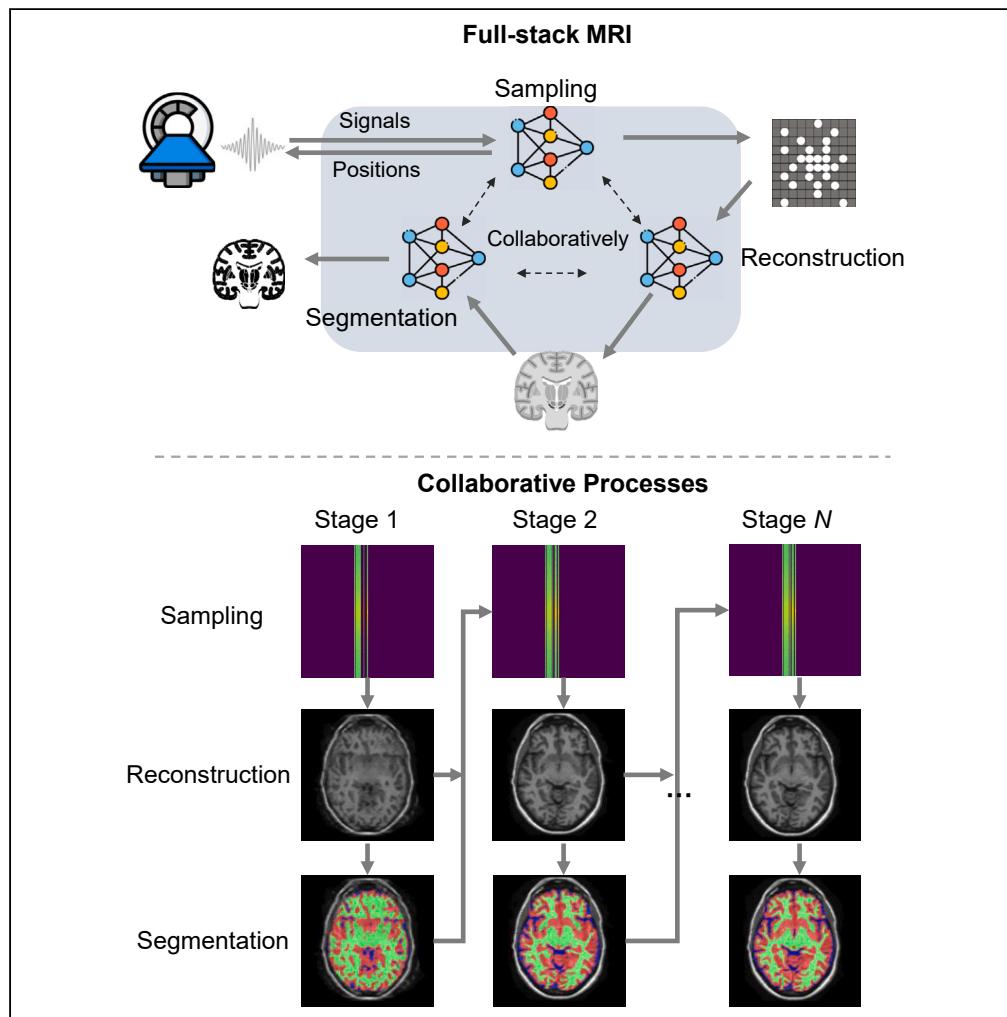


Article

Promoting fast MR imaging pipeline by full-stack AI



Zhiwen Wang,
Bowen Li, Hui Yu,
..., Jiliu Zhou,
Hongming Shan, Yi
Zhang

hmshan@fudan.edu.cn (H.S.)
yzhang@scu.edu.cn (Y.Z.)

Highlights

A novel AI paradigm for the full-stack fast MRI imaging

One network to simultaneously address sampling, reconstruction, and segmentation

Semantic information can be used to guide the acquisition of k-space raw data

Significant performance benefits for all tasks validated on multiple MRI datasets

Wang et al., iScience 27,
108608
January 19, 2024 © 2023 The
Author(s).
<https://doi.org/10.1016/j.isci.2023.108608>



Article

Promoting fast MR imaging pipeline by full-stack AI

Zhiwen Wang,¹ Bowen Li,¹ Hui Yu,¹ Zhongzhou Zhang,¹ Maosong Ran,¹ Wenjun Xia,¹ Ziyuan Yang,¹ Jingfeng Lu,² Hu Chen,¹ Jiliu Zhou,¹ Hongming Shan,^{3,*} and Yi Zhang^{2,4,*}

SUMMARY

Magnetic resonance imaging (MRI) is a widely used imaging modality in clinics for medical disease diagnosis, staging, and follow-up. Deep learning has been extensively used to accelerate k-space data acquisition, enhance MR image reconstruction, and automate tissue segmentation. However, these three tasks are usually treated as independent tasks and optimized for evaluation by radiologists, thus ignoring the strong dependencies among them; this may be suboptimal for downstream intelligent processing. Here, we present a novel paradigm, full-stack learning (FSL), which can simultaneously solve these three tasks by considering the overall imaging process and leverage the strong dependence among them to further improve each task, significantly boosting the efficiency and efficacy of practical MRI workflows. Experimental results obtained on multiple open MR datasets validate the superiority of FSL over existing state-of-the-art methods on each task. FSL has great potential to optimize the practical workflow of MRI for medical diagnosis and radiotherapy.

INTRODUCTION

Magnetic resonance imaging (MRI) is a noninvasive diagnostic imaging technique that enables the study of low-contrast soft tissue structures without imposing harmful radiation risks. However, the long acquisition time of MRI results in increased costs, patient discomfort, motion artifacts, and some time-restricted scenarios. Many clinical indicators can be imaged with compressed sensing-based MRI (CS-MRI) to accelerate the imaging process without significantly compromising the image quality or diagnostic accuracy.¹ In fact, minimizing the amount of required MRI-measured data is a commonly accepted practice, and CS-MRI has been a hot research topic in the medical imaging field for almost two decades.^{2–6} However, a reduction in the measured data might lead to missed important anatomical information and can introduce artifacts into the reconstructed images, which may adversely jeopardize the subsequent analysis or diagnosis if these problems are not overcome.

Benefiting from rapid developments of both algorithms and computer hardware, significant advances have emerged from intelligent fast imaging^{5–7} and accurately automated medical image analysis^{8–10} via deep learning¹¹ (DL). Specifically, magnetic resonance (MR) images need to be measured partially and reconstructed by scanners or processed and analyzed by servers. While most images in the past are sampled and reconstructed for human examination visually, in the era of intelligent computing, an increasing number of visual images can simultaneously serve human perception and power machine vision-based intelligent systems. Currently, existing imaging methods are mainly designed for either medical image reconstruction or medical image analysis, leading to fragmented visual processing stages.

Traditionally, various DL reconstruction paradigms have been proposed for imaging pipelines due to their data-driven superiority. These DL frameworks can be categorized into the following categories: (1) CS-MRI,¹² (2) compressed sampling learning^{13–16} (CSL), and (3) compressed sensing-based multitask learning^{17–20} (CS-MTL). DL-based CS-MRI methods form hand-crafted trajectories that are developed to achieve the desired visual fidelity; then, the reconstructed images can be fed into machine analysis tasks. In this way, artifacts and missing measurements can be visually processed in the data or/and image domains, yet data incompatibility between hand-crafted trajectories and scene-specific scans can be encountered, resulting in redundant sampling. Over the past year, we have started to see the development of learnable trajectory methods based on DL, called CSL approaches, which significantly reduce redundancy in such a way that reliable images can be reconstructed. However, the resultant methods formed by both CS-MRI and CSL often suffer from downstream tasks loss, inducing inaccurate medical image analyses and distorted high-level semantics that indicate structure distortion and detail loss in image reconstruction. DL-based CS-MTL paradigm aims to jointly learn reconstruction vision tasks and analysis vision tasks to link this disconnection, and this technique has achieved remarkable improvements. Unfortunately, such a CS-MTL paradigm that is conventionally trained with hand-crafted trajectories, also disconnects reconstruction–analysis vision from the measurement of raw data, thereby incurring fragmented observation

¹School of Computer Science, Sichuan University, Chengdu, Sichuan, China

²School of Cyber Science and Engineering, Sichuan University, Chengdu, Sichuan, China

³Institute of Science and Technology for Brain-inspired Intelligence, Fudan University, Shanghai, China

⁴Lead contact

*Correspondence: hmshan@fudan.edu.cn (H.S.), yzhang@scu.edu.cn (Y.Z.)

<https://doi.org/10.1016/j.isci.2023.108608>



and imaging stages for both visions, leading to a high level of risk for MR imaging pipelines. The main motivation of this study is to identify the opportunities and challenges involved in developing a novel full-stack MTL framework for collaborative and scalable modes to achieve improved imaging performance with efficient expression in the data domain for both image reconstruction and analysis vision, and establish a foundation for an MRI pipeline in which both visions are connected for all tasks by big data. Toward that goal, our method of choice is a novel full-stack MTL approach implemented with a novel neural network framework. In addition to the improved performance achieved on all tasks, the implication is clear: if the full-stack collaborative MTL approach performs favorably or comparably to the fragmented processing mechanism, the use of the full-stack paradigm in collaborative manner in the medical imaging pipeline might yield a practical advantage over traditional processing solutions. As the real scanning and imaging processes are nonintegrable, we simulate the full-stack approach to compare various DL methods for three traditional paradigms on fair ground. We stress that our intent is to demonstrate the superiority of full-stack collaborative MTL pipelines over traditional imaging paradigms implemented by mainstream methods instead of developing a specific neural network for improving the performance of specific tasks.

With the rapid development of MTL techniques, joint reconstruction–analysis vision tasks have recently produced great results for image processing pipelines. At present, MTL-based methods only start from reconstruction vision tasks and extend to machine vision tasks with a preset observation. In this context, a conventional MTL paradigm may cause a disconnection between the subsampled raw data and reconstruction–analysis vision, leading to rather limited priori information for neural networks. Here, we present the necessity of reconstruction–analysis collaborative sampling and formulate a new problem of full-stack learning (FSL). The novelty of this work is related to the exploration of a brand new MTL paradigm for a collaborative MR imaging pipeline by starting from observation to human perception to automated machine analysis in a DL manner so that all tasks in the imaging process effectively and efficiently benefit each other in the whole pipeline. As the imaging process progresses from sampling to reconstruction to analysis, the results of each step are suboptimally processed by different devices. To a significant degree, this process can be stepwise mimicked through DL from the scanner with a preset accelerated rate to the reconstructor to the analyst. A novel aspect of our approach is that it combines all task-specific networks into a single network. More precisely, we instantiate FSL by a trial exemplar architecture, a multitask learning network framework, and we simultaneously optimize learnable trajectories, reconstruction vision tasks and analysis vision tasks. This allows obtaining the optimal results of each step to minimize the risks involved in medical imaging. That is, each network performs at its own maximum in the imaging pipeline: (1) a subsampling trajectory fully and coherently expressed by reconstruction–analysis vision; (2) MR images with correct tissue details, plausible high-level semantics, and latent representations of machine vision; and (3) truly trustworthy machine predictions. To implement the FSL approach in a practical manner, reconstruction tasks and analysis tasks are much more efficient and accurate, and their interactions therefore do not have a great divide. This is both reconstruction- and analysis-friendly, and also expected to proactively provide insights into bridging cross-device and cross-domain research from radiology practice, computer vision, and machine learning when AI meets medical big data.

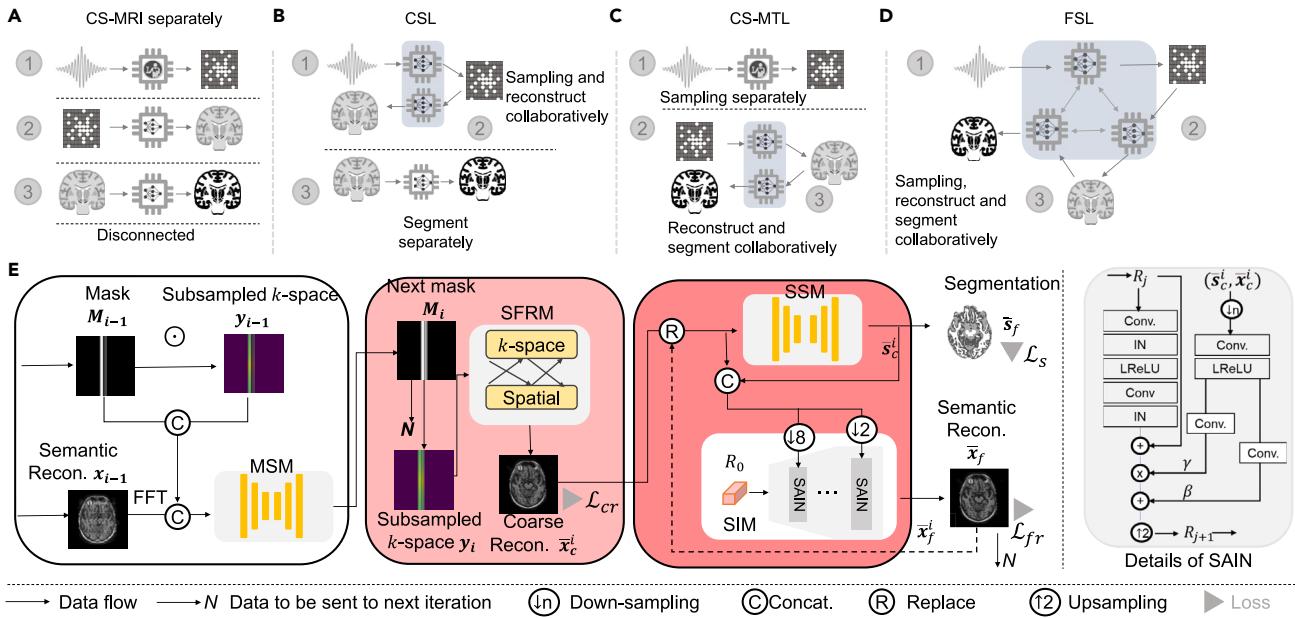
RESULTS

Concept of FSL

Conceptually, if data or signals are sufficiently sampled, an imaging workflow can be performed separately ([Figure 1A](#)). In CSL, signals are measured in a learnable way so that sampling and imaging can be optimized jointly by pixel references ([Figure 1B](#)), which can substantially enhance image reconstruction for examining and thereby improving the imaging results¹⁴ but poses disadvantages such as the loss of anatomical structures and machine vision representation data, as well as the use of a suboptimal workflow for the imaging pipeline. Multitask learning approaches^{17,18} have been developed ([Figure 1C](#)), wherein joint learning for reconstruction–analysis tasks is responsible for semantically aware inverse problem or analysis-friendly image reconstruction; however, a hand-crafted sampling measurement trajectory is maintained.

As an alternative, we introduce FSL ([Figure 1D](#)), which is instantiated by a trial exemplar multitask learning framework ([Figure 1E](#)) that optimizes the sampling, reconstruction and segmentation processes jointly in an end-to-end manner. Although our components possess a common design, the way they are assembled is very innovative in three aspects. (1) This method utilizes an easy yet novel end-to-end differentiable approach for this problem to fully explore the mutual influence among the sequential tasks and further improve the performance of each task simultaneously, which not only improves the visual quality of the subsampled MR images via semantic awareness but also boosts the accuracy of clinical tasks. (2) In the downstream task our training strategy can cope with both subsampled inputs and fully sampled inputs without any fine-tuning. (3) We observe that semantic sampling makes the reconstructor generalizable to other datasets with different distributions. We select both the CS-MRI and medical image segmentation tasks to exemplify the potential medical value of FSL.

In the instantiated framework of FSL, an autoencoder to predict next subsampling probability map and sampling operation ([Figure S1](#)) to binarize the map in a desired sampling rate is first employed to build a measurement sensing module (MSM) due to its differentiable compressive sampling method. A preset sampling rate trajectory can be predicted progressively by inputting the output to autoencoder, followed by a uniform distribution matrix that achieves Monte Carlo searching and a scaling layer that adjusts prediction to preset sampling rate. Although the processing paradigms for reconstruction–analysis vision have apparent differences, differentially, they are the same learning methods in deep neural networks. Naturally, we are able to utilize the differentiability of these three tasks to conduct measurement and jointly support reconstruction and analysis vision tasks. As a result, spatial-frequency reconstruction module (SFRM) and semantic segmentation module (SSM) ([Figure S3](#)) are then employed sequentially to realize the image reconstruction and analysis tasks. We also propose a semantic interaction module (SIM) ([Figure 1E](#)) to ensure that both features of reconstruction and segmentation interact sufficiently. SIM outputs a prior semantic reconstructed image, which is sent to MSM that can sense the spatial intense and semantic representation and predicts next

**Figure 1. Concept of FSL**

(A) Illustration of the concept of CS-MRI with sampling, medical image reconstruction and analysis tasks in separate, disconnected workflows.

(B) Principle of CSL.

(C) CS-MTL, where the reconstruction task is adapted to the analysis task but with sampling locations set by human experience.

(D) Principles of FSL with all tasks adapted to each other.

(E) Schematic of a trial exemplar FSL consisting of a collaborative multitask network that uses a differentiable approach. Imaging data are individually used by the submodules in a forward manner, and the sampling model collects reconstruction and analysis vision data to predict the next subsampled trajectory.

subsampled trajectory. This sequential and interactive method can also greatly improve the semantic information contained in the reconstructed images so that they can be efficiently examined by humans and machines.

Evaluation with 2D trajectories

We first compare our method with three conventional paradigms (Figures 1A–1C, detailed in the Compared methods section), including DL-based CS-MRI, CS-MTL, and CSL. We evaluate our proposed FSL under 5% and 10% sampling rates on the ACDC, OASIS, and NC-ISBI datasets.

As a conventional DL-based CS-MRI paradigm, CS-MRI₁ reconstructs images with low performance (Figures 2A–2C) and a limited de-aliasing effect, where significant remaining aliasing artifacts can be observed (Figures 2D, 2F, and 2H). This leads to a decline in the segmentation accuracy of this method (Figures 2A–2C, 2D, 2F, and 2H). On the other hand, CS-MRI₂ is trained with a dual-domain model to interact information between the spatial and k-space domains, and performs well in terms of the various evaluation metrics (Figures 2A–2C); however, its unreliable reconstructed structure and blurred tissues cause performance degradation on the segmentation task (Figures 2D, 2F, and 2H). Although CSL works quite well in terms of quantitative metrics (Figures 2A–2C) due to its data-adapted trajectories, it still suffers from a distorted structure (Figures 2D, 2F, and 2H), which affects the downstream tasks (Figures 2D, 2F, and 2H). Moreover, CSL suffers from severe overfitting (Figure 2B) on OASIS dataset. Guided by fully sampled segmentation-aware tasks that provide more semantics and machine vision priors, CS-MRI₁ and CS-MTL reconstruct images that reserve relatively plausible structures (from a qualitative perspective) but with clearly over-smoothed reconstruction details and relatively low performance. By comparison, as FSL utilizes a learnable measurement to integrate reconstruction-analysis vision, it yields better performance (Figures 2A–2C) in subsampling artifact removal, better reliability in maintaining structural information, and more accurate segmentation results (Figures 2D, 2F, and 2H). Even when the sampling rate results in a large loss of raw data, the image structure can still be satisfactorily reserved by FSL. The corresponding sampling trajectories are placed in Figures 2E, 2G, and 2I.

Evaluation with 1D trajectories

Due to hardware requirements, Cartesian sampling trajectories are considered to validate the performance of the proposed FSL. CS-MRI₁, CS-MTL and CSL are unable to completely restore reliable structures and predict accurate segmentations (Figures 3D, 3F, and 3H) under different rates in the Cartesian-sampled images. Although CS-MRI₁ and CS-MTL try to recover more reliable structures, they have lower performance (Figures 3A–3C) and oversmoothed details constrained (Figures 3D, 3F, and 3H) by limited partial measurements (Figures 3E, 3G, and 3I). In contrast, the proposed FSL method performs well in medical segmentation and restores better structures and details.

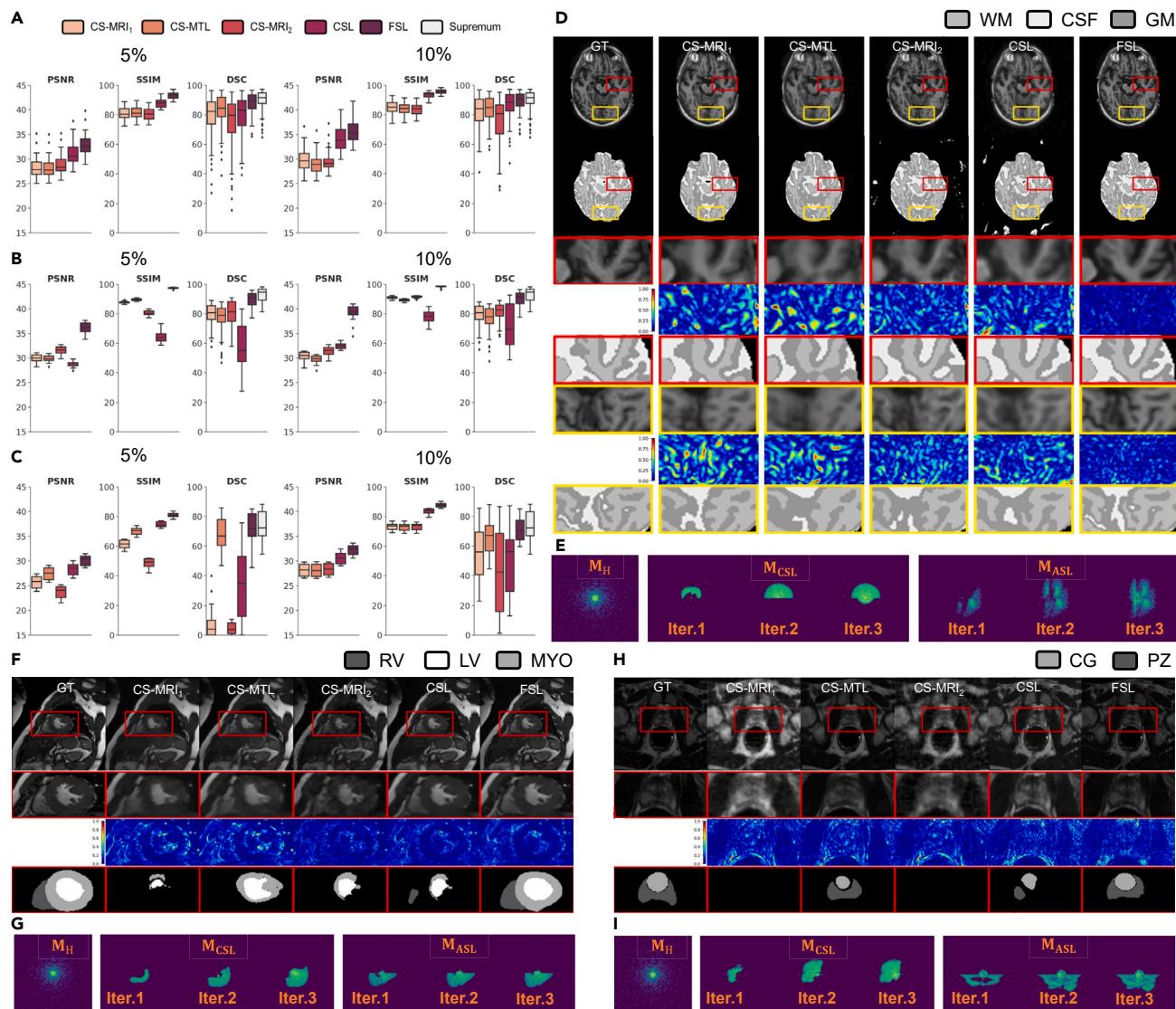


Figure 2. Comparison of FSL with different imaging paradigms based on ACDC, OASIS, and NCI-ISBI datasets for learning 2D sampling trajectories

(A–C) Quantitative results with 5% (left) and 10% (right) sampling rates using a variety of paradigms based on ACDC, OASIS and NCI-ISBI datasets respectively. (D) (F) (H) Representative visual results (rate = 5%) of reconstructed images (first row), segmentation maps (second row), zoom-in reconstructed images (third row) in the first ROI, zoom-in reconstructed error maps using NMSE (fourth row) in the first ROI, zoom-in segmentation maps (fifth row) in the first ROI. (D) Zoom-in reconstructed images (sixth row) in the second ROI, zoom-in reconstructed error maps using NMSE (seventh row) in the second ROI, and zoom-in segmentation maps (eighth row) in the second ROI. The ground truth (GT) is shown in the first column, and the images reconstructed using CS-MRI₁, CS-MTL, CS-MRI₂, CSL and FSL are provided in the second to sixth columns. (E) (G) (I) Sampling k-space amplitude images by corresponding trajectories of hand-tailed (left panel, rate = 5%), CSL learned (middle panel) over three iterations (from left to right) and FSL learned (right panel) over three iterations (from left to right). FSL removes more artifacts, preserves more anatomical details and predicts more accurate segmentations that are in agreement with the ground truth in ROI.

Generalization and robustness

In clinic practice, the importance of generalization and robustness of a DL-based imaging model has increased. We conduct detailed experiments and analysis of the differences between the three conventional paradigms and the proposed FSL model on two externally unseen datasets from fastMRI and M&Ms to study the generalization and on noise-corrupted OASIS testing datasets to study the robustness.

To investigate the collaboration of FSL, we perform a generalization experiment from externally same imaging modality scans, including M&Ms (cardiac MRI with segmentation reference) at 8% 1D sampling case and fastMRI (Brain MRI without segmentation reference) at 10% 1D sampling case. We find that the results of our method have the best or most competitive scores in comparison with those of other approaches in terms of various metrics (Figures 4A and 4C). In particular, FSL outperforms CSL by a large margin. Moreover, our methods still achieve

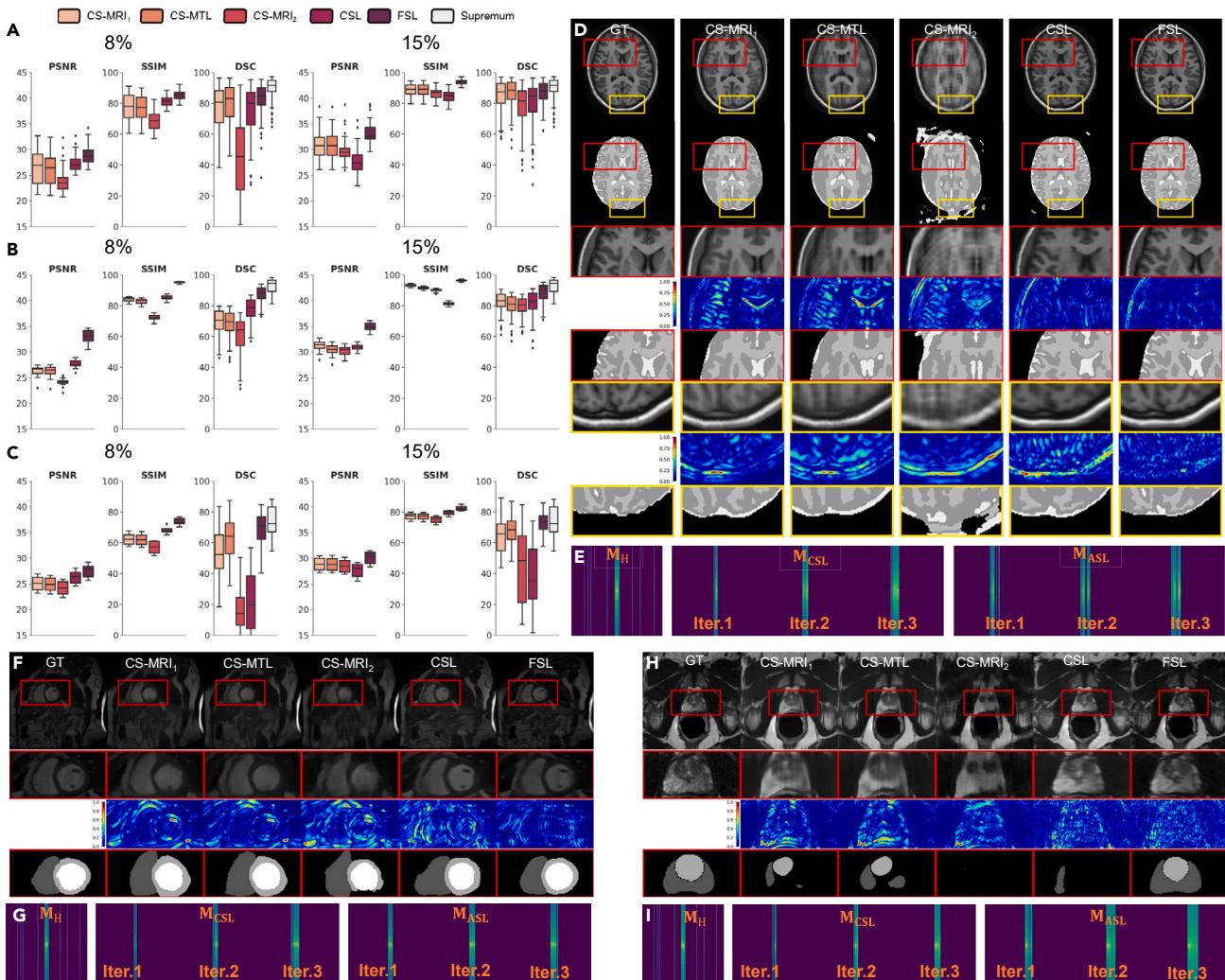


Figure 3. Comparison of FSL with different imaging paradigms based on ACDC, OASIS, and NCI-ISBI datasets for learning 1D sampling trajectories

(A–C) Quantitative results with 8% (left) and 15% (right) sampling rates using a variety of paradigms based on ACDC, OASIS and NCI-ISBI datasets respectively. (D) (F) (H) Representative visual results (rate = 5%) of reconstructed images (first row), segmentation maps (second row), zoom-in reconstructed images (third row) in the first ROI, zoom-in reconstructed error maps using NMSE (fourth row) in the first ROI, zoom-in segmentation maps (fifth row) in the first ROI. (D) Zoom-in reconstructed images (sixth row) in the second ROI, zoom-in reconstructed error maps (seventh row) in the second ROI, and zoom-in segmentation maps (eighth row) in the second ROI. The ground truth (GT) is shown in the first column, and the images reconstructed using CS-MRI₁, CS-MTL, CS-MRI₂, CSL, and FSL are provided in the second to sixth columns. (E) (G) (I) Sampling k-space amplitude images by corresponding trajectories of hand-tailed (left panel, rate = 5%), CSL learned (middle panel) over three iterations (from left to right) and FSL learned (right panel) over three iterations (from left to right). FSL outperforms the other methods in removing artifacts, in preserving anatomical details and in predicting accurate segmentation in ROI.

great performance on the fastMRI and M&Ms datasets in removing subsampled artifacts and recovering more real anatomical structures and textures (Figures 4B and 4D). The results suggest that our method has significant generalization in terms of artifact removal, detailed texture recovery, anatomical structure recovery and segmentation task.

As for robustness, experiments were performed by adding in silico Rician noise to the OASIS testing datasets at 8% 1D sampling case. Silico Rician noise-corrupted subsampled MR images are generated by adding Gaussian noise to the real and imaginary parts. Then we evaluate the performance of the three conventional paradigms and the proposed model in Figure 4E. As the noise level σ increases from 0.5 to 2.0, our method outperforms all other paradigms by a large margin. Moreover, we can observe that conventional imaging paradigms suffer from severe noise distortion (second to fourth columns in Figure 4F) and have unreliable segmentation maps, which is because hand-tailed trajectories or lack of semantics-guidance cannot sample efficient semantic information for supporting anatomical structure. Compared with these methods, FSL can acquire images with more reliable visual quality and details, closing to corresponding groundtruth. The results suggest that FSL interacts semantic representation, supporting trustworthy tissue structure and offsetting the noise corruption.

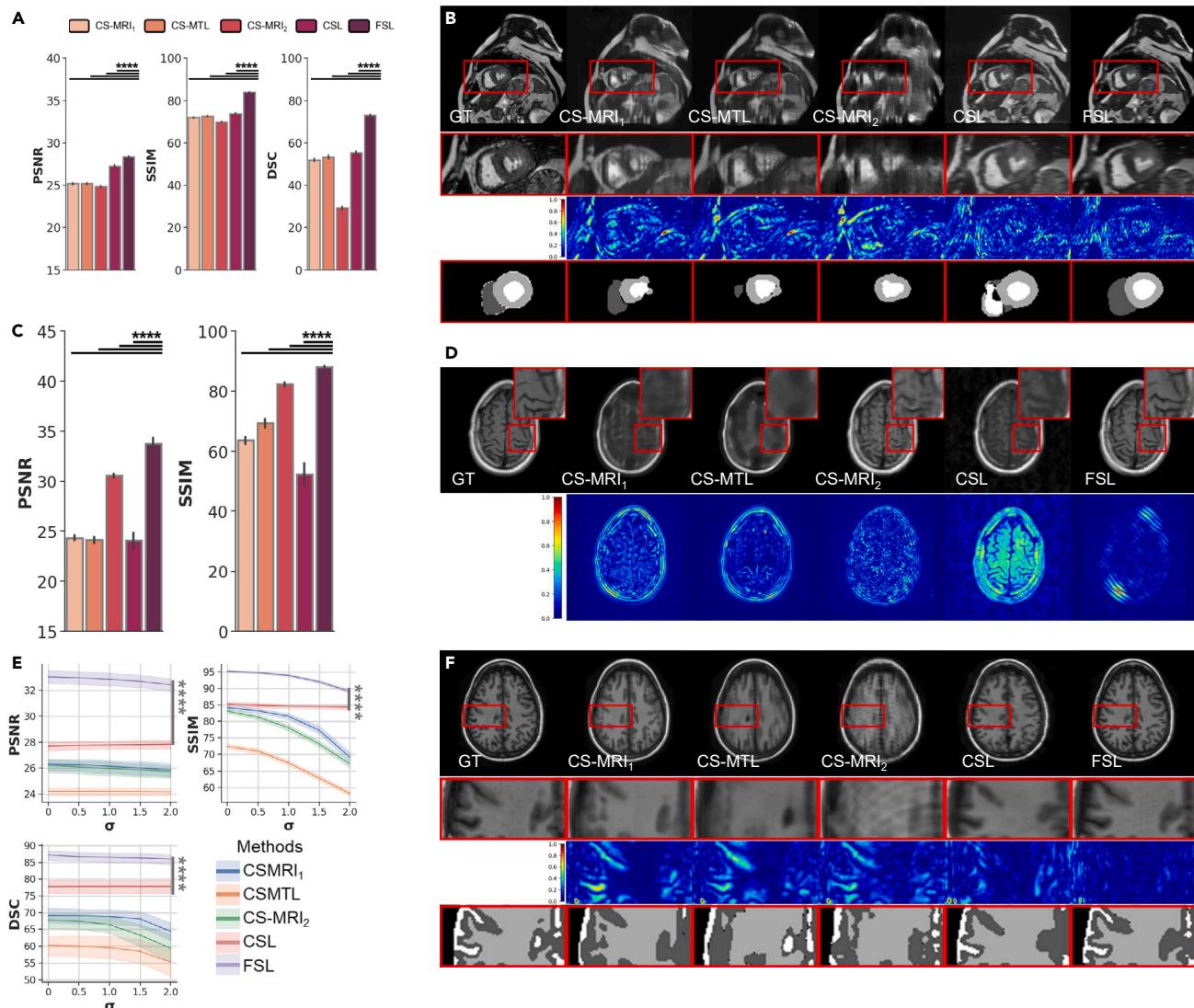


Figure 4. Generalization and robustness comparison of the reconstructed images for different externally unseen testing datasets at different sampling rates in different trajectories with a variety of paradigms

(A and B) Quantitative and qualitative generalization results (rate = 8%, 1D trajectories) on M&Ms dataset for evaluating generalization, where representative visual results in (B) are reconstructed images (first row), zoom-in reconstructed images (second row) of ROI, zoom-in reconstructed error maps using NRMSE (third row) of ROI, and zoom-in segmentation maps (fourth row) of ROI.

(C and D) Quantitative and qualitative results (rate = 10%, 2D trajectories) on fastMRI T1 test for evaluating generalization, where representative visual results in (D) are reconstructed images (first row), error maps using NRMSE (second row).

(E and F) Quantitative and qualitative generalization results on OASIS for evaluating robustness. (E) PSNR (left top), SSIM (right-top), DSC (left bottom) comparison of the robustness of different paradigms at different Rician noise levels $\sigma = 0, 0.5, 1.0, 1.5, 2.0$, where representative visual results in (F) of reconstructed images (first row), zoom-in reconstructed images (second row) of ROI, zoom-in reconstructed error maps using NRMSE (third row) of ROI, zoom-in segmentation maps (fourth row) of ROI. (B), (D), (F) The ground truths are shown in the first column, and the results using CS-MRI₁, CS-MTL, CS-MRI₂, CSL, and FSL are shown in the second to sixth columns. These samples validate that FSL is robust on sampling, reconstruction and segmentation for images with different noise and generalizable from different scanners.

Algorithm investigation

The effectiveness of coarse- and semantic-reconstruction for segmentation

In the interaction between reconstruction and segmentation, a fundamental difficulty in directly segmenting reconstructed images is the negative effects of subsampled artifacts and unreliable tissue of reconstructed images. Here, we use coarse reconstruction to remove subsampled artifacts and use semantic reconstruction to restore reliable anatomical tissue images. In Figures 5A and 5C, we implement FSL without SFRM for coarse-reconstruction (#3 in Figure 5B) and without SIM for semantic-reconstruction (#2 in Figure 5B). The performance

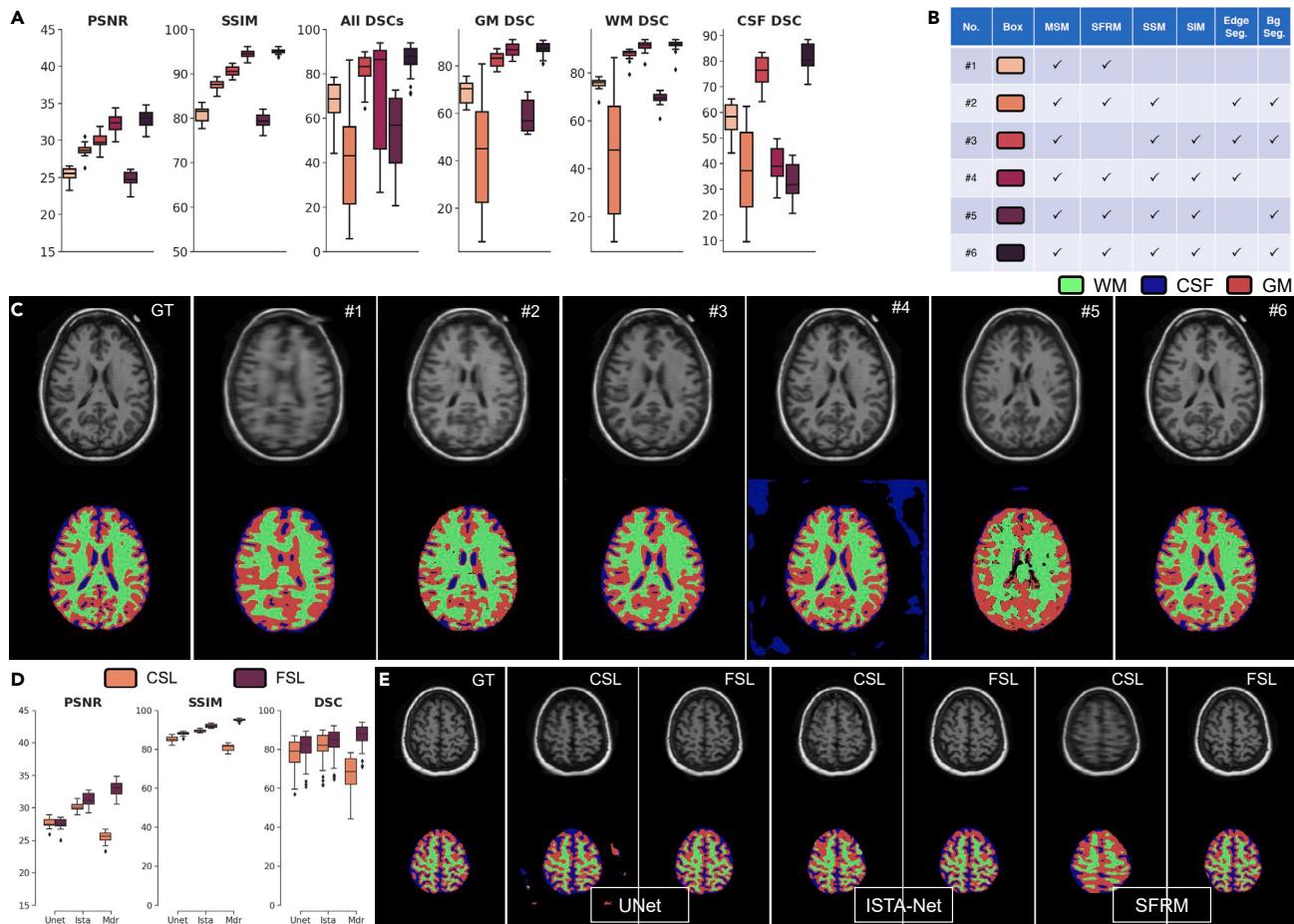


Figure 5. Ablation comparison of FSL with different modules and reconstruction backbones

(A), (C) Quantitative and Qualitative ablation results (rate = 8%, 1D trajectories) on the OASIS dataset for evaluating the effectiveness of each module. (A) PSNR (first column), SSIM (second column), DSC (third column), gray matter (GM) DSC (fourth column), white matter (WM) DSC (fifth column) and CSF DSC (sixth column) comparisons of FSL by different combinations (legends placed in (B)). (B) Different ablation combinations of FSL. (C) Representative visual results of reconstructed images (first row), and segmentation maps (second row). The ground truths are shown in the first column, and the results using ablation combination #1 to #6 (explained in (B)) are shown in the second to sixth columns.

(D) and (E), Quantitative and Qualitative results (rate = 8%, 1D trajectories) on OASIS test dataset for evaluating FSL paradigm where representative visual results in (E) of reconstructed images (first row) and segmentation maps (second row) of CSL and FSL on different reconstruction backbones on UNet (second panel), ISTA-Net (third panel) and SFRM (fourth panel). The results of the ablation experiments show that all three modules of FSL play a critical role, and that edge segmentation enhances both reconstruction and segmentation, while background segmentation enables the model to make predictions in the ROI region (A)-(C). FSL can improve the performance of different backbone networks, compared to CSL, and SFRM is well suited for the FSL paradigm (D), (E).

drops clearly demonstrate the benefits brought by coarse- and semantic-reconstruction. Moreover, we can observe the method encounters obvious performance drop without FSL. The proposed FSL learns structure-robust features from both pixel and semantic level, which plays an important role in segmentation results.

The effectiveness of high-level semantics for reconstruction

Then, we evaluate the effect of the proposed SSM and SIM at different semantic information guided by both segmentation classes of edge and background for reconstruction in Figure 5A. We can observe that the proposed edge segmentation (#4 in Figure 5B) brings a finer image reconstruction while the performance drops (#5 in Figure 5B) demonstrate the benefits brought by edge segmentation. Moreover, with background segmentation in FSL, results possess more anatomical features in ROI (#6 in Figure 5B), where the segmentation results are allowed for more accurate prediction in turn, due to the reliable reconstructed images. As shown in Figure 5C, we visualize the reconstruction results without (#1 in Figure 5B) or with SSM and SIM (#6 in Figure 5B). The method with SSM and SIM achieves visual quality with less subsampled artifacts and reliable bio-tissue structures. Therefore, it can be inferred that anatomical edge segmentation plays a vital role as a bridge between the reconstruction and segmentation processes of the multitask network.

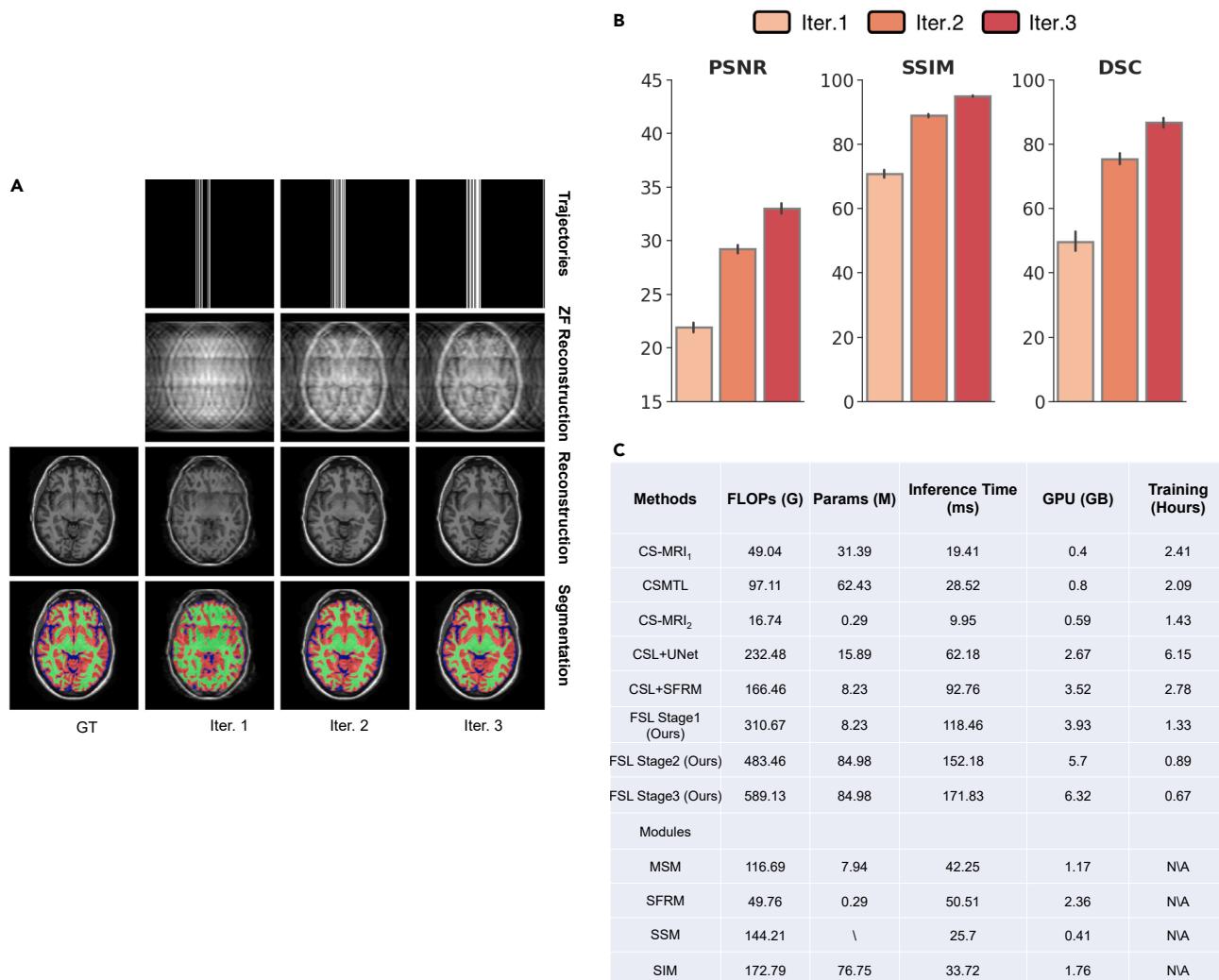


Figure 6. Experimental observation of iteration process and training strategies of FSL

(A and B) Qualitative and Quantitative results (rate = 8%, 1D trajectories) on the OASIS dataset for evaluating the effectiveness of each iteration. (A) Representative visual results of sampling trajectories (first row), zero-filling (ZF) reconstructed images (second row), reconstructed images (third row), and predicted segmentation (fourth row). The ground truths are shown in the first column. (B) PSNR (left), SSIM (middle), and DSC (right) comparisons of FSL over three progressively imaging iterations.

(C) Model performance of CS-MRI₁, CS-MRI₂, CS-MTL, CSL and FSL (different training stages and modules). GFLOPs calculation, inference time, training parameters, GPU cost and training time adopt with 240×240 resolution as inputs, batch size = 1. The SSIM and DSC are reported with inference on the OASIS testing dataset (rate = 8%, 1D trajectories).

The effectiveness of full-stack learning paradigm

To demonstrate the efficiency of our proposed FSL paradigm, we study the FSL extensions on two representative reconstruction backbones, including UNet,¹² ISTA-Net,²¹ Figures 5D and 5E summarize the numerical and visual results of on OASIS dataset at 10% 2D sampling rate of different reconstruction backbones equipped with different sampling paradigms, from which we observed that our imaging paradigms can cooperate effectively for these reconstruction algorithms and consistency outperforms the CSL paradigms. Note that, FSL with SFRM is still the top-performing FSL algorithm, while ISTA-Net with our FSL paradigm can reach a comparable performance using the same training strategy. These results collectively show the adaptive collaboration (yielded by our FSL) is of great benefit to MR imaging pipeline in terms of both better restoration and segmentation accuracy.

In addition, to verify progressive collaboration and semantic sampling strategy in our proposed method, how sampling, reconstruction and segmentation results change was analyzed in each sampling inference. The obtained results are shown in Figures 6A and 6B, from which we observe that the performance are improved along iterations simultaneously for both tasks. Better reconstruction would ease the difficulty

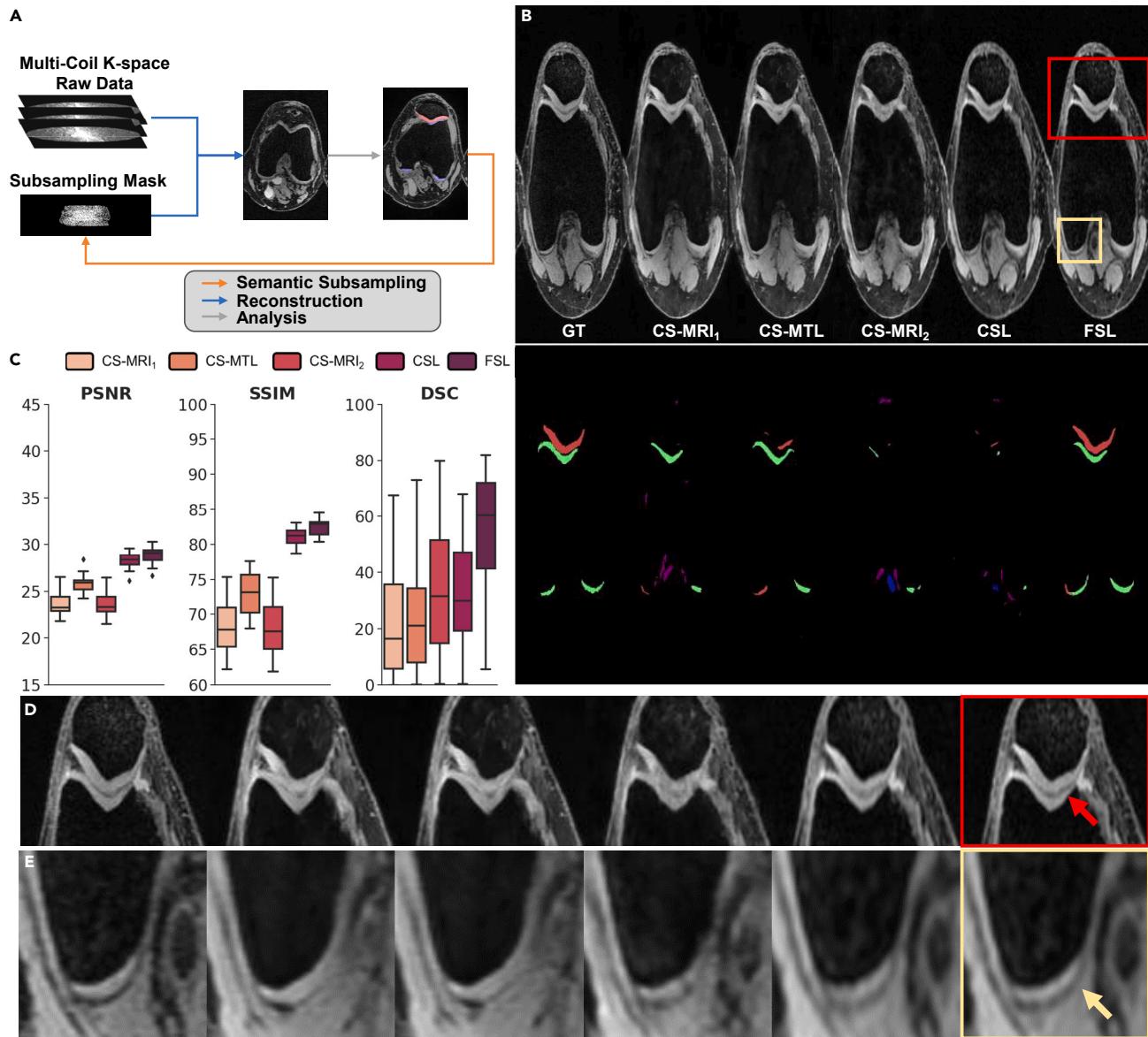


Figure 7. Comparison of FSL with different imaging paradigms on clinical SKMTEA dataset

(A) FSL applied to clinical data with multi-coil raw data.

(B and C) Qualitative and quantitative results (ratio = 10%, 2D trajectories) using a variety of paradigms, respectively.

(D and E) Enlarged ROIs, corresponding one-to-one with each column in (B).

of segmentation task; anatomical semantic guidance would collaborate reconstruction in turn; Knowledge-prior k-space trajectories can be obtained by inferencing MR images rich in semantics.

Evaluation on real-clinic multicoil raw data

To assess the performance of multitask on FSL paradigm for real scenarios, experiments were conducted on SKMTEA²² with real-clinical multi-coil raw knee data, reconstructed images, and segmentation labels. We evaluated our FSL and other methods on SKMTEA at 10% sampling ratio. The experimental results are depicted in Figure 7. It is observed that: (1) The proposed FSL achieves substantial improvements in PSNR, SSIM, and DSC when compared with CSL; (2) From a visual perspective, FSL demonstrates superior reconstruction and segmentation results in the ROI area. These findings suggest that integrating three tasks of sampling, reconstruction, and segmentation into our FSL effectively benefits each other's performance in real-world conditions. The proposed FSL shows a stable performance and strong generalization capability in real-world scenarios.

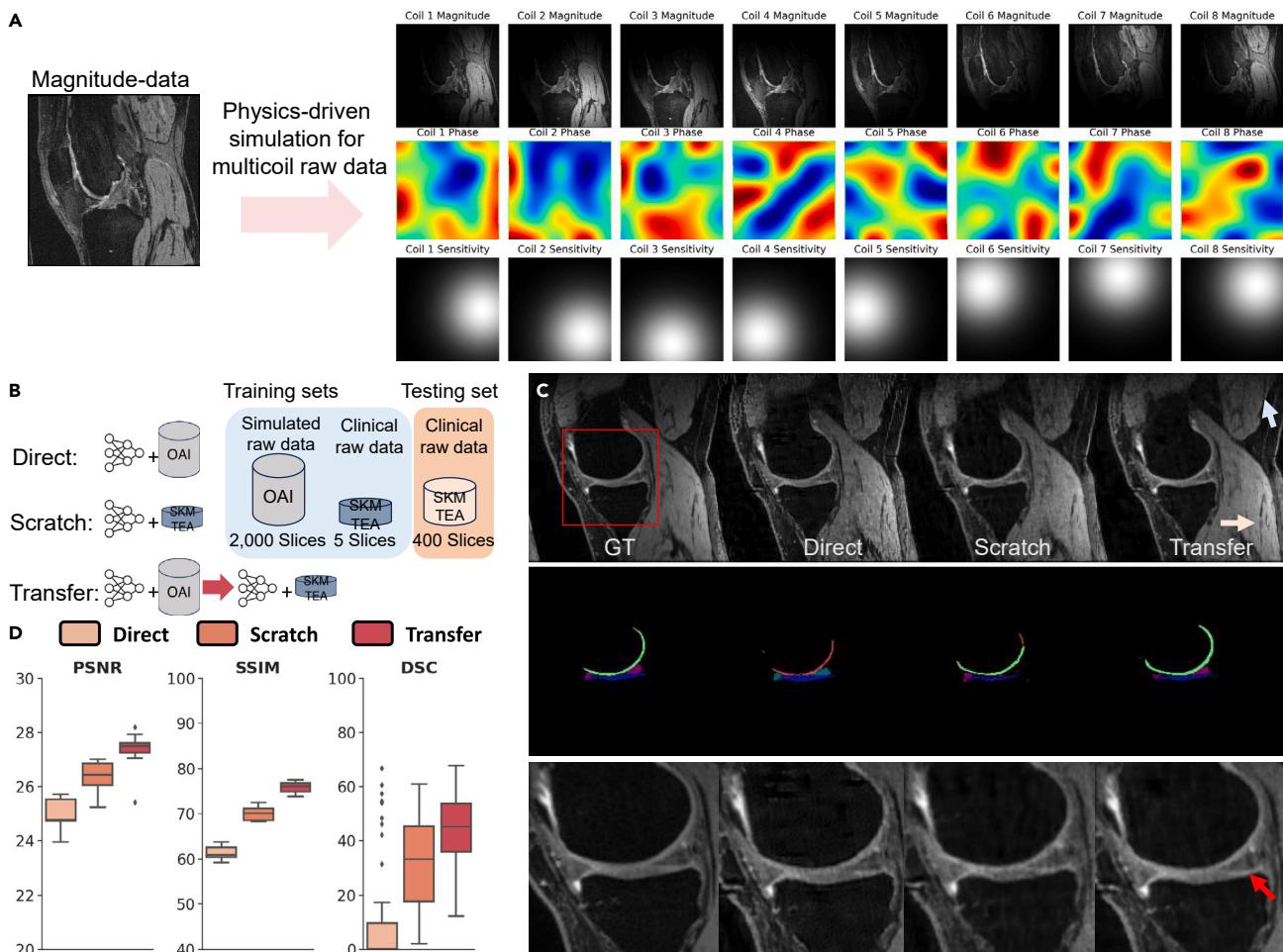


Figure 8. Investigation in transfer learning of FSL

(A) Physics-driven multicoil simulation on images with only magnitude values, accompanied by segmentation labels.
 (B) Three methods for evaluating transfer learning and few-shot learning. Direct: trained on OAI simulated raw data; Scratch: trained on clinical SKMTEA data with only 5 slices; Transfer: network weights trained on OAI simulated raw data used as initial values, followed by training on clinical SKMTEA data with only 5 slices. In the testing stage, all methods were tested on clinical SKMTEA data for sampling, reconstruction, and segmentation.
 (C and D) Qualitative and quantitative results (ratio = 10%, 2D trajectories) on clinical SKMTEA dataset.

Investigation of transfer learning and few-shot learning

Multitask data is very difficult to collect. To overcome the lack of multitask data for reconstruction and segmentation, we conducted experiments with 2D trajectory at 10% sampling ratio. We first pre-trained FSL model on 2,000 physics-driven simulation knee magnitude images with segmentation labels from OAI,²³ as depicted in Figure 8A; then, transfer learning (finetuning in our experiments) and few-shot learning (5 slices in our experiments) were applied to the pre-trained model on the real-world clinical SKMTEA.²² The finetuned model was noted as ‘Transfer’. As demonstrated in Figure 8B, two other compared methods were conducted: ‘Direct’, which was trained on OAI simulated data; and ‘Scratch’, which was trained on clinical SKMTEA data using only 5 slices. In the testing stage, all methods were evaluated on the clinical SKMTEA data. Experimental results in Figure 8 show that ‘Transfer’ outperforms ‘Direct’ and ‘Scratch’ in both reconstruction and segmentation performance qualitatively (Figure 8C) and quantitatively (Figure 8D). The results demonstrate that our FSL can leverage the knowledge acquired from one dataset to enhance performance on other datasets with limited labeled data through transfer learning and few-shot learning, thereby addressing the challenges arising from the lack of multitask datasets.

DISCUSSION

Our data suggest that reconstruction and analysis vision largely regard their measurement positions as an efficient collaboration that is capable of enhancing the representation of reconstruction-analysis vision in the whole imaging pipeline. Selecting a correct position with regard to arranging the measurements of reconstruction-analysis vision and collaborating the three tasks (measurement, reconstruction, medical segmentation) are associated with higher performance. Through the three comparative studies in our experiments, we find that FSL

achieves performance that is better than or comparable to that of other methods despite the lack of semantic information. We therefore confirm that the key paradigm of our proposal, i.e., FSL, contributes to achieving such high performance. In conventional CS-MRI paradigms, powerful models are designed to improve visual perception. Although such paradigms improve perception, our experiments and previous works suggest that some information is lost in the representation of reconstruction vision, while more is lost in analysis vision with a data-independent sampling trajectory. From the results, CSL approaches illustrate that even if sampling is associated with image texture details, without semantic association, k-space sampling trajectories cannot be fully optimized. In our experiments, CS-MTL shows that even with a segmentation task guide, reconstruction fails without a sampling mapping of texture and semantic information.

We presented an FSL approach, which is a full-stack framework for reconstruction and analysis tasks with learnable observation in a medical imaging workflow. We have demonstrated the full-stack collaborative training of a measuring inference method in the challenging imaging task of CS-MRI and the subsequent segmentation process. Furthermore, we have showcased quantitative and qualitative superiorities of FSL over the conventional MRI paradigms on externally unseen datasets and noise-corrupted datasets. Our work serves as the first step toward the implementation of next-generation imaging pipeline methods in medical imaging workflows. It can be applied to multitask research in many fields, such as video coding²⁴ and radio astronomical observations,²⁵ allowing for the collaboration of up- and downstream tasks and mastery over compressed sensing data with the theoretical and empirical guarantees of visually intermediate checks and a human-machine-friendly dataflow. FSL represents a targeted evolution of our previous work toward all-in-one focused deployment. Although we focus on a segmentation task for the presented case study, FSL is highly adaptable to a variety of medical imaging analysis workflows employing different network architectures, datasets and more. To date, no work has been published on the role of a learnable observational FSL model—an area where if feasible, its use as a collaborator for reconstruction-analysis vision would be an incredible breakthrough in our mastery over imaging pipelines.

In conclusion, we presented a novel FSL paradigm for medical image reconstruction and analysis tasks that are integrated by a learnable observational model on medical imaging data, and we showcase the paradigm in a clinically relevant real-life case study. Further research and development will enable a visual- and machine-friendly sequential design of our framework, the validation of our findings on sequential imaging, and the further widespread utilization of FSL techniques in healthcare and beyond. Importantly, the possibility of contributing to changing the way we collaborate with machines is a very exciting proposition.

Limitations of the study

In FSL, although we have achieved multi-task success on raw data, there is still a lack of multitask datasets that support both reconstruction and analysis vision in evaluating the performance of sampling learning methods. Specifically, common medical image datasets are mainly designed for image reconstruction tasks and lack high-level labels, while previous attempts toward sampling-based learning focused on pixel supervision and could not restore machine semantics. This paper presents the first attempt toward the use of compressed sampling to support reconstruction and segmentation vision simultaneously and provides a preliminary and new exploration direction that imposes higher demands on data. In this paper, we choose mainly those MRI datasets for two reasons. First, they have important research value in functional analysis and diagnostics. More importantly, these parts of humans have been well studied in the medical imaging community, and there are rich related segmentation and detection models that can provide labels for high-level tasks. Nevertheless, in future studies, due to the lack of suitable datasets, it is compromised to use the frequency data simulated from the spatial images as the raw data, which might not be an accurate simulation of the acquisition process. Hence, we call for the release of more expert-level annotations of medical diagnostics for raw data, and how to study unsupervised semantics for machine tasks is also our future research direction.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- METHOD DETAILS
 - Datasets
 - Evaluation metrics
 - Compressed sensing MRI
 - Trial exemplar architecture of FSL
 - Measurement sensing module
 - Spatial-frequency reconstruction module
 - Semantic segmentation module
 - Semantic interaction module
 - Compared methods

- Training strategy
- Implementation details

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2023.108608>.

ACKNOWLEDGMENTS

This work was supported in part by: The National Natural Science Foundation of China under Grant 62271335 and 61871277. In part by the Sichuan Science and Technology Program under Grant 2021JDJQ0024. In part by the Sichuan University "From 0 to 1" Innovative Research Program under Grant 2022SCUH0016.

AUTHOR CONTRIBUTIONS

Y.Z. and Z.W. conceived the study. Z.W. designed the experiments and the algorithmic pipeline under the supervision of Y.Z. and H.S. Z.W., B.L., and M.R. improved the reconstruction module. Z.W., H.Y., and Z.Z. improved the segmentation module. Z.W., Z.L., and W.X. improved the overall framework. Z.W. performed the experiments. Z.W., B.L., Z.Y., and J.L. analyzed the data, with conceptual advice from Y.Z., H.S., H.C., and J.Z. Z.W. wrote the manuscript under the supervision of Y.Z. and H.S. All authors discussed the results and commented on the manuscript. Y.Z. and H.S. supervised the research.

DECLARATION OF INTERESTS

The authors declare that they have no competing interests.

Received: March 15, 2023

Revised: October 17, 2023

Accepted: November 29, 2023

Published: December 2, 2023

REFERENCES

1. Lustig, M., Donoho, D.L., Santos, J.M., and Pauly, J.M. (2008). Compressed sensing MRI. *IEEE Signal Process. Mag.* 25, 72–82.
2. Wang, S., Su, Z., Ying, L., Peng, X., Zhu, S., Liang, F., Feng, D., and Liang, D. "Accelerating Magnetic Resonance Imaging via Deep Learning". 2016 IEEE 13th international symposium on biomedical imaging (ISBI). IEEE, pp. 514-517
3. Yang, Y., Sun, J., Li, H., and Xu, Z. (2020). ADMM-CSNet: A deep learning approach for image compressive sensing. *IEEE Trans. Pattern Anal. Mach. Intell.* 42, 521–538.
4. Yushchenko, M., Sarracanie, M., and Salameh, N. (2022). Fast acquisition of propagating waves in humans with low-field MRI: Toward accessible MR elastography. *Sci. Adv.* 8, eab04739.
5. Chen, Y., Schönlieb, C.-B., Lio, P., Leiner, T., Dragotti, P.L., Wang, G., Rueckert, D., Firmin, D., and Yang, G. (2022). AI-based reconstruction for fast MRI—A systematic review and meta-analysis. *Proc. IEEE* 110, 224–245.
6. Heaton, H., Fung, S.W., Lin, A.T., Osher, S., and Yin, W. (2022). Wasserstein-based projections with applications to inverse problems. *SIAM J. Math. Data Sci.* 4, 581–603.
7. Toi, P.T., Jang, H.J., Min, K., Kim, S.-P., Lee, S.-K., Lee, J., Kwag, J., and Park, J.-Y. (2022). *In vivo* direct imaging of neuronal activity at high temporospatial resolution. *Science* 378, 160–168.
8. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., and Maier-Hein, K.H. (2021). nnU-Net: a selfconfiguring method for deep learning-based biomedical image segmentation. *Nat. Methods* 18, 203–211.
9. Wang, S., Li, C., Wang, R., Liu, Z., Wang, M., Tan, H., Wu, Y., Liu, X., Sun, H., Yang, R., et al. (2021). Annotation-efficient deep learning for automatic medical image segmentation. *Nat. Commun.* 12, 5915.
10. Witkowski, J., Heacock, L., Reig, B., Kang, S.K., Lewin, A., Pysarenko, K., Patel, S., Samreen, N., Rudnicki, W., Łuczyńska, E., et al. (2022). Improving breast cancer diagnostics with deep learning for MRI. *Sci. Transl. Med.* 14, eab04802.
11. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444.
12. Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* 18 (Springer), pp. 234–241.
13. Pineda, L., Basu, S., Romero, A., Calandra, R., and Drozdzał, M. (2020). Active MR k-space sampling with reinforcement learning. In *Medical Image Computing and Computer Assisted Intervention-MICCAI 2020: 23rd International Conference, Lima, Peru, October 4-8, 2020, Proceedings, Part II* 23 (Springer), pp. 23–33.
14. Bahadir, C.D., Wang, A.Q., Dalca, A.V., and Sabuncu, M.R. (2020). Deep-learning-based optimization of the under-sampling pattern in MRI. *IEEE Trans. Comput. Imaging* 6, 1139–1152.
15. Yin, T., Wu, Z., Sun, H., Dalca, A.V., Yue, Y., and Bouman, K.L. (2021). End-to-end sequential sampling and reconstruction for MRI. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2105.06460>.
16. Van Gorp, H., Huijben, I., Veeling, B.S., Pezzotti, N., and Van Sloun, R.J. (2021). Active deep probabilistic subsampling. In *International Conference on Machine Learning (PMLR)*, pp. 10509–10518.
17. Sun, L., Fan, Z., Ding, X., Huang, Y., and Paisley, J. (2019). Joint CS-MRI reconstruction and segmentation with a unified deep network. In *Information Processing in Medical Imaging: 26th International Conference, IPMI 2019, Hong Kong, China, June 2-7, 2019, Proceedings* 26 (Springer), pp. 492–504.
18. Calíva, F., Leynes, A.P., Shah, R., Bharadwaj, U.U., Majumdar, S., Larson, P.E., and Pedoia, V. (2020). Breaking speed limits with simultaneous ultra-fast MRI reconstruction and tissue segmentation. In *Medical Imaging with Deep Learning (PMLR)*, pp. 94–110.
19. Zhang, Y., and Yang, Q. (2022). A survey on multi-task learning. *IEEE Trans. Knowl. Data Eng.* 34, 5586–5609.
20. Liu, D., Wen, B., Jiao, J., Liu, X., Wang, Z., and Huang, T.S. (2020). Connecting image denoising and high-level vision tasks via deep learning. *IEEE Trans. Image Process.* 29, 3695–3706.
21. Zhang, J., and Ghanem, B. (2018). ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1828–1837.
22. Desai, A.D., Schmidt, A.M., Rubin, E.B., Sandino, C.M., Black, M.S., Mazzoli, V., Stevens, K.J., Boutin, R., Re, C., Gold, G.E., et al. (2021). SKM-TEA: A Dataset for

- Accelerated MRI Reconstruction with Dense Image Labels for Quantitative Clinical Evaluation. In Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2).
23. Peterfy, C.G., Schneider, E., and Nevitt, M. (2008). The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. *Osteoarthritis Cartilage* 16, 1433–1441.
24. Li, X., Shi, J., and Chen, Z. (2021). Task-driven semantic coding via reinforcement learning. *IEEE Trans. Image Process.* 30, 6307–6320.
25. Pino, C., Sortino, R., Sciacca, E., Riggi, S., and Spampinato, C. (2021). Semantic segmentation of radioastronomical images. In Progress in Artificial Intelligence and Pattern Recognition: 7th International Workshop on Artificial Intelligence and Pattern Recognition, IWAIPR 2021, Havana, Cuba, October 5–7, 2021, Proceedings 7 (Springer), pp. 393–403.
26. Bernard, O., Lalande, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P.-A., Cetin, I., Lekadir, K., Camara, O., Gonzalez Ballester, M.A., et al. (2018). Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? *IEEE Trans. Med. Imag.* 37, 2514–2525.
27. Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., and Buckner, R.L. (2007). Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *J. Cognit. Neurosci.* 19, 1498–1507.
28. Bloch, N., and Madabhushi, A. (2015). NCI-ISBI 2013 Challenge: Automated Segmentation of Prostate Structures.
29. Campello, V.M., Gkontra, P., Izquierdo, C., Martin-Isla, C., Sojoudi, A., Full, P.M., Maier-Hein, K., Zhang, Y., He, Z., Ma, J., et al. (2021). Multi-centre, multi-vendor and multi-disease cardiac segmentation: the M&Ms challenge. *IEEE Trans. Med. Imag.* 40, 3543–3554.
30. Muckley, M.J., Riemschneider, B., Radmanesh, A., Kim, S., Jeong, G., Ko, J., Jun, Y., Shin, H., Hwang, D., Mostapha, M., et al. (2021). Results of the 2020 fastMRI challenge for machine learning MR image reconstruction. *IEEE Trans. Med. Imag.* 40, 2306–2317.
31. Guo, P., Wang, P., Zhou, J., Jiang, S., and Patel, V.M. (2021). Multi-institutional collaborations for improving deep learning-based magnetic resonance image reconstruction using federated learning. *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn.* 2021, 2423–2432.
32. Chen, Y., Schönlieb, C.-B., Liò, P., Leiner, T., Dragotti, P.L., Wang, G., Rueckert, D., Firmin, D., and Yang, G. (2022). AI-Based Reconstruction for Fast MRI—A Systematic Review and Meta-Analysis. *Proc. IEEE* 110, 224–245.
33. Huang, J., Ding, W., Lv, J., Yang, J., Dong, H., Del Ser, J., Xia, J., Ren, T., Wong, S.T., and Yang, G. (2022). Edge-enhanced dual discriminator generative adversarial network for fast MRI with parallel imaging using multi-view information. *Appl. Intell.* 52, 14693–14710.
34. Huang, J., Xing, X., Gao, Z., and Yang, G. (2022). Swin deformable attention u-net transformer (sdaut) for explainable fast mri. In International Conference on Medical Image Computing and Computer-Assisted Intervention (Springer), pp. 538–548.
35. Huang, J., Aviles-Rivero, A.I., Schönlieb, C.B., and Yang, G. (2023). ViGU: Vision GNN U-Net for fast MRI. Preprint at arXiv. <https://doi.org/10.48550/arXiv.2302.10273>.
36. Ran, M., Xia, W., Huang, Y., Lu, Z., Bao, P., Liu, Y., Sun, H., Zhou, J., and Zhang, Y. (2021). Md-recon-net: A parallel dual-domain convolutional neural network for compressed sensing mri. *IEEE Trans. Radiat. Plasma Med. Sci.* 5, 120–135.
37. Rad, M.S., Bozorgtabar, B., Marti, U.-V., Basler, M., Ekenel, H.K., and Thiran, J.-P. (2019). Srobb: Targeted perceptual loss for single image super-resolution. In Proceedings of the IEEE/CVF international conference on computer vision, pp. 2710–2719.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Automated Cardiac Diagnosis Challenge (ACDC) dataset	University Hospital of Dijon, France	https://www.creatis.insa-lyon.fr/Challenge/acdc/
the Open Access Series of Imaging Studies (OASIS) dataset	Washington University, USA	http://www.oasis-brains.org/#data
NCI-ISBI 2013 Challenge - Automated Segmentation of Prostate Structures dataset	The National Cancer Institute, USA	https://wiki.cancerimagingarchive.net/display/public/nci-isbi+2013+challenge+-+automated+segmentation+of+prostate+structures
the Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation (M&Ms) dataset	Universitat de Barcelona, Spain	https://www.ub.edu/mnms/
fastMRI dataset	New York University, USA	https://fastmri.org/
The SKM-TEA dataset	Stanford AIMI Center, USA	https://github.com/StanfordMIMI/skm-tea
The Osteoarthritis Initiative (OAI) dataset	National Institutes of Health, USA	https://nda.nih.gov/oai/
Software and algorithms		
Our proposed FSL	This paper	https://github.com/wangzhiwen-scu/FSL
Other		
Quadro RTX 8000	NVIDIA Corporation, USA	https://www.nvidia.com

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Yi Zhang (yzhang@scu.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All data are available in the main text. We analyzed multiple published MRI datasets throughout the evaluations. These data are available as follows: (1) Automated Cardiac Diagnosis Challenge (ACDC) dataset (<https://www.creatis.insa-lyon.fr/Challenge/acdc/>); (2) the Open Access Series of Imaging Studies (OASIS) dataset (<http://www.oasis-brains.org/#data>); (3) NCI-ISBI 2013 Challenge - Automated Segmentation of Prostate Structures dataset (<https://wiki.cancerimagingarchive.net/display/public/nci-isbi+2013+challenge+-+automated+segmentation+of+prostate+structures>); (4) the Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation (M&Ms) dataset (<https://www.ub.edu/mnms/>); (5) fastMRI dataset (<https://fastmri.org/>); (6) The SKM-TEA dataset (<https://github.com/StanfordMIMI/skm-tea>); and (7) The Osteoarthritis Initiative (OAI) (<https://nda.nih.gov/oai/>).
- The FSL model was developed using standard libraries in the open-source platform PyTorch 1.7.1. The source codes of FSL, several representative pre-trained models as well as some example images for testing are publicly accessible via: <https://github.com/wangzhiwen-scu/FSL>. The code base is made available for non-commercial and academic purposes. The datasets and the code are publicly available as of the date of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](mailto:yzhang@scu.edu.cn) upon request (yzhang@scu.edu.cn).

METHOD DETAILS

Datasets

We start the exploration of our FSL framework on cardiac, brain and prostate MR images, as they are usually highly structured, share unified tissue components, provide strong priors for reconstruction-analysis tasks, and lead to a very important research area of medical image analysis. We train FSL by using five training datasets, ACDC, OASIS, NCI-ISBI, OAI and SKMTEA, and evaluate it on themselves' testing datasets, respectively. In addition, we added Rician noise to test the robustness of the model and tested the generalization of the model on fastMRI

and M&Ms datasets. There are seven datasets in the multi-task experiments: ACDC (cardiac MRI, for training and testing), OASIS (brain MRI for training and testing), NCI-ISBI (prostate MRI for training and testing), M&Ms (cardiac MRI for testing), fastMRI (brain MRI for testing), OAI (knee MRI for multicoil data training), and SKMTEA (knee MRI for multicoil data testing). Since there is no large design for both MR image reconstruction and segmentation, subsampled k-space data are obtained by executing preset subsampled trajectories on the segmentation benchmark. Here are details for each dataset.

- ACDC.²⁶ We collect the cardiac T1-weighted MRI scans from Automated Cardiac Diagnosis Challenge (ACDC), including 50 subjects for training, and 50 subjects for evaluation. The cardiac tissues of each patient are labeled with three types, including Left-Ventricle (LV), Right Ventricle (RV) and Myocardium (MYO).
- OASIS.²⁷ The brain dataset from the Open Access Series of Imaging Studies (OASIS) is used to evaluate the proposed method. Each patient aged 18 to 96 is provided with T1-weighted MRI modalities. Each scan had approximately 180 slices with expert annotations for cortical gray matter (GM), white matter (WM), and cerebrospinal fluid in the extracerebral space (CSF). We collect 20 patients for training, and 20 patients for evaluation.
- NCI-ISBI.²⁸ The abdominal T2-weighted MRI dataset from the NCI-ISBI 2013 Challenge - Automated Segmentation of Prostate Structures, comprises of 20 scans for training and 10 scans for evaluation, where each scan had approximately 30 slices with expert annotations, including central gland (CG) and peripheral zone (PZ).
- M&Ms.²⁹ We test our FSL model for generalization on an external cardiac MRI dataset from the Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation (M&Ms). The training dataset is acquired on two scanners at multi-centers and consists of 150 patients. The dataset of each patient is provided with T1 modalities where three organs were annotated, including LV, RV and MYO. In our experiment, all patients are used for testing. M&Ms and ACDC are slightly different w.r.t their scanning parameter settings.
- FastMRI.³⁰ It is a brain benchmark dataset for MRI reconstruction. Since no ground truth is available for dataset segmentation, we test our FSL model only for reconstruction generalization on the external fastMRI Brain subset. The brain subset data are acquired on 3T and 1.5T scanners and consist of 6970 samples, and the test dataset consists of 860 samples. The subset of each patient is provided with four MRI modalities, including FLAIR, T1 weighted, T1 weighted with contrast agent (T1 POST) and T2 weighted. In our experiments, we choose randomly 80 emulated³¹ single-coil (ESC) brain T1-weighted MRI scans for evaluation. FastMRI and OASIS are greatly different w.r.t their scanning parameter settings.
- SKMTEA.²² The SKMTEA multitask dataset contains multicoil k-space raw data, reconstructed images, and expert-derived segmentation labels from 155 knee qMRI scans. Each scan has two echoes, and we chose the first one for our experiments, where $k_x \times k_y \times k_z = 512 \times 512 \times 160$ for each k-space volume with voxel spacing [0.3125, 0.3125, 0.8] mm. Constrained by the storage cost of k-space data, we randomly chose 12 scans for training and 8 scans for testing from SKMTEA. For Sec. *Evaluation on real-clinic multicoil raw data*, note that only the middle 300 slices $k_x \times k_z$ were used in the transverse plane. For Sec. *Investigation of transfer learning and few-show learning*, the middle 80 slices $k_x \times k_y$ were used in the sagittal plane. The selection of the middle slices was to include most of the labels. Expert-derived segmentation labels in SKMTEA are provided: patellar cartilage, femoral cartilage, lateral tibial cartilage, medial tibial cartilage, lateral meniscus, and medial meniscus.
- OAI.²³ The Osteoarthritis Initiative (OAI) dataset contains 240 double-echo DESS knee MRI scans. Each image data volume size is $384 \times 384 \times 160$ and its voxel spacing is $0.3646 \times 0.3646 \times 0.7$ mm. Only the reconstructed sagittal images and expert-derived segmentation labels are provided. The segmentation labels include femoral cartilage, medial tibial cartilage, lateral tibial cartilage, patellar cartilage, lateral meniscus, and medial meniscus. 20 scans were selected randomly and simulated into multicoil raw data for training.

Evaluation metrics

We evaluate our algorithm against the above state-of-the-art MRI methods in terms of image quality and segmentation accuracy on open-source datasets and use the peak-signal-to-noise ratio (PSNR), structural similarity index (SSIM) as reconstruction evaluation metrics, and Dice similarity coefficient (DSC) as segmentation evaluation metrics.

Compressed sensing MRI

MR imaging can be formulated as the following inverse problem:

$$\mathbf{y} = \mathbf{Ax} + \mathbf{n} \quad (\text{Equation 1})$$

where $\mathbf{x} \in \mathbb{C}^{M \times M}$ stands for a 2D MR image, $\mathbf{y} \in \mathbb{C}^{M \times N}$, represents fully-sampled measurements in k-space (i.e., Fourier space), \mathbf{n} stands for measured error and system noise, and $\mathbf{A} = \mathbf{F}\mathbf{M}$ is the measurement process, where \mathbf{F} is Fourier transform matrix and $\mathbf{M} \in \{1\}^{M \times N}$ is a binary sampling trajectory. In compressed sensing (CS) MRI, only a partial observation $\mathbf{M} \in \{0, 1\}^{M \times N}$ is implemented by a certain sampling rate α , which equals the matrix size of MRI divided by observed measurement.

Recent advances in compressed sensing (CS) magnetic resonance imaging (MRI) focus on deep learning based MRI reconstruction from undersampled data, effectively reducing artifacts and improving image quality.³² A pioneer work is to perform CS-MRI reconstruction with convolutional neural network (CNN).² Similarly, in,³³ the authors adopt Generative Adversarial Networks (GANs) to further refine the reconstructed images and improve the perceptual quality and diagnostic viability. Different from conventional network architectures, Transformer models³⁴ were introduced to capture long-range dependencies and hierarchical structures in image data and showed improved

reconstruction of intricate anatomical structures. In,³⁵ Graph Neural Networks (GNNs) serve as a powerful tool to model implicit relations in image data, delivering perceptible improvements in reconstructing images from undersampled k -space data. However, these methods suffer from huge computational and GPU memory consumption. In contrast, the dual-domain model, MD-Recon-Net,³⁶ pays more attention to the fidelity of k -space data and utilizes frequency-domain and image-domain information in a lightweight yet high-performance manner. Since FSL requires multiple iterations of different tasks, leading to a multiplicative increase in computational load and GPU memory usage, the light-weight dual-domain model is suitable for our FSL paradigm.

Trial exemplar architecture of FSL

We first give an overview of our FSL approach (Figure 1D). Figure 1E illustrates the proposed trial exemplar architecture of FSL method. Our framework combines segmentation and reconstruction with trajectory learning, which involves anatomical structural sampling. The main modules of the proposed framework, including the measurement sensing module (**MSM**), spatial-frequency reconstruction module (**SFRM**), semantic segmentation (**SSM**) and semantic interaction module (**SIM**), are first sequentially described in detail. Then, other issues regarding the proposed architecture, especially those related the training strategy and loss function, are presented.

Measurement sensing module

For MR image reconstruction, the sampling trajectory is commonly empirically hand-tailored to avoid subsampling and aliasing, e.g., radial, Cartesian, or spatial Boolean matrices. However, a fixed trajectory ignores the fact that different anatomical structure and reconstruction model may be suitable for different subsampling trajectories, which will miss the semantic measurements of trapping into the local optimum. In response to this challenge, we proposed MSM to progressively sampling the efficiently semantic measurements (Figure S1). Given a fully-sampled measurement \mathbf{y} and sampling trajectory \mathbf{M}_i (a center preset sampling \mathbf{M}_0 as the initial trajectory), in the i th iteration, the process can be written as:

$$\mathbf{y}_{i-1} = \mathbf{M}_{i-1} \odot \mathbf{y} \quad (\text{Equation 2})$$

where \odot is the Hadamard product. The MSM outputs a sampling trajectory \mathbf{M}_i as a partial observation, which is progressively predicted by MSM with the $i-1$ th iteration subsampling reconstructed image $\bar{\mathbf{y}}_{i-1}$.

The goal of the MSM is to predict the sampling pattern for specific issue datasets in the k -space:

$$\mathbf{M}_i = \text{MSM}(\bar{\mathbf{y}}_{i-1}, \mathbf{M}_{i-1}). \quad (\text{Equation 3})$$

To predict a probabilistic observation matrix in the k -space, we adopt a similar architecture¹⁵ for our MSM. Inspired by LOUPE¹⁴ and SeqMRI,¹⁵ the architecture and details of the MSM are shown in Figure S1. Here, we show two different predicted trajectories for specific prior tissue knowledge in Figures 2E, 2G, 3E, 3G, and 3I.

We conduct two different sampling trajectory cases experiments for sampling, reconstruction and segmentation. The specific details are outlined below.

Task1: 2D trajectory (Gaussian-like) learning

Constrained by equipment, a Gaussian-like trajectory is still calculated in this paper to demonstrate the high performance of our approach. The top- n largest values in \mathbf{T}_{MC} (Figure S1) are replaced by Boolean values to produce the final sampling pattern \mathbf{M}_i , and n is chosen according to the predetermined sampling rate α , where $n = \alpha \cdot M \cdot N$. Accordingly, the Booleanizing operation can be written as:

$$(\mathbf{M}_{kl}) = \begin{cases} 1, & \text{if } (\mathbf{T}_{MC})_{kl} \text{ is in top } -n \\ 0 & \text{otherwise.} \end{cases} \quad (\text{Equation 4})$$

Task2: 1D trajectory (Cartesian) learning

Cartesian trajectory learning is a common sampling method for structural MRI in conventional spin echo (CSE or SE) and fast spin echo (FSE) pulse sequences, but this approach suffers from slow data acquisition. The fractional number of excitations (NEX) accelerates MRI by reducing the number of encoded phases.

In this case, a subsampling trajectory \mathbf{M}_i is searched in a 2D sampling space $\mathbf{T}_{MC}^{M \times N}$ (Figure S1). First, the weights \mathbf{T}_{MC} of each phase encoding are obtained by summing the frequency encoding, $\sum_0^N \mathbf{T}_{MC}^{M \times N}$. Then, the binary process is the same as that the without constrained trajectory. The process above can be written defined as:

$$(\mathbf{M}_k) = \begin{cases} 1, & \text{if } (\mathbf{T}_{MC})_k \text{ is in top } -n \\ 0 & \text{otherwise.} \end{cases} \quad (\text{Equation 5})$$

Finally, we extend the binary result to a 2D matrix (along the frequency encoding axis). Since we do not have the labels for sampling pattern learning, we propose to merge the MSM into the FSDM, SSM and SIM described below. When the cascaded network converges, the output \mathbf{M}_i of the MSM will be optimal.

Spatial-frequency reconstruction module

Recently, extensive network models have been proposed for MRI reconstruction, and in this work, we carefully modify a Spatial-Frequency Reconstruction Module (SFRM). Letting $SFRM(\cdot)$ denotes the SFRM, \mathbf{F}^{-1} is the inverse Fourier transform matrix, \mathbf{y}_i is the sampling measurement, and $\mathbf{x}_i = \mathbf{F}^{-1}\mathbf{y}_i$ is the zero-filling image, and the coarse reconstructed image $\bar{\mathbf{x}}$ can be obtained in the i th iteration as:

$$\bar{\mathbf{x}}_c^i = SFRM(\mathbf{x}_i, \mathbf{y}_i) \quad (\text{Equation 6})$$

The SFRM branches of the FSL are trained as a coarse reconstructing integral with the objective function defined as:

$$\mathcal{L}_{cr} = \|\bar{\mathbf{x}}_c^i - \mathbf{x}_{gt}\|_2^2 \quad (\text{Equation 7})$$

where $\|\cdot\|_2^2$ is a L2 loss used to measure the similarity between the reconstructed image and the corresponding label. The architecture of the SFRM adopts the interactive representation infusing between k-space and image domains as the backbone,³⁶ and this kind of backbone has demonstrated competitive performance in artifact reduction for ours FSL.

Semantic segmentation module

In recent years, prior structural or edge knowledge has been exploited by gradient guidance in deep learning-based methods, which is a regularization function derived from structural priors. Model-based methods sometimes draw prior structures from limited data in a time-consuming way. On the other hand, current deep learning-based methods cannot exploit artificially issued structural information in the k-space. Furthermore, both of the above approaches cannot provide a feedback structure prior for entering the k-space. In this work, we utilize hand-crafted prior segmentation with structural knowledge, which is realized by a Semantic Segmentation Module (SSM) of a supervised neural network, to exploit the optimal structural priors in the k-space to improve the reconstruction and segmentation robustness of the model and prevent overfitting:

$$\bar{\mathbf{s}}_c^i = SSM(\bar{\mathbf{x}}_c^i) \quad (\text{Equation 8})$$

where $\bar{\mathbf{s}}_c^i$ is the coarse segmentation in the i th iteration. Recently, many networks have been proposed for automatic tissue segmentation^{8,12}. Since the UNet-like architecture has demonstrated excellent medical image segmentation performance, in this section, we also choose the same network structure shown in Figure S2 as our SSM for simplicity.

Semantic interaction module

As mentioned above, the interaction between the imaging pipeline i.e., sampling, reconstruction, and segmentation is inadequate in MSM, SFRM and SSM, which will result in limited performance in the whole pipeline, since it only communicates information in a gradient manner. Therefore, we take a further step towards exploring all stages interaction learning in the pipeline. Specifically, we propose a novel Semantic Interaction Module (SIM) (see the right panel in Figure 1E), which adopts spatially-adaptive normalization layers to effectively integrate the semantic label into the interaction process from low to high image scales. As a result, the model can learn to guide the SFRM process to adjust semantic information and further benefit both the reconstruction and segmentation tasks. In the i th iteration, we take the coarse reconstruction $\bar{\mathbf{x}}_c^i$ and coarse segmentation $\bar{\mathbf{s}}_c^i$ as input to SIM to get a fine reconstruction, $\bar{\mathbf{x}}_f^i$, which will be input again to SSM to get a fine segmentation, $\bar{\mathbf{s}}_f^i$. This fine reconstruction $\bar{\mathbf{x}}_f^i$ will be inputted to MSM for predicting the next sampling trajectory in the next iteration. We use SIM to obtain a fine reconstructed image $\bar{\mathbf{x}}_f^i$:

$$\bar{\mathbf{x}}_f^i = SIM(\bar{\mathbf{x}}_c^i, \bar{\mathbf{s}}_c^i) \quad (\text{Equation 9})$$

$$\bar{\mathbf{s}}_f^i = SSM(\bar{\mathbf{x}}_f^i) \quad (\text{Equation 10})$$

The FSL samples P times to progressively refine the reconstruction and segmentation: $\bar{\mathbf{x}}^i, \bar{\mathbf{s}}^i = (\bar{\mathbf{x}}_f^i, \bar{\mathbf{s}}_f^i)$. $(\bar{\mathbf{x}}_f^P, \bar{\mathbf{s}}_f^P)$ are the final refined reconstructed image and semantic segmentation maps, respectively. We use L2 loss functions to make SIM convergence:

$$\mathcal{L}_{fr} = \|\bar{\mathbf{x}}_f - \mathbf{x}_{gt}\|_2^2 \quad (\text{Equation 11})$$

Cross-entropy loss is utilized for the SSM:

$$\mathcal{L}_{seg} = - \sum_k \mathbf{s}_{gt}^k \ln \bar{\mathbf{s}}_f^k \quad (\text{Equation 12})$$

for K tissue class labels, where \mathbf{s}_{gt}^k is the target label and $\bar{\mathbf{s}}_f^k$ is the softmax segmentation probability for the k^{th} class. In addition, under unstable observations, the model needs to take the upstream and downstream tasks into account, but the characteristics of the learning process for the upstream and downstream tasks are very different, and this requires a loss function to build a bridge between the two. In a recent work, anatomical edge³⁷ provides a bridge between low-level and high-level vision. While the L1 loss in Equation 11 enforces holistic supervision from the fully sampled ground-truth MR image, the boundaries between the key tissues (e.g., cortical gray matter, white matter and ventricles) of brains may be easily obscured as they are distorted under a certain sampling ratio. The use of the L1 loss alone is likely to fail to restore true

tissue boundaries from artifacts. Thus, inspired by SROBB,³⁷ we propose organ edge detection and foreground/background extraction as extra segmentation labels to explore strong semantic mapping in the k-space, generate semantically reconstructed results and improve the accuracy of downstream tasks. The process of label creation is shown in [Figure S4](#). The organ edge and background enforce the FSL to explore more semantic details with fewer subsampled artifacts and generate latent representations for downstream tasks with respect to MR images.

Compared methods

We use a trail exemplar FSL and conduct an ablation-like study as our basic experiment to demonstrate that a) semantic exploration in the k-space is crucial for high-quality image and anatomy reservation and b) the guidance of representative semantic features is the key to providing a sampling- and learning-based reconstruction task. With respect to a), we train a Unet¹² as baseline CS-MRI₁. In addition, for further guidance by semantic representation, we follow Liu et al.²⁰ and pretrain the segmentation networks as CS-MTL that are pretrained on the fully sampled image, which essentially is FSL without a sampler, i.e., joint reconstruction and segmentation with fixed trajectories – with the guidance of semantic information but without semantic exploration in the k-space. With regard to b), we train a joint learning network for trajectory optimizing and reconstruction, i.e., seqMRI¹⁵ as CSL, without the guidance of semantic features. Moreover, we also introduce a state-of-the-art model MD-Recon-Net³⁶ as CS-MRI₂ to demonstrate that an excellent model might fail when the sampling trajectory, distribution of the dataset, and reconstructor are mismatched.

CS-MRI₁: We build a variant of our FSL as a baseline that learns reconstruction with fixed trajectories. We train this model using the same training strategy as that of our FSL method.

CS-MTL: Inspired by Liu et al.,²⁰ We build a variant of our FSL as a baseline that learns reconstruction and segmentation simultaneously with fixed trajectories and a pretrained segmentation network. We train this model using the same training settings as that of our FSL method.

CS-MRI₂: The MRI dual-domain reconstruction network MD-Recon-Net is a recently proposed dual-domain reconstruction with fixed trajectories. It contains two parallel and interactive branches that simultaneously operate on the k-space and spatial-domain data.

CSL: seqMRI is a recently proposed sampling pattern learning model driven by reconstruction.

For CS-MRI₁, CS-MRI₂, and CSL, we train them with paired MR images for reconstruction test. In testing, the inputs are the reconstructed images in the pretrained segmentation network same with CS-MTL and FSL. In our next experiments, we empirically study the behaviors of FSL. We pay special attention to what may contribute to the model's mutual benefits among the three subnetworks.

Training strategy

We train the proposed framework on the three described tasks. Considering their diversity under unstable measurements, it is difficult to jointly train the whole network synchronously. The proposed model adopts the following progressive training strategy:

- Stage1: We first fix the pretrained SSM and train the MSM and SFRM using L2 loss and CE loss for 3,000 iterations.
- Stage2: We then fix MSM, SFRM and SSM and train the SIM using the L2 loss and CE loss for 3,000 iterations.
- Stage3: Finally, we jointly optimize both MSM and SIM by minimizing the L2 loss and CE loss for 3,000 iterations.

We demonstrate that adopting such a training strategy guarantees that the learned sampling pattern can acquire as much useful information as possible for the subsequent reconstruction and segmentation tasks. More specifically, the proposed FSL can be easily adapted to different clinical tasks, and we can substitute the SSM with any other task network. Our approach not only facilitates the training effort while utilizing the ReconNet to fit clinical tasks and keeping the SSM performing accurately for subsampled MR images but also enables the sampler to learn more clinically useful features from the k-space data.

Implementation details

Then, the hybrid loss function for the proposed joint learning network is formulated as:

$$\mathcal{L} = \mathcal{L}_{cr} + \lambda_1 \mathcal{L}_{sseg} + \lambda_2 \mathcal{L}_{fr} \quad (\text{Equation 13})$$

where λ_1, λ_2 are the weighting parameters used to balance the tradeoff between the two components. The number of the iteration i for FSL is set to be 3. All implementations are based on PyTorch. All models are trained using one Quadro RTX 8000 GPU, and the batch size is set to 12. Uniform random initialization is used for the sampler and Xavier initialization for SFRM and SSM. The ADAM optimizer is adopted with an initial learning rate of $5e^{-4}$. λ_1 and λ_2 are empirically set to $1e^{-1}$ and $1e^{10}$, respectively. We use a batch size of 12 and apply random rotation-based data augmentation within $\pm 10^\circ$. The whole training process takes approximately 0.5 days on the above GPU card. To preprocess the data, we used PyTorch real-time data augmentation. To save computation, all images were truncated to 240×240 in k-space from their different original size. We simulate the sampling process by using trajectories.