (SSIM) than the SOTA unsupervised model while having a competitive peak signal-to-noise ratio (PSNR).

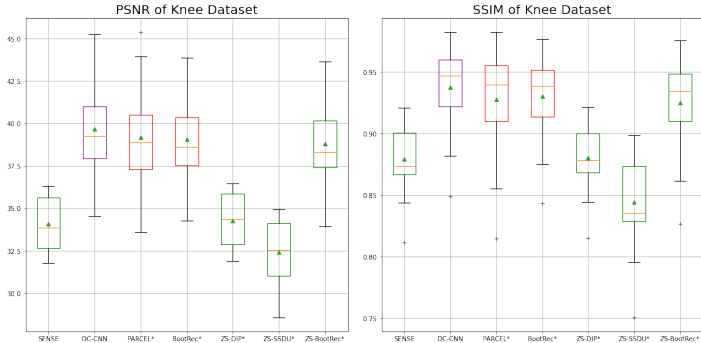


Figure 7: PSNR and SSIM of different models. Purple boxes indicate supervised models; red boxes show unsupervised models; green boxes indicate methods without training data.

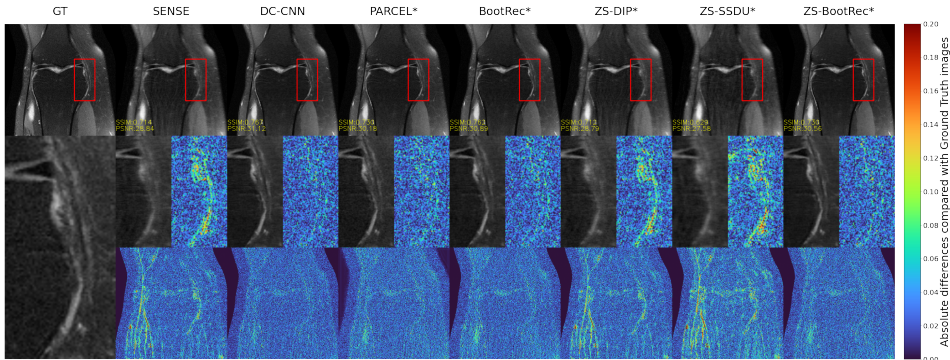


Figure 8: Visual examples of reconstruction of fastMRI validation set. The error maps are amplified by 5 times for better presentation.

6 CONCLUSION AND FUTURE WORK

In conclusion, as an attempt to provide a theoretical foundation and direct design of self-supervised learning algorithms, we propose a new paradigm for unsupervised compressed sensing MRI reconstruction. Unsupervised MRI reconstruction is modeled as parameter estimation, then we can wrap existing reconstruction methods to form estimators. Based on this insight, several designs including aggregation function, equivalent sample distribution, virtual sample set, and pseudo resampling trick are proposed to connect re-undersampling in self-supervised learning with Bootstrap sampling. Our flexible framework can not only estimate the MSE of arbitrary reconstructions without accessing ground truth images but also train self-supervised models for better performance.

We believe our paradigm may also inspire some new insights into transforming unsupervised learning. For example, if we define corresponding domains, all augmentations on self-supervised learning may be transformed to re-undersampling and thus can be analyzed with our framework. Also, the proposed aggregation function and estimators are not fully studied and have a large space for improvement with future efforts. The training process of the derived loss function may suffer from collapsing and exploding, which can also be further addressed for better solutions.

REFERENCES

- Anastasios N. Angelopoulos, Amit Pal Singh Kohli, Stephen Bates, Michael I. Jordan, Jitendra Malik, Thayer Alshaabi, Srigoikul Upadhyayula, and Yaniv Romano. Image-to-image regression with distribution-free uncertainty quantification and applications in imaging. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato (eds.), *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, pp. 717–730. PMLR, 2022. URL <https://proceedings.mlr.press/v162/angelopoulos22a.html>.
- Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey E. Hinton. A simple framework for contrastive learning of visual representations. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, volume 119 of *Proceedings of Machine Learning Research*, pp. 1597–1607. PMLR, 2020. URL <http://proceedings.mlr.press/v119/chen20j.html>.
- Xinlei Chen and Kaiming He. Exploring simple siamese representation learning. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, pp. 15750–15758. Computer Vision Foundation / IEEE, 2021. doi: 10.1109/CVPR46437.2021.01549. URL https://openaccess.thecvf.com/content/CVPR2021/html/Chen_Exploring_Simple_Siamese_Representation_Learning_CVPR_2021_paper.html.
- Yutong Chen, Carola-Bibiane Schonlieb, Pietro Lio, Tim Leiner, Pier Luigi Dragotti, Ge Wang, Daniel Rueckert, David Firmin, and Guang Yang. AI-Based Reconstruction for Fast MRI—A Systematic Review and Meta-Analysis. *Proceedings of the IEEE*, 110(2):224–245, February 2022. ISSN 0018-9219, 1558-2256. doi: 10.1109/JPROC.2022.3141367. URL <https://ieeexplore.ieee.org/document/9703109/>.
- Hyungjin Chung, Eunju Cha, Leonard Sunwoo, and Jong Chul Ye. Two-stage deep learning for accelerated 3d time-of-flight mra without matched training data. *Medical Image Analysis*, 71: 102047, 2021.
- D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.
- Vineet Edupuganti, Morteza Mardani, Shreyas Vasanawala, and John Pauly. Uncertainty quantification in deep mri reconstruction. *IEEE Transactions on Medical Imaging*, 40(1):239–250, 2020.
- Canberk Ekmekci and Mujdat Cetin. Uncertainty quantification for deep unrolling-based computational imaging. *IEEE Transactions on Computational Imaging*, 8:1195–1209, 2022.
- Zalan Fabian, Reinhard Heckel, and Mahdi Soltanolkotabi. Data augmentation for deep learning based accelerated MRI reconstruction with limited data. In Marina Meila and Tong Zhang (eds.), *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, pp. 3057–3067. PMLR, 2021. URL <http://proceedings.mlr.press/v139/fabian21a.html>.
- Jakob Gawlikowski, Cedrique Rovile Njiteucheu Tassi, Mohsin Ali, Jongseok Lee, Matthias Humt, Jianxiang Feng, Anna M. Kruspe, Rudolph Triebel, Peter Jung, Ribana Roscher, Muhammad Shahzad, Wen Yang, Richard Bamler, and Xiao Xiang Zhu. A survey of uncertainty in deep neural networks. *CoRR*, abs/2107.03342, 2021. URL <https://arxiv.org/abs/2107.03342>.
- Jean-Bastien Grill, Florian Strub, Florent Althé, Corentin Tallec, Pierre H. Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Ávila Pires, Zhaohan Guo, Mohammad Gheshlaghi Azar, Bilal Piot, Koray Kavukcuoglu, Rémi Munos, and Michal Valko. Bootstrap your own latent - A new approach to self-supervised learning. In Hugo Larochelle, Marc’Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin (eds.), *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020. URL <https://proceedings.neurips.cc/paper/2020/hash/f3ada80d5c4ee70142b17b8192b2958e-Abstract.html>.

- Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross B. Girshick. Masked autoencoders are scalable vision learners. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, June 18-24, 2022*, pp. 15979–15988. IEEE, 2022. doi: 10.1109/CVPR52688.2022.01553. URL <https://doi.org/10.1109/CVPR52688.2022.01553>.
- Ramin Jafari, Pascal Spincemaille, Jinwei Zhang, Thanh D. Nguyen, Xianfu Luo, Junghun Cho, Daniel Margolis, Martin R. Prince, and Yi Wang. Deep neural network for water/fat separation: Supervised training, unsupervised training, and no training. *Magnetic Resonance in Medicine*, 85(4):2263–2277, 2021. doi: <https://doi.org/10.1002/mrm.28546>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.28546>.
- Alex Kendall and Yarin Gal. What uncertainties do we need in bayesian deep learning for computer vision? In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pp. 5574–5584, 2017. URL <https://proceedings.neurips.cc/paper/2017/hash/2650d6089a6d640c5e85b2b88265dc2b-Abstract.html>.
- Yilmaz Korkmaz, Salman Ul Hassan Dar, Mahmut Yurt, Muzaffer Özbey, and Tolga Çukur. Unsupervised MRI reconstruction via zero-shot learned adversarial transformers. *IEEE Trans. Medical Imaging*, 41(7):1747–1763, 2022. doi: 10.1109/TMI.2022.3147426. URL <https://doi.org/10.1109/TMI.2022.3147426>.
- Kang Lin and Reinhard Heckel. Vision transformers enable fast and robust accelerated MRI. In Ender Konukoglu, Bjoern H. Menze, Archana Venkataraman, Christian F. Baumgartner, Qi Dou, and Shadi Albarqouni (eds.), *International Conference on Medical Imaging with Deep Learning, MIDL 2022, 6-8 July 2022, Zurich, Switzerland*, volume 172 of *Proceedings of Machine Learning Research*, pp. 774–795. PMLR, 2022. URL <https://proceedings.mlr.press/v172/lin22a.html>.
- M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly. Compressed sensing mri. *IEEE Signal Processing Magazine*, 25(2):72–82, 2008.
- Angshul Majumdar. An autoencoder based formulation for compressed sensing reconstruction. *Magnetic resonance imaging*, 52:62–68, 2018.
- Gyutaek Oh, Byeongsu Sim, Hyungjin Chung, Leonard Sunwoo, and Jong Chul Ye. Unpaired deep learning for accelerated MRI using optimal transport driven cycleGAN. *IEEE Trans. Computational Imaging*, 6:1285–1296, 2020. doi: 10.1109/TCI.2020.3018562. URL <https://doi.org/10.1109/TCI.2020.3018562>.
- OpenAI. GPT-4 technical report. *CoRR*, abs/2303.08774, 2023. doi: 10.48550/arXiv.2303.08774. URL <https://doi.org/10.48550/arXiv.2303.08774>.
- Klaas P Pruessmann, Markus Weiger, Markus B Scheidegger, and Peter Boesiger. Sense: sensitivity encoding for fast mri. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 42(5):952–962, 1999.
- Jo Schlemper, Jose Caballero, Joseph V. Hajnal, Anthony N. Price, and Daniel Rueckert. A deep cascade of convolutional neural networks for dynamic mr image reconstruction. *IEEE Transactions on Medical Imaging*, 37(2):491–503, 2018a. doi: 10.1109/TMI.2017.2760978.
- Jo Schlemper, Daniel Coelho de Castro, Wenjia Bai, Chen Qin, Ozan Oktay, Jinming Duan, Anthony N. Price, Joseph V. Hajnal, and Daniel Rueckert. Bayesian deep learning for accelerated MR image reconstruction. In Florian Knoll, Andreas K. Maier, and Daniel Rueckert (eds.), *Machine Learning for Medical Image Reconstruction - First International Workshop, MLMIR 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Proceedings*, volume 11074 of *Lecture Notes in Computer Science*, pp. 64–71. Springer, 2018b. doi: 10.1007/978-3-030-00129-2_8. URL https://doi.org/10.1007/978-3-030-00129-2_8.

- Ortal Senouf, Sanketh Vedula, Tomer Weiss, Alex Bronstein, Oleg Michailovich, and Michael Zibulevsky. Self-supervised learning of inverse problem solvers in medical imaging. In Qian Wang, Fausto Milletari, Hien V. Nguyen, Shadi Albarqouni, M. Jorge Cardoso, Nicola Rieke, Ziyue Xu, Konstantinos Kamnitsas, Vishal Patel, Badri Roysam, Steve Jiang, Kevin Zhou, Khoa Luu, and Ngan Le (eds.), *Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data*, pp. 111–119, Cham, 2019. Springer International Publishing. ISBN 978-3-030-33391-1.
- Vanika Singhal and Angshul Majumdar. Reconstructing multi-echo magnetic resonance images via structured deep dictionary learning. *Neurocomputing*, 408:135–143, 2020.
- Martin Uecker, Peng Lai, Mark J Murphy, Patrick Virtue, Michael Elad, John M Pauly, Shreyas S Vasanaawala, and Michael Lustig. Espirit—an eigenvalue approach to autocalibrating parallel mri: where sense meets grappa. *Magnetic resonance in medicine*, 71(3):990–1001, 2014.
- Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- Ke Wang, Anastasios Angelopoulos, Alfredo De Goyeneche, Amit Kohli, Efrat Shimron, Stella Yu, Jitendra Malik, and Michael Lustig. Rigorous uncertainty estimation for mri reconstruction. In *Proceedings of the Proceedings of the 30th Annual Meeting of ISMRM*, volume 749, 2022a.
- S. Wang, R. Wu, C. Li, J. Zou, Z. Zhang, Q. Liu, Y. Xi, and H. Zheng. Parcel: Physics-based unsupervised contrastive representation learning for multi-coil mr imaging. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 14(8):1–12, oct 2022b. ISSN 1557-9964. doi: 10.1109/TCBB.2022.3213669.
- Shanshan Wang, Taohui Xiao, Qiegen Liu, and Hairong Zheng. Deep learning for fast MR imaging: A review for learning reconstruction from incomplete k-space data. *Biomed. Signal Process. Control.*, 68:102579, 2021. doi: 10.1016/j.bspc.2021.102579. URL <https://doi.org/10.1016/j.bspc.2021.102579>.
- Burhaneddin Yaman. Zero-shot self-supervised learning for mri reconstruction. In *International Conference on Learning Representations*, 2022.
- Burhaneddin Yaman, Seyed Amir Hossein Hosseini, Steen Moeller, Jutta Ellermann, Kâmil Uğurbil, and Mehmet Akçakaya. Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data. *Magnetic resonance in medicine*, 84(6):3172–3191, 2020.
- Yan Yang, Jian Sun, Huibin Li, and Zongben Xu. Deep admm-net for compressive sensing MRI. In Daniel D. Lee, Masashi Sugiyama, Ulrike von Luxburg, Isabelle Guyon, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain*, pp. 10–18, 2016. URL <https://proceedings.neurips.cc/paper/2016/hash/1679091c5a880faf6fb5e6087eb1b2dc-Abstract.html>.
- Jure Zbontar, Florian Knoll, Anuroop Sriram, Matthew J. Muckley, Mary Bruno, Aaron Defazio, Marc Parente, Krzysztof J. Geras, Joe Katsnelson, Hersh Chandarana, Zizhao Zhang, Michal Drozdal, Adriana Romero, Michael G. Rabbat, Pascal Vincent, James Pinkerton, Duo Wang, Nafissa Yakubova, Erich Owens, C. Lawrence Zitnick, Michael P. Recht, Daniel K. Sodickson, and Yvonne W. Lui. fastmri: An open dataset and benchmarks for accelerated MRI. volume abs/1811.08839, 2018. URL <http://arxiv.org/abs/1811.08839>.
- Bo Zhou, Jo Schlemper, Neel Dey, Seyed Sadegh Mohseni Salehi, Kevin Sheth, Chi Liu, James S. Duncan, and Michal Sofka. Dual-domain self-supervised learning for accelerated non-cartesian mri reconstruction. *Medical Image Analysis*, 81:102538, 2022. ISSN 1361-8415. doi: <https://doi.org/10.1016/j.media.2022.102538>. URL <https://www.sciencedirect.com/science/article/pii/S1361841522001852>.
- Juan Zou, Cheng Li, Sen Jia, Ruoyou Wu, Tingrui Pei, Hairong Zheng, and Shanshan Wang. Self-collearn: Self-supervised collaborative learning for accelerating dynamic MR imaging. *CoRR*, abs/2208.03904, 2022. doi: 10.48550/arXiv.2208.03904. URL <https://doi.org/10.48550/arXiv.2208.03904>.

A NOTATIONS

Symbol	Description
\boldsymbol{x}	The fully-sampled spatial MRI image. (Ground truth for supervision. Inaccessible for self-supervision.)
\boldsymbol{y}	Undersampled k-space data computed with Equation 2. (The first-hand acquired data in medical practice.)
\boldsymbol{U}	Undersampling mask. (Indicating positions to fill in the acquisition of MRI.)
\boldsymbol{s}	The observation, composed of \boldsymbol{y} and \boldsymbol{U} . (Given information after MRI acquisition.)
$\hat{\boldsymbol{x}}$	The output of the reconstruction model with Equation 3.
\cdot^*	Values from one-time collected sample set containing only one observation.
$\cdot^{(i)}$	Values in the i_{th} observation of sample set containing multiple observations.
\cdot^B	Values from the Bootstrap resampled set. \cdot^{B_k} means the k_{th} set.
\cdot^V	Values from the virtual sample set.

B IMPLEMENTATION OF ALGORITHMS

Algorithm 1: Pseudo Resampling Function**Input:** Mask \boldsymbol{U}^* , Mask probability $\boldsymbol{P}_{\boldsymbol{U}^*}$, Size of virtual sample set n **Result:** Bootstrap observation mask \boldsymbol{U}^{B^*}

```

1 def main ( $\boldsymbol{U}^*$ ,  $\boldsymbol{P}_{\boldsymbol{U}^*}$ ,  $n$ ) :
2   calculate  $\boldsymbol{P}_{\boldsymbol{U}}$  according to Equation 11;
3   initialize  $\boldsymbol{P}_{\boldsymbol{U}^{B^*}}$  as a zeroed tensor with the same shape of  $\boldsymbol{P}_{\boldsymbol{U}^*}$ ;
4   fill  $\boldsymbol{P}_{\boldsymbol{U}^{B^*}}$  with Equation 15;
5   draw random variable  $\boldsymbol{V}$  with uniform distribution in  $[0, 1)$  and the same shape  $\boldsymbol{U}^*$ ;
6   return the 0-1 converted result of  $\boldsymbol{U}^{B^*} = toInteger(\boldsymbol{V} < \boldsymbol{P}_{\boldsymbol{U}^{B^*}})$ 

```

Algorithm 2: Bootstrap Estimation MSE of Reconstruction**Input:** Observed k-space \boldsymbol{y} , Mask \boldsymbol{U}^* , Mask probability $\boldsymbol{P}_{\boldsymbol{U}^*}$, Size of virtual sample set n ,
Number of Bootstrap samples m , Reconstructor f **Result:** Estimated MSE l_{mse}

```

1 compute the original estimation  $\hat{\boldsymbol{x}} = f(\boldsymbol{y}, \boldsymbol{U}^*)$ ;
2  $boot\_mse\_list \leftarrow \emptyset$ ;
3 for  $k \leftarrow 1$  to  $m$  do
4   get resampling mask  $\boldsymbol{U}^{B_k}$  with Algorithm 1;
5   resample the k-space with  $\boldsymbol{y}^{B_k} = \boldsymbol{U}^{B_k}$ ;
6   compute the corresponding estimation  $\hat{\boldsymbol{x}}^{B_k} = f(\boldsymbol{y}^{B_k}, \boldsymbol{U}^{B_k})$ ;
7   compute the mean square error between the estimation and the original estimation
    $m\hat{se}^{B_k} = (\hat{\boldsymbol{x}}^{B_k} - \hat{\boldsymbol{x}})^2$ ;
8   append  $m\hat{se}^{B_k}$  to  $boot\_mse\_list$ ;
9 average the values to get the Bootstrap estimation of MSE:  $mean(boot\_mse\_list)$ ;

```

C DERIVATION OF DISTRIBUTION OF VIRTUAL SAMPLE SET OBSERVATIONS

From an intuitive aspect, for a given position i , the size of the virtual sample set n and the corresponding equivalent sample distribution parameter P_{U_i} , the times of the position being selected in n observations in the virtual sample set, represented as C_i , follows a Binomial distribution $\mathcal{B}(n, P_{U_i})$. Thus, with Bayes' Theorem, the conditional probability can be computed in Equation 17.

$$\begin{aligned}
& Pr(\mathbf{U}_i^V = 1 | \mathbf{U}_i^* = 1) \\
&= \frac{Pr(\mathbf{U}_i^* = 1 | \mathbf{U}_i^V = 1) Pr(\mathbf{U}_i^V = 1)}{Pr(\mathbf{U}_i^* = 1)} \\
&= \frac{Pr(\mathbf{U}_i^* = 1 | \mathbf{U}_i^V = 1) Pr(\mathbf{U}_i^V = 1)}{1 - (1 - P_{U_i})^n} \\
&= \frac{Pr(\mathbf{U}_i^V = 1)}{1 - (1 - P_{U_i})^n} \\
&= \frac{Pr(\mathbf{U}_i^V = 1 | C_i > 0) Pr(C_i > 0) + Pr(\mathbf{U}_i^V = 1 | C_i = 0) Pr(C_i = 0)}{1 - (1 - P_{U_i})^n} \\
&= \frac{Pr(\mathbf{U}_i^V = 1 | C_i = k) Pr(C_i > 0)}{1 - (1 - P_{U_i})^n} \\
&= \frac{\sum_{k=1}^{+\infty} Pr(\mathbf{U}_i^V = 1 | C_i = k) Pr(C_i = k)}{1 - (1 - P_{U_i})^n} \\
&= \frac{\sum_{k=1}^n \frac{k}{n} Pr(C_i = k)}{1 - (1 - P_{U_i})^n} \\
&= \frac{\sum_{k=1}^n \frac{k}{n} \binom{n}{k} P_{U_i}^k P_{U_i}^{n-k}}{1 - (1 - P_{U_i})^n}
\end{aligned} \tag{17}$$

The computation is too complicated and hard to use, so we provide a simplified version by analyzing three special cases:

1. $C_i = 0$, which means that the position is not selected in all observations, thus U^V will always be zero.
2. $P_{U_i} = 1$, common in the MRI acquisition to keep more low-frequency information. In this case, every observation in the virtual sample set contains the position.
3. P_{U_i} is very small. In this case, the probability of being selected again ($C_i > 1$) in the observations is small and can be neglected. If $C_i \neq 0$, we can assume $U^V \sim \mathcal{B}(U^V; 1, 1/n)$.

Note that only $P_{U_i^*}$ is given and P_{U_i} is computed with Equation 11, and n is a hyper-parameter that can be set manually, so for better approximation we can choose larger n when $P_{U_i^*}$ is not small enough. As a result, the third case can actually cover all halfway situations by setting proper n . Some examples of C_i distributions are shown in Figure 9. It can be found that the conditional distributions (golden bars) have a much higher probability in $C_i = 1$, which means nearly all sampled positions are sampled only once.

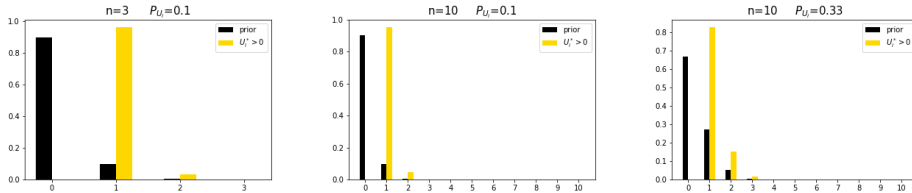


Figure 9: Examples of C_i 's distribution with different n and $P_{U_i^*}$. The golden bars show distributions when $C_i > 0$ is conditioned. The corresponding P_{U_i} is around 0.0345, 0.0105 and 0.0393.

Combining the situations, we can get the distribution of the mask.

$$P_{U_i^V} = Pr(\mathbf{U}_i^V = 1) = \begin{cases} 1 & P_{U_i} = 1 \\ 1/n & U_i^* \neq 0 \& P_{U_i} \neq 1 \\ 0 & U_i^* = 0 \end{cases} \quad (18)$$

D THE PIPELINE OF BOOTSTRAP SELF-SUPERVISED TRAINING

By referring to Algorithm 2, the Bootstrap estimated MSE can serve as a loss function for self-supervised training. The PyTorch-style code snippet of the training process is shown below. We use stop-gradient and inverse consistency to enable and stabilize the training.

```

1 n = 5
2 m = 20
3
4 with torch.no_grad():
5     x_organial=model(y*mask,mask)
6     y_original=fft(y_organial)
7 for _ in range(m):
8     re_mask=generate_re_mask(mask,p_maskin,n)
9     x_bootstrap=model(y*re_mask,re_mask)
10    y_bootstrap=fft(x)
11    x_bootstrap=ifft(y_bootstrap*(~mask)+y_organial*mask)
12    loss=nn.MSELoss(x_bootstrap,x_organial)/m
13    loss.backward()
14 optimizer.step()

```

E BRIEFS OF OTHER ZERO-SHOT MODELS

Besides the basic SENSE reconstruction that zero-fills the missing k-space, the differences between the models mainly come from the different loss functions. All the models are based on the same backbone structure of 8-layer DC-CNN. Note that due to the memory limit, the weight of the model is shared among the z-axis.

For the DIP models, the loss function is limited to keeping the consistency of all the knowing positions. Due to the data consistency layers in the model, the loss in k-space will always be 0 and nothing can be learned, so the loss is defined in the image domain as Equation 19, where f_{SENSE} is the SENSE reconstruction.

$$\mathcal{L}_{DIP}(\hat{\mathbf{x}}, \mathbf{y}^*) = mse(f_{SENSE}(\mathbf{y}^*), \hat{\mathbf{x}}) \quad (19)$$

The SSDU model applies re-undersampling in training and defines a normalized $l_1 - l_2$ loss in k-space as Equation 20. As for the super-parameters, we choose $K = 10$ and the re-undersampling ratio of 40% following the optimal setting of original paper(Yaman et al., 2020), and a uniform re-undersampling strategy is used. Also, Gaussian re-undersampling is conducted as described in the paper instead of deriving re-undersampling masks based on virtual sample set. The settings are modified to follow our basic setting of sharing weights among different slices.

$$\mathcal{L}_{SSDU}(\hat{\mathbf{x}}, \mathbf{y}^*) = \frac{\|\mathbf{y}^* - \mathcal{F}\hat{\mathbf{x}}\|_2}{\|\mathbf{y}^*\|_2} + \frac{\|\mathbf{y}^* - \mathcal{F}\hat{\mathbf{x}}\|_1}{\|\mathbf{y}^*\|_1} \quad (20)$$

F IMPLEMENTATION METHODS

Given a reconstructor f and an acquisition \mathbf{y}^* and \mathbf{U}^* of the spatial anatomy \mathbf{x} , a reconstruction and estimation result can be obtained by $\hat{\mathbf{x}} = f(\mathbf{y}^*, \mathbf{U}^*) = f_{AF}((\mathbf{y}^*, \mathbf{U}^*))$. MSE and its provide a simple but effective evaluation of the differences between $\hat{\mathbf{x}}$ and the target \mathbf{x} .

To conduct the bootstrapping in reconstructing a single MRI observation, although it may be complicated to state the transformation of the problem, the resulting solution is rather simple and intuitive. The construction of a virtual sample set and the computation of aggregation function become invisible parts of the pipeline. The only extra thing we need to do is to compute a re-undersampling mask based on the selected cardinality n , the actual mask U and its distribution parameter P_U . We provide the pseudo code of the computation in Algorithm 1. As bootstrapping enables the computation of various statistic properties without acquiring more sample sets, statistics like MSE can be estimated with only a single observation of partial measurements acquired CS MRI. The pseudo code of unsupervised computation of MSE can be found in Algorithm 2 based on Equation 13.

It can be seen that besides the input and the reconstructor, the most important hyper-parameters of the algorithms are the cardinal of the virtual sample set n , which decides the ratio and positions of re-undersampling and the times of Bootstrap sampling m , which increases the accuracy of Bootstrap estimators as it grows.

G QUANTITATIVE COMPARISON OF UNCERTAINTY QUANTIFICATION

With residual errors as UQ notion, we compare our methods with residual magnitude prediction (Angelopoulos et al., 2022), a method that directly enables a separated prediction head to predict the errors and we name it Residual Regression (RR) for simplification. The model is trained on fastMRI single-coil dataset with the same setting from the original paper except for a smaller batch size (48 v.s. 78) due to the limited GPU memory. After training, we directly apply Algorithm 2 of BootRec in the trained model to get our quantification, and square the predictions from RR. We also apply Stein’s Unbiased Risk Estimator (SURE) to estimate the trained model’s MSE, with settings from (Edupuganti et al., 2020).

Evaluation of UQ is still an open problem, with various metrics focusing on different notions and aspects. Considering the task, we conduct inter-image evaluations and intra-image evaluations. For inter-image evaluation, we average the pixels’ estimated MSE for every image and collect the ground truth MSE (GT MSE) of the reconstruction. The visualization can be found in Figure 10 and the quantitative results of correlation coefficients are summarized in Table 1. It can be seen that BootRec estimated MSE continuously presents better coefficients compared with predictions from RR head and has the highest values in Spearman and Kendall Coefficients. For intra-image evaluation, the distribution of the errors is more important than the absolute values as it indicates areas of hallucination. For the inter-image evaluation, we display an example from validation set for comparison of error distributions in Figure 11. The GT MSE shows high error levels in border areas, which is reasonable as the pixel values change rapidly, and the Bootstrap estimated error shows a similar distribution. The RR and SURE error maps put too much uncertainty inside the tissue, especially for the RR results. SURE estimations performed well in inter-image correlation but failed in inter-image distribution of errors, with one possible explanation that the residual variance of specific position is hard to compute.

	Pearson	Spearman	Kendall
RR	0.866	0.864	0.692
SURE	0.941	0.752	0.561
Bootstrap	0.888	0.920	0.756

Table 1: Correlation Coefficients computed corresponding to Figure 10. All values have a negligible p-value (we get zero in computation). The Bootstrap estimator achieves best correlation in Spearman and Kendall Coefficients.

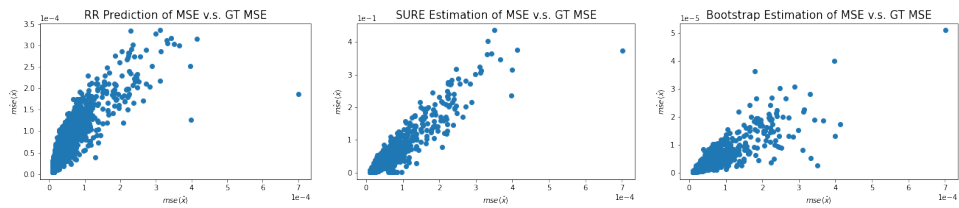


Figure 10: Inter-image Correlation of predicted/estimated MSE in the validation set.

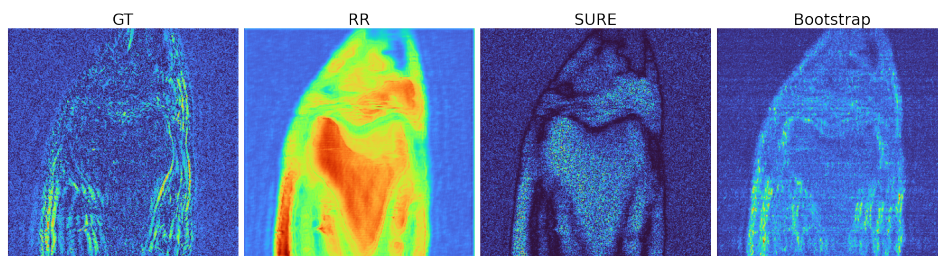


Figure 11: Intra-image distribution of predicted/estimated errors in the validation set (square roots are computed for MSE results). Note that the values are normalized across the image.