LEARNING FROM K-SPACE: ACQUISITION-INFORMED PRETRAINING FOR LABEL-EFFICIENT SEGMENTATION FROM UNDERSAMPLED 3D MRI

Anonymous authors
Paper under double-blind review

ABSTRACT

Accurate three-dimensional (3D) brain tumor segmentation from MRI faces two practical challenges: costly voxel-wise annotations and accelerated acquisition through undersampling. Existing MRI segmentation paradigms either operate on reconstructed images without direct access to acquisition-domain information or incorporate k-space as an additional input without explicitly modeling its task-relevant structure, which can degrade segmentation performance under limited supervision. To address this, we propose an acquisition-informed self-supervised learning framework for effective label-efficient 3D brain tumor segmentation. Specifically, our network learns directly from undersampled k-space together with the corresponding zero-filled reconstructed image, combining acquisition-domain frequency information with essential spatial context without requiring a high-quality MRI reconstruction. To utilize k-space more effectively, our empirical analysis reveals structured pathology-associated spectral deviations that exhibit meaningful spatial correspondence with tumor regions. Consequently, we formulate these spectral residuals as a soft spectral prior during pretraining to guide representation learning. Finally, we introduce an auxiliary lesion-presence branch during fine-tuning that softly modulates dense predictions to effectively suppress unsupported false positives. Extensive experiments on multimodal brain tumor MRI demonstrate consistent improvements under limited supervision, highlighting the value of pathology-aware k-space learning for label-efficient segmentation from accelerated MRI.

1 INTRODUCTION

Three-dimensional (3D) brain tumor segmentation from multimodal MRI typically integrates T1-weighted, contrast-enhanced T1 (T1ce), T2-weighted, and T2-FLAIR (FLAIR) sequences to delineate whole tumor (WT), tumor core (TC), and enhancing tumor (ET) Menze et al. (2014); de Verdier et al. (2024); Bonato et al. (2025). Reliable segmentation of these tumor subregions is essential for clinical assessment and treatment planning, yet two practical challenges remain: annotation scarcity and acquisition acceleration. First, dense voxel-wise annotation requires substantial expert effort, motivating label-efficient segmentation with limited labeled data Chaitanya et al. (2020); Zhou et al. (2021). Second, MRI signals are acquired in the frequency domain (k-space) and are routinely undersampled to accelerate scanning Lustig et al. (2007); Wang et al. (2016). Conventional "reconstruct-then-segment" pipelines optimize an intermediate reconstruction stage for image fidelity rather than segmentation-relevant pathological structures Tolpadi et al. (2023). This decoupled approach may introduce residual artifacts and over-smoothing Wang et al. (2024), while also leaving the raw acquisition-domain signals underexploited once the image is reconstructed. Together, making downstream segmentation highly challenging under limited supervision.

As illustrated in Figure 1,, we formulate label-efficient segmentation directly from undersampled acquisitions using k-space together with its zero-filled reconstruction. Computed via a deterministic inverse Fourier transform (IFFT) of the zero-filled k-space, the zero-filled image provides essential spatial context. Because it does not rely on an auxiliary learned reconstruction module, this approach inherently avoids model-induced hallucinations and reconstruction biases. However, learning effective representations from these undersampled dual-domain inputs remains challenging under label

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

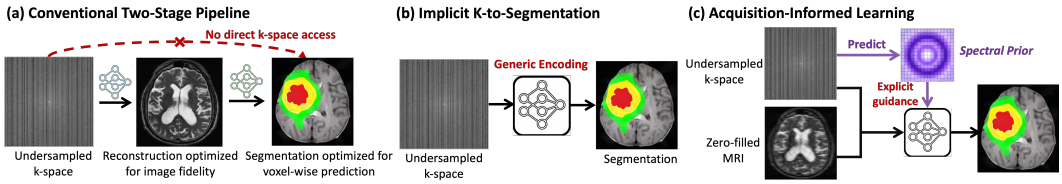

Figure 1: **Comparison of segmentation paradigms from undersampled MRI acquisition.** Conventional pipelines segment reconstructed images without direct k-space access; direct k-space methods use acquisition-domain signals as generic input features; our framework instead performs acquisition-informed learning by jointly leveraging undersampled k-space and zero-filled MRI, while utilizing the spectral prior from k-space to provide explicit guidance for segmentation.

scarcity, as limited supervision may be insufficient to distinguish pathology-relevant frequency patterns from undersampling and anatomical variations. We therefore leverage unlabeled fully sampled acquisitions during offline pretraining to provide additional frequency-domain supervision.

Self-supervised learning (SSL) Chaitanya et al. (2020); Zhou et al. (2021); He et al. (2022); Tang et al. (2022); Munk et al. (2024); Wald et al. (2025) provides a natural way to exploit such unlabeled data during pretraining and reduce reliance on scarce task-specific annotations. A key question, then, is how to utilize this k-space information effectively. Existing studies have incorporated k-space into deep networks (Huang et al., 2019; Xiao et al., 2021; Li et al., 2024; Yen et al., 2024), but largely rely on the network to discover task-relevant frequency patterns implicitly (Figure 1b). Since k-space encodes image structure globally through complex frequencies, implicitly discovering pathology-relevant patterns can be data-inefficient under limited supervision. We therefore ask whether pathological changes produce structured (k)-space signatures that can serve as an explicit prior for label-efficient segmentation.

To answer this, we conduct an empirical analysis of whether brain tumors are associated with structured changes in MRI k-space (Section 3). We characterize deviations from the smooth spectral energy decay of fully sampled k-space as *spectral residuals*, and examine whether prominent residual frequency bands correspond to lesion regions. Their image-space projections achieve a Localization Hit Rate (LHR) of 0.775, but only moderate spatial precision (0.340), indicating that these spectral deviations provide pathology-sensitive yet spatially coarse cues. We therefore formulate them as a *soft spectral prior* for representation learning rather than deterministic spatial pseudo-labels.

Building on this observation, we propose an acquisition-informed framework for label-efficient 3D brain tumor segmentation from undersampled k-space and its zero-filled reconstruction. During pretraining, fully sampled acquisitions are retrospectively undersampled to form the network inputs, while the original measurements provide reconstruction and spectral supervision. We jointly optimize image reconstruction and spectral-prior prediction to encourage the encoder to capture both anatomical structure and pathology-associated frequency patterns. Under limited supervision, models trained on undersampled inputs are prone to mistaking aliasing artifacts for pathological tissues, causing severe false positives. Since global frequency representations capture whole-lesion existence cues robustly, we introduce an auxiliary presence branch during fine-tuning. The resulting presence probabilities softly modulate dense predictions, suppressing unsupported false positives under limited supervision.

Our main contributions are threefold:

- We identify a key limitation of existing MRI segmentation paradigms: acquisition-domain information is either inaccessible after reconstruction or only implicitly modeled. To utilize k-space more effectively, we further reveal structured pathology-associated spectral deviations in k-space and formulate them as a lesion-relevant prior for segmentation.

- We propose an acquisition-informed framework that explicitly integrates this spectral prior into label-efficient tumor segmentation. Spectral-prior prediction complements reconstruction during self-supervised pretraining, while a presence-aware mechanism leverages predicted spectral evidence to calibrate segmentation and suppress unsupported false positives.

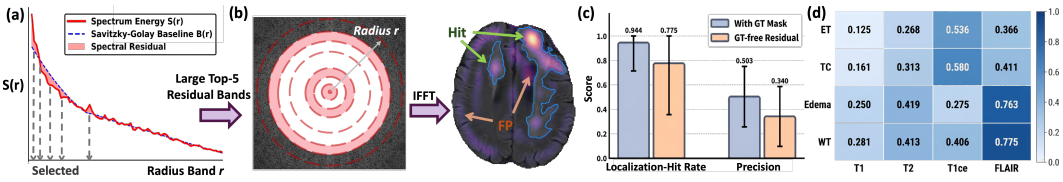

Figure 2: **Empirical evidence of lesion-relevant spectral structure in k-space.** **(a)** GT-free spectral characterization. **(b)** The image-domain projections of the top-5 selected bands can localize the lesion. **(c)** Comparison between GT-guided oracle and GT-free residual-based band selection on WT. **(d)** Modality-region sensitivity matrix of GT-free LHR across modalities and tumor subregions.

- Extensive experiments on undersampled multi-modal MRI demonstrate consistent improvements over competitive end-to-end and pretraining-based baselines, together with substantial reductions in false-positive predictions.

2 Related Work

Self-Supervised and Label-Efficient Medical Image Segmentation. Self-supervised pretraining is widely used to reduce the dependence of medical image segmentation on dense annotations. Existing methods learn transferable representations through reconstruction-based pretext tasks (Zhou et al., 2021; Haghighi et al., 2021), contrastive learning (Chaitanya et al., 2020), and masked modeling or self-distillation for volumetric images (Tang et al., 2022; Xie et al., 2022; Wu et al., 2024). Recent large-scale studies further demonstrate the effectiveness of masked autoencoding for label-efficient 3D segmentation, including AMAES (Munk et al., 2024) and related large-scale pretraining efforts (Wald et al., 2025). These approaches primarily derive supervision from reconstructed images and thus make limited use of acquisition-domain information. Our work instead incorporates pathology-associated spectral structure from k-space as an additional self-supervised learning signal.

Learning from k-Space for Downstream MRI Analysis. Recent studies have explored learning directly from raw or undersampled k-space for downstream MRI analysis, including segmentation (Schlemper et al., 2018; Huang et al., 2019; Zhang et al., 2024; 2025; Rotkopf et al., 2026), pathology detection (Xiao et al., 2021), and task-adaptive sampling (Yen et al., 2024). KMAE (Li et al., 2024) further introduces masked autoencoding in the frequency domain and demonstrates transferable k-space representations across classification, regression, and segmentation tasks. These studies establish the utility of acquisition-domain information, but primarily exploit k-space as model input or a representation-learning space, leaving pathology-relevant spectral structure to be learned implicitly. In contrast, we explicitly characterize lesion-associated spectral deviations and convert them into a spectral prior for acquisition-informed pretraining and label-efficient 3D segmentation.

3 Empirical Motivation: Spectral Signatures of Pathology in k-Space

A central premise of our method is that pathological changes leave identifiable signatures in k-space. To investigate this premise, we conduct a preliminary empirical study with two objectives: first, to verify that these k-space signatures produce spatial responses to pathological regions when mapped back to the image domain; and second, to show that such signatures can be detected by a specific annotation-free procedure.

3.1 Spectral Characterization Formulation

Spectral Energy Distribution. Given the raw 2D k-space data k of an MRI slice x, we characterize its frequency content by aggregating spectral amplitudes along the radial axis. Specifically, we

define the spectral energy distribution (SED) as

$$S(r) = \frac{1}{|\Omega_r|} \sum_{(k_x,k_y)\in\Omega_r} \log\big(|k(k_x,k_y)| + \epsilon\big), \tag{1}$$

where Ω_r denotes the set of k-space coordinates in the r-th radial frequency band, and ϵ is a small constant for numerical stability. The resulting $S(r)$ summarizes how spectral energy is distributed across frequencies.

Savitzky–Golay Baseline and Spectral Residual. To identify lesion-related spectral signatures as deviations from the smooth frequency decay of normal anatomy, we apply a Savitzky–Golay filter (Savitzky & Golay, 1964) to $S(r)$. Specifically, within a local window of radius h, we fit a degree-p polynomial by

$$\mathbf{a}^*(r) = \arg\min_{\mathbf{a}} \sum_{j=-h}^{h} \left[S(r+j) - \sum_{q=0}^{p} a_q j^q \right]^2, \qquad B(r) = a_0^*(r), \quad R(r) = S(r) - B(r), \tag{2}$$

where $B(r)$ is the smooth baseline given by the zero-order coefficient, and $R(r)$ is the corresponding spectral residual. By construction, $R(r)$ is computed directly from k-space without lesion annotations. In the following, we examine whether these residuals reveal pathology-related spectral signatures and whether such signatures can be identified in an annotation-free manner.

3.2 Empirical Evidence for Lesion-Relevant Spectral Signatures

We map selected frequency bands back to the image domain via inverse Fourier transform (IFFT) and analyze their spatial responses. We consider two settings: a GT-guided oracle that identifies lesion-relevant frequency bands using annotations, and a GT-free setting that selects the top-5 bands with the largest $|R(r)|$. Figure 2 summarizes the analysis pipeline and results.

To quantify localization quality, let $\mathcal{C}_i$ be the set of response components for case i and G_i the corresponding lesion mask. A component is considered as correct if its center lies inside G_i or if it overlaps more than 20% of the lesion area. Denoting the set of correct components by $\mathcal{C}_i^* = \{c \in \mathcal{C}_i \mid \mathrm{Center}(c) \in G_i \vee |c \cap G_i|/|G_i| > 0.2\}$, we measure whether at least one component correctly localizes the lesion using the localization hit rate (LHR), together with the spatial precision of the responses:

$$\mathrm{LHR} = \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}(|\mathcal{C}_i^*| > 0), \qquad \mathrm{Precision} = \frac{1}{N}\sum_{i=1}^{N} \frac{\sum_{c\in\mathcal{C}_i^*} |c|}{\sum_{c\in\mathcal{C}_i} |c|}. \tag{3}$$

Observation 1: Pathological changes leave spatially informative spectral signatures in k-space. Under the GT-guided oracle, selected frequency bands produce responses that strongly overlap the lesion, reaching an LHR of 0.944 and a precision of 0.503 on WT (Figure 2(c)). This provides empirical evidence that pathology-related information is encoded in structured regions of k-space and remains spatially informative after projection to the image domain.

Observation 2: Large spectral residuals detect lesion-relevant signatures without annotation and exhibit modality sensitivity. We next compare the GT-free selection based on the top-5 $|R(r)|$ values under the same evaluation protocol. Although no lesion annotations are used, the projected responses retain substantial lesion localization on WT, achieving an LHR of 0.775 without lesion annotations (Figure 2(c)). More importantly, this localization is modality-dependent (Figure 2(d)): T1ce yields the strongest responses for ET and TC, with LHRs of 0.536 and 0.580, whereas FLAIR is most sensitive to Edema and WT, reaching 0.763 and 0.775, respectively. The modest precision of the GT-free responses (0.340 on WT) further indicates that the lesion-relevant evidence is informative but not spatially specific. Therefore, we use the residual-derived signal as a modality-specific *soft spectral prior*, rather than as a direct pseudo-label.

4 Methodology

4.1 Overview and Problem Formulation

Our goal is label-efficient 3D brain tumor segmentation from accelerated MRI, where only undersampled k-space and its zero-filled reconstruction are available at inference. We therefore formulate

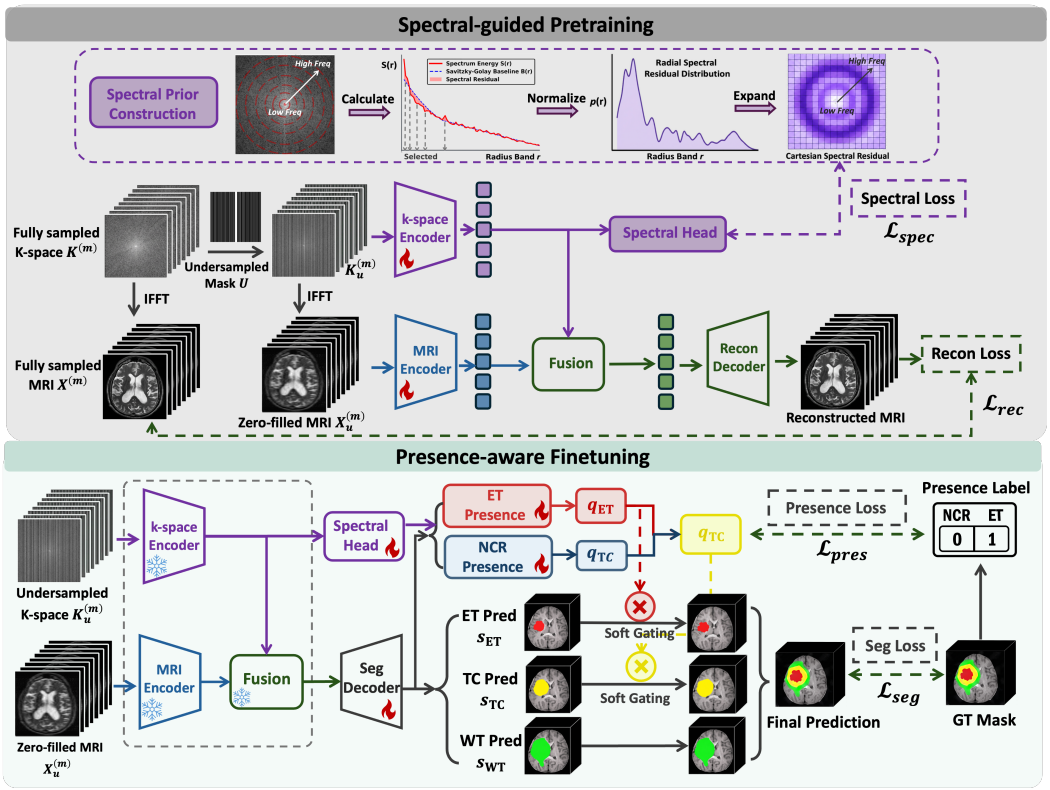

Figure 3: **Overview of the proposed K2S framework.** During spectral-guided pretraining, undersampled k-space and zero-filled MRI are jointly encoded, while a spectral prior constructed from fully sampled k-space supervises the spectral prediction together with an image reconstruction objective. During presence-aware fine-tuning, the pretrained encoders are adapted for segmentation using voxel-wise supervision and slice-level presence prediction, where the predicted spectral information provides guidance for the final tumor segmentation.

the task in a dual-domain setting that directly preserves acquisition-domain information without requiring a separately reconstructed high-quality MRI as input. The model is pretrained on unlabeled data and subsequently adapted using limited voxel-wise annotations; throughout both stages, the model receives only undersampled observations, while fully sampled acquisitions are used solely to construct self-supervised pretraining targets.

Let $\mathcal{D}_u$ and $\mathcal{D}_l$ denote the unlabeled pretraining set and label-scarce fine-tuning set, respectively. Each subject contains four co-registered MRI modalities, $\mathcal{M} = \{\mathrm{T1}, \mathrm{T1ce}, \mathrm{T2}, \mathrm{FLAIR}\}$. For modality $m \in \mathcal{M}$, let $K^{(m)}$ denote the fully sampled k-space measurement and $X^{(m)} = \mathcal{F}^{-1}(K^{(m)})$ the corresponding MRI volume. Given a random undersampling mask U, we define

$$\tilde{K}^{(m)} = U \odot K^{(m)}, \qquad \tilde{X}^{(m)} = \mathcal{F}^{-1}(\tilde{K}^{(m)}),$$

and use $\{(\tilde{K}^{(m)}, \tilde{X}^{(m)})\}_{m\in\mathcal{M}}$ as the dual-domain input. During pretraining, reconstruction and spectral-guided supervision jointly learn anatomical and pathology-associated representations. During fine-tuning, these representations are adapted to predict the nested tumor regions $\mathbf{Y} = \{Y_{\mathrm{WT}}, Y_{\mathrm{TC}}, Y_{\mathrm{ET}}\}$, with $Y_{\mathrm{WT}} \supseteq Y_{\mathrm{TC}} \supseteq Y_{\mathrm{ET}}$, while presence-aware modulation suppresses unsupported positive predictions under limited supervision.

4.2 Spectral-Guided Pretraining

As illustrated in Fig. 3, the MRI and k-space encoders first extract domain-specific representations from the paired undersampled inputs:

$$\mathbf{F}_x = E_x\left(\{\tilde{X}^{(m)}\}_{m\in\mathcal{M}}\right), \qquad \mathbf{F}_k = E_k\left(\{\tilde{K}^{(m)}\}_{m\in\mathcal{M}}\right).$$

The two domain-specific representations are progressively fused within the MRI encoder: projected k-space features are first injected through a learnable residual connection, followed by a deeper cross-attention module $\Phi_{\text{fuse}}(\cdot)$ in which MRI features serve as queries and k-space features as keys and values Vaswani et al. (2017). The resulting shared representation $\mathbf{Z} = \Phi_{\text{fuse}}(\mathbf{F}_x, \mathbf{F}_k)$ is decoded as $\hat{\mathbf{X}} = D_{\text{rec}}(\mathbf{Z})$ and supervised by the fully sampled MRI $\mathbf{X}$, providing anatomical reconstruction supervision. In parallel, a spectral head $\hat{\mathbf{p}} = H_{\text{spec}}(\mathbf{F}_k)$ predicts the lesion-sensitive spectral prior $\mathbf{p}$ derived from fully sampled k-space as described in Sec. 3.1. This objective encourages the learned k-space representation to preserve pathology-associated spectral information complementary to image-domain anatomy.

Cartesian Spectral Prior. The radial spectral residual $R_z^{(m)}(r)$ in Sec. 3.1 already captures lesion-related deviations from the smooth anatomical spectral trend. Here, instead of using this one-dimensional radial residual directly, we further lift it to the Cartesian frequency grid under a radial-symmetry assumption that is structurally aligned with the 2D k-space feature layout. This expansion preserves the lesion-sensitive evidence carried by $R_z^{(m)}(r)$ while providing a spatially aligned supervision map for Cartesian k-space Baddour (2011); Bernal et al. (2020).

Specifically, we partition the 2D frequency plane into N non-overlapping Cartesian regions $\{\mathcal{A}_n\}_{n=1}^{N}$. Each Cartesian coordinate (u, v) is assigned the residual magnitude at its corresponding radial frequency $r_{u,v}$, and the values are aggregated and normalized over regions:

$$p_{z,n}^{(m)} = \frac{\sum_{(u,v)\in\mathcal{A}_n} |R_z^{(m)}(r_{u,v})| + \epsilon}{\sum_{j=1}^{N} \left(\sum_{(u,v)\in\mathcal{A}_j} |R_z^{(m)}(r_{u,v})| + \epsilon \right)}. \tag{4}$$

This yields a normalized Cartesian spectral prior $\mathbf{p}_z^{(m)} = \{p_{z,n}^{(m)}\}_{n=1}^{N}$, where $r_{u,v}$ denotes the radial frequency of coordinate (u, v) and ϵ is a small constant. The prior thus converts lesion-sensitive radial residuals into a topology-aligned supervisory target for the k-space branch.

Pretraining Objectives. The overall pretraining objective combines image reconstruction and spectral prediction:

$$\mathcal{L}_{\text{pre}} = \mathcal{L}_{\text{rec}} + \lambda_{\text{spec}} \mathcal{L}_{\text{spec}}. \tag{5}$$

The reconstruction decoder $\mathcal{D}_{\text{rec}}$ predicts the fully sampled MRI from the fused representation:

$$\mathcal{L}_{\text{rec}} = \frac{1}{|\mathcal{M}|} \sum_{m\in\mathcal{M}} \left\| \mathcal{D}_{\text{rec}}(\mathbf{Z})^{(m)} - X^{(m)} \right\|_1. \tag{6}$$

In parallel, a lightweight spectral head $\mathcal{H}_{\text{spec}}$ predicts the slice-wise Cartesian spectral distribution $\hat{\mathbf{p}}_z^{(m)}$ directly from the k-space features. As shown in 3.2, T1ce and FLAIR exhibit relatively stronger lesion-related spectral sensitivity than native T1 and T2. We therefore restrict spectral supervision to $\mathcal{M}_{\text{spec}} = \{\text{T1ce}, \text{FLAIR}\}$ using the Kullback–Leibler divergence:

$$\mathcal{L}_{\text{spec}} = \frac{1}{|\mathcal{M}_{\text{spec}}| D} \sum_{m\in\mathcal{M}_{\text{spec}}} \sum_{z=1}^{D} D_{\text{KL}} \left(\mathbf{p}_z^{(m)} \,\|\, \hat{\mathbf{p}}_z^{(m)} \right). \tag{7}$$

T1 and T2 remain part of the dual-domain representation and are supervised through the reconstruction pathway and cross-domain feature fusion.

4.3 Presence-Aware Fine-Tuning and Inference

During fine-tuning, relatively small tumor subregions such as TC and ET are particularly susceptible to spurious false-positive predictions under limited supervision Li et al. (2023). We therefore augment voxel-wise segmentation with slice-level presence estimation to provide complementary evidence of whether each tumor component is supported.

Presence-Aware Fine-Tuning. As in pretraining, the model takes both undersampled k-space and its zero-filled reconstruction as input. Their dual-domain representations are fused and passed to a 3D segmentation decoder, replacing the reconstruction decoder, to predict voxel-wise probability maps $s_c(v) \in [0, 1]$ for $c \in \{\text{WT}, \text{TC}, \text{ET}\}$. In parallel, the retained spectral pathway extracts

frequency-derived features from the undersampled k-space, which are combined with spatial features from the segmentation pathway by a lightweight presence head to estimate slice-level probabilities $q_{z,c} \in [0,1]$ for $c \in \{\mathrm{NCR}, \mathrm{ET}\}$.

The segmentation branch is optimized with

$$\mathcal{L}_{\mathrm{seg}} = \mathcal{L}_{\mathrm{BCE}} + \mathcal{L}_{\mathrm{Dice}}, \tag{8}$$

where binary cross-entropy (BCE) penalizes voxel-wise classification errors, while Dice directly optimizes region overlap and alleviates foreground–background imbalance Isensee et al. (2021). Slice-level presence labels are derived from the segmentation masks:

$$y_{z,\mathrm{ET}} = \mathbb{I}\left(\sum_{v \in z} Y_{\mathrm{ET}}(v) > 0\right), y_{z,\mathrm{NCR}} = \mathbb{I}\left(\sum_{v \in z} \left[Y_{\mathrm{TC}}(v) - Y_{\mathrm{ET}}(v)\right] > 0\right). \tag{9}$$

To address the imbalance between tumor-present and tumor-absent slices, we combine grouped BCE with a margin-based ranking objective over positive and hard-negative slices Huang et al. (2024):

$$\mathcal{L}_{\mathrm{pres}} = \sum_{c \in \{\mathrm{NCR},\mathrm{ET}\}} \left[\mathcal{L}_{\text{grouped-BCE}}(\mathbf{q}_c, \mathbf{y}_c) + \lambda_{\mathrm{rank}} \mathcal{L}_{\mathrm{rank}}(\mathbf{q}_c, \mathbf{y}_c)\right]. \tag{10}$$

The grouped BCE balances positive and negative slice groups, while $\mathcal{L}_{\mathrm{rank}}$ encourages positive slices to receive higher confidence than hard negatives. The overall fine-tuning objective is

$$\mathcal{L}_{\mathrm{ft}} = \mathcal{L}_{\mathrm{seg}} + \lambda_{\mathrm{pres}} \mathcal{L}_{\mathrm{pres}}. \tag{11}$$

Soft Gating Inference. Because $\mathrm{TC} = \mathrm{NCR} \cup \mathrm{ET}$, its slice-level presence confidence is defined as

$$q_{z,\mathrm{TC}} = \max(q_{z,\mathrm{NCR}}, q_{z,\mathrm{ET}}). \tag{12}$$

At inference, TC and ET predictions are softly modulated by their corresponding presence evidence:

$$\tilde{s}_{\mathrm{WT}}(v) = s_{\mathrm{WT}}(v), \qquad \tilde{s}_{\mathrm{TC}}(v) = q_{z(v),\mathrm{TC}}\, s_{\mathrm{TC}}(v), \qquad \tilde{s}_{\mathrm{ET}}(v) = q_{z(v),\mathrm{ET}}\, s_{\mathrm{ET}}(v), \tag{13}$$

where $z(v)$ denotes the axial slice containing voxel v. This coupling uses slice-level presence confidence to attenuate unsupported TC and ET responses while retaining the voxel-wise localization provided by the segmentation branch.

5 Experiments

5.1 Experimental Setup

We conduct all experiments on the BraTS 2024 brain tumor dataset de Verdier et al. (2024) using four MRI modalities (T1ce, FLAIR, T1, and T2). The original BraTS 2024 annotations contain four voxel-wise tumor labels. Following the BraTS evaluation convention, we derive three nested segmentation targets from the original four voxel-wise labels: WT = NCR/NET $\cup$ ED $\cup$ ET, TC = NCR/NET $\cup$ ET, and ET = ET, with the resection-cavity label excluded from evaluation. We use 100 subjects for label-free pretraining, 50 labeled subjects for fine-tuning, 20 subjects for validation, and 20 held-out subjects for testing.

Since the BraTS dataset does not provide raw complex-valued measurements, our experiments rely on image-derived synthetic k-space. Specifically, we construct retrospective fully sampled k-space using a centered orthonormal 2D FFT independently for each axial slice. Then, we apply random Cartesian undersampling at $R = 4$ while retaining 8% of the central phase-encoding lines to construct the corresponding undersampled k-space. The same sampling protocol is used throughout pretraining and fine-tuning, with deterministic masks used for evaluation. We train on contiguous 3D windows and use a six-stage 3D nnUNet-based PlainConvUNet as the image backbone. We set the spectral-prior loss weight to 0.25 during pretraining, and during fine-tuning, use $\mathcal{L}_{\mathrm{seg}} + 0.1\,(\mathcal{L}_{\text{grouped-BCE}} + 0.5\mathcal{L}_{\mathrm{ranking}})$.

We compare our method with representative end-to-end 3D segmentation architectures, including 3D U-Net (Çiçek et al., 2016), nnUNet (Isensee et al., 2021), MedNeXt (Roy et al., 2023), and

Table 1: Comparison with representative end-to-end and pretraining-based brain tumor segmentation methods on the BraTS 2024 benchmark.

Paradigm	Method	WT		TC			ET			Average		
		Dice ↑	HD95 ↓	Dice ↑	NE Dice ↑	HD95 ↓	Dice ↑	NE Dice ↑	HD95 ↓	Dice ↑	NE Dice ↑	HD95 ↓
End-to-end	3D U-Net	0.661	26.717	0.261	0.414	35.575	0.218	0.345	36.197	0.380	0.473	32.829
	nnUNet	0.702	**18.345**	0.440	0.484	16.129	0.317	0.445	16.297	0.487	0.544	16.924
	MedNeXt	0.671	26.935	0.457	0.429	18.900	0.429	0.381	19.597	0.519	0.493	21.811
	SwinUNETR	0.711	18.609	0.297	0.495	38.077	0.261	0.434	38.480	0.423	0.547	31.722
Pretrain + FT	AMAES (U-Net-XL)	0.668	19.129	0.281	0.398	28.149	0.227	0.341	30.137	0.392	0.469	25.805
	AMAES (MedNeXt-L)	0.672	24.263	0.285	0.451	34.073	0.256	0.406	34.222	0.404	0.509	30.853
	KMAE	0.635	27.974	0.243	0.406	61.496	0.207	0.345	61.917	0.362	0.462	50.462
	Ours	**0.727**	19.724	**0.559**	**0.514**	**8.813**	**0.518**	**0.447**	**9.543**	**0.601**	**0.563**	**12.693**

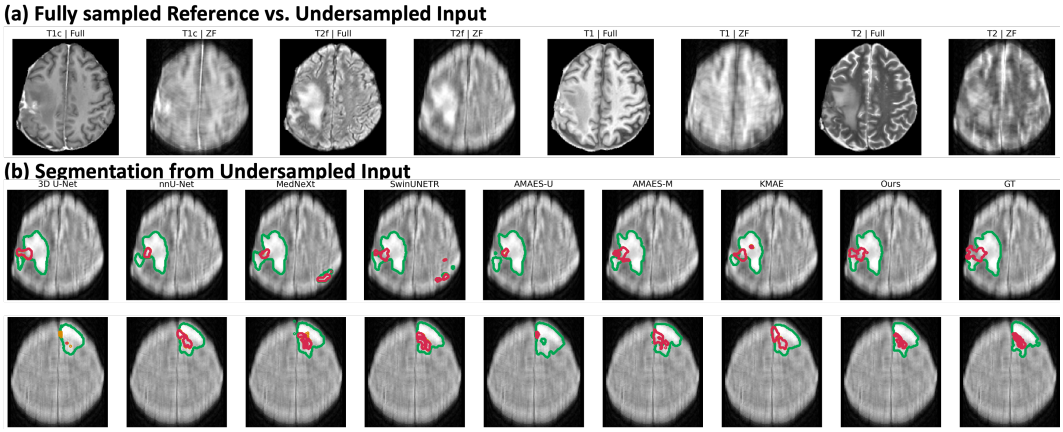

Figure 4: **Qualitative comparison under undersampled MRI acquisition. (a)** Fully sampled and zero-filled (ZF) undersampled images across four modalities. **(b)** Segmentation results of different methods on two representative cases using the undersampled inputs. Colored contours indicate WT (green), TC (yellow), and ET (red), with the ground truth shown in the last column.

SwinUNETR (Hatamizadeh et al., 2021), as well as self-supervised pretraining methods, including AMAES (Munk et al., 2024) and the k-space-based KMAE (Li et al., 2024). For AMAES, we directly fine-tune the publicly released pretrained MedNeXt-L and U-Net-XL checkpoints on our dataset. Our method jointly takes the resulting zero-filled images and complex undersampled k-space as input, whereas image-only baselines are trained on the same zero-filled images. For KMAE, which takes only undersampled k-space as input, we freeze its pretrained reconstruction network and replace the original single-layer 3D convolutional segmentation head with a full 3D nnUNet, enabling stable optimization under our limited-label setting. We evaluate segmentation performance using the Dice score and the 95th-percentile Hausdorff distance (HD95), following the BraTS evaluation protocol Menze et al. (2014). Subject Dice is averaged over all test subjects, whereas NE Dice and HD95 are computed on cases with nonempty ground-truth regions.

5.2 Comparison

Table 1 summarizes the comparison with representative end-to-end and pretraining-based methods. Our method achieves the strongest overall performance across the three evaluation metrics. Relative to the strongest end-to-end baseline for each metric, it improves the average subject-level Dice by 8.2% and reduces the average nonempty HD95 by 25.0%, with particularly clear gains on the more challenging TC and ET regions. It also consistently surpasses the pretraining-based AMAES and k-space-based KMAE, supporting the effectiveness of pathology-associated spectral information as an acquisition-informed signal for label-efficient tumor segmentation. The qualitative examples in Fig. 4 further show that, despite substantial degradation in the undersampled inputs, our method produces tumor delineations that better align with the ground truth while exhibiting fewer spurious predictions than the competing methods.

Table 2: Ablation study of the proposed spectral-guided pretraining and presence-aware adaptation.

Variant	WT		TC				ET				Average		
	Dice ↑	HD95 ↓	Dice ↑	NE Dice ↑	HD95 ↓	Absent FP ↓	Dice ↑	NE Dice ↑	HD95 ↓	Absent FP ↓	Dice ↑	NE Dice ↑	HD95 ↓
Recon-only	0.718	21.520	0.477	0.461	34.251	0.500	0.437	0.395	34.954	0.500	0.544	0.525	30.242
+ Spectral Prior	0.725	19.934	0.302	0.503	21.639	1.000	0.258	0.430	22.451	1.000	0.428	0.553	21.341
+ Presence Head	**0.727**	**19.724**	0.502	0.503	14.844	0.500	0.458	0.430	15.303	0.500	0.562	0.553	16.624
+ Soft Gating	**0.727**	**19.724**	**0.559**	**0.514**	**8.813**	**0.375**	**0.518**	**0.447**	**9.543**	**0.375**	**0.601**	**0.563**	**12.693**

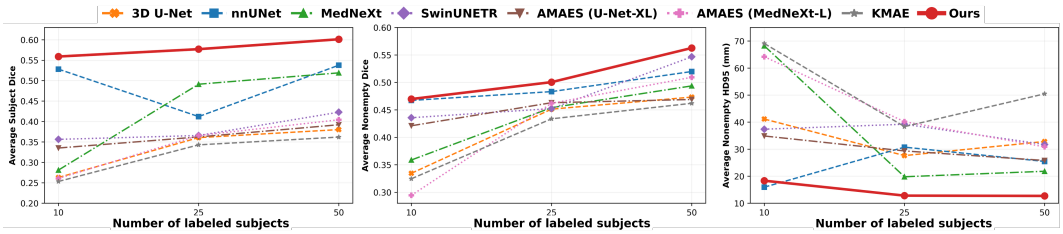

Figure 5: **Label-efficiency analysis under different annotation budgets.** Average subject-level Dice, NE Dice, and NE HD95 are reported using 10, 25, and 50 labeled subjects for fine-tuning.

5.3 Ablation Study

Table 2 shows that the two key components address different aspects of label-efficient segmentation. Introducing the spectral prior during pretraining improves positive-case segmentation, increasing the mean nonempty Dice from 0.525 to 0.553 and reducing nonempty HD95 from 30.242 to 21.341 mm. This supports the role of the spectral objective in preserving pathology-associated acquisition-domain information beyond reconstruction supervision alone. However, the higher false-positive rate on absent cases indicates that such spectral evidence is insufficient to determine lesion existence by itself. The presence-aware adaptation complements this representation by incorporating slice-level existence evidence into voxel-wise prediction, reducing the TC/ET absent-case false-positive rate from 1.000 to 0.375 and raising the mean subject Dice from 0.428 to 0.601. Together, the results support the intended division of roles in our framework: spectral-guided pretraining enriches lesion-relevant representations, while presence-aware fine-tuning improves their reliability under limited supervision.

5.4 Label Efficiency Analysis

To further evaluate performance under different annotation budgets, we fine-tune all methods using 10, 25, and 50 labeled subjects. As shown in Fig. 5, our method consistently achieves the highest average subject-level Dice across all label budgets, demonstrating strong robustness in the low-label regime. It also maintains competitive nonempty Dice and substantially lower nonempty HD95 as the number of labeled subjects increases, indicating that the proposed acquisition-informed pretraining and presence-aware adaptation improve both overall segmentation reliability and positive-case localization with limited supervision.

6 Conclusion

We presented an acquisition-informed framework for label-efficient brain tumor segmentation from undersampled multi-modal MRI. Our study reveals that pathology-associated spectral structure in k-space can provide lesion-relevant cues beyond conventional image-domain representations. We exploit this information through a soft spectral prior during self-supervised pretraining and further introduce presence-aware adaptation to couple spectral evidence with dense segmentation, reducing unsupported false-positive predictions under limited supervision. Experiments demonstrate consistent improvements over both end-to-end and pretraining-based baselines, particularly for challenging tumor subregions. These results highlight the value of explicitly modeling acquisition-domain information for robust and label-efficient MRI segmentation.

STATEMENTS

AI USE STATEMENT

Generative AI tools, including ChatGPT, Gemini and coding assistants, were used as assistive tools during this work for literature discussion, refinement of experimental design, code development and debugging, organization of experimental results, and drafting and polishing of the manuscript. The research question, empirical observations motivating this work, core methodological ideas, and proposed framework were developed by the authors; generative AI was used to assist with their refinement, implementation, and presentation rather than to autonomously originate the proposed method.

All reported experiments were conducted by executing the described training and evaluation pipelines on the stated MRI data. All AI-assisted manuscript content was reviewed and revised by the authors. The authors made the final methodological, experimental, and reporting decisions and take full responsibility for the content of this work.

ETHICS STATEMENT

This study uses de-identified brain MRI data released through the BraTS 2024 challenge. No new patient data were collected, and the authors had no access to personally identifiable information. The data were used in accordance with the applicable BraTS data-use terms.

REPRODUCIBILITY STATEMENT

Details of the data splits, preprocessing, model architectures, training settings, and evaluation protocols are provided in the main paper. To facilitate independent verification of the reported results, we provide the model checkpoints for all methods evaluated in our experiments, together with the corresponding testing code and evaluation data, in an anonymous repository: `https://anonymous.4open.science/r/Learning-from-K-space_ICLR_36898`.

REFERENCES

Natalie Baddour. Two-dimensional fourier transforms in polar coordinates. In *Advances in Imaging and Electron Physics*, volume 165, pp. 1–45. Elsevier, 2011.

Jose Bernal, Maria del C Valdés-Hernández, Javier Escudero, Linda Viksne, Anna K Heye, Paul A Armitage, Stephen Makin, Rhian M Touyz, and Joanna M Wardlaw. Analysis of dynamic texture and spatial spectral descriptors of dynamic contrast-enhanced brain magnetic resonance images for studying small vessel disease. *Magnetic Resonance Imaging*, 66:240–247, 2020.

Beatrice Bonato, Loris Nanni, and Alessandra Bertoldo. Advancing precision: a comprehensive review of mri segmentation datasets from brats challenges (2012–2025). *Sensors*, 25(6):1838, 2025.

Krishna Chaitanya, Ertunc Erdil, Neerav Karani, and Ender Konukoglu. Contrastive learning of global and local features for medical image segmentation with limited annotations. *Advances in neural information processing systems*, 33:12546–12558, 2020.

Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pp. 424–432. Springer, 2016.

Maria Correia de Verdier, Rachit Saluja, Louis Gagnon, Dominic LaBella, Ujjwall Baid, Nourel Hoda Tahon, Martha Foltyn-Dumitru, Jikai Zhang, Maram Alafif, Saif Baig, et al. The 2024 brain tumor segmentation (brats) challenge: Glioma segmentation on post-treatment mri. *arXiv preprint arXiv:2405.18368*, 2024.

Fatemeh Haghighi, Mohammad Reza Hosseinzadeh Taher, Zongwei Zhou, Michael B Gotway, and Jianming Liang. Transferable visual words: Exploiting the semantics of anatomical patterns for self-supervised learning. *IEEE Transactions on Medical Imaging*, 40(10):2857–2868, 2021.

Ali Hatamizadeh, Vishwesh Nath, Yucheng Tang, Dong Yang, Holger R Roth, and Daguang Xu. Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In *International MICCAI brainlesion workshop*, pp. 272–284. Springer, 2021.

Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. Masked autoencoders are scalable vision learners. In *2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pp. 15979–15988. IEEE, 2022.

Qiaoying Huang, Xiao Chen, Dimitris Metaxas, and Mariappan S Nadar. Brain segmentation from k-space with end-to-end recurrent attention network. In *International conference on medical image computing and computer-assisted intervention*, pp. 275–283. Springer, 2019.

Wentao Huang, Xiaoling Hu, Shahira Abousamra, Prateek Prasanna, and Chao Chen. Hard negative sample mining for whole slide image classification. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 144–154. Springer, 2024.

Fabian Isensee, Paul F Jaeger, Simon AA Kohl, Jens Petersen, and Klaus H Maier-Hein. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2):203–211, 2021.

Bo Li, Xinge You, Qinmu Peng, Jing Wang, and Chuanwu Yang. Region-related focal loss for 3d brain tumor mri segmentation. *Medical physics*, 50(7):4325–4339, 2023.

Ruochen Li, Jiazhen Pan, Youxiang Zhu, Juncheng Ni, and Daniel Rueckert. Classification, regression and segmentation directly from k-space in cardiac mri. In *International Workshop on Machine Learning in Medical Imaging*, pp. 31–41. Springer, 2024.

Michael Lustig, David Donoho, and John M Pauly. Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007.

Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging*, 34 (10):1993–2024, 2014.

Asbjørn Munk, Jakob Ambsdorf, Sebastian Llambias, and Mads Nielsen. Amaes: Augmented masked autoencoder pretraining on public brain mri data for 3d-native segmentation. *arXiv preprint arXiv:2408.00640*, 2024.

Lukas T Rotkopf, Marco Schlimbach, Julius C Holzschuh, Heinz-Peter Schlemmer, Jens Kleesiek, and Moritz Rempe. Robustness of breast lesion segmentation under mri undersampling improves with k-space-aware deep learning. *arXiv preprint arXiv:2605.22327*, 2026.

Saikat Roy, Gregor Koehler, Constantin Ulrich, Michael Baumgartner, Jens Petersen, Fabian Isensee, Paul F Jaeger, and Klaus H Maier-Hein. Mednext: transformer-driven scaling of convnets for medical image segmentation. In *International conference on medical image computing and computer-assisted intervention*, pp. 405–415. Springer, 2023.

Abraham Savitzky and Marcel JE Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8):1627–1639, 1964.

Jo Schlemper, Ozan Oktay, Wenjia Bai, Daniel C Castro, Jinming Duan, Chen Qin, Jo V Hajnal, and Daniel Rueckert. Cardiac mr segmentation from undersampled k-space using deep latent representation learning. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 259–267. Springer, 2018.

Yucheng Tang, Dong Yang, Wenqi Li, Holger R Roth, Bennett Landman, Daguang Xu, Vishwesh Nath, and Ali Hatamizadeh. Self-supervised pre-training of swin transformers for 3d medical image analysis. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20698–20708. IEEE, 2022.

Aniket A Tolpadi, Upasana Bharadwaj, Kenneth T Gao, Rupsa Bhattacharjee, Felix G Gassert, Johanna Luitjens, Paula Giesler, Jan Nikolas Morshuis, Paul Fischer, Matthias Hein, et al. K2s challenge: from undersampled k-space to automatic segmentation. *Bioengineering*, 10(2):267, 2023.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Tassilo Wald, Constantin Ulrich, Stanislav Lukyanenko, Andrei Goncharov, Alberto Paderno, Maximilian Miller, Leander Maerkisch, Paul Jaeger, and Klaus Maier-Hein. Revisiting mae pretraining for 3d medical image segmentation. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5186–5196. IEEE, 2025.

Shanshan Wang, Zhenghang Su, Leslie Ying, Xi Peng, Shun Zhu, Feng Liang, Dagan Feng, and Dong Liang. Accelerating magnetic resonance imaging via deep learning. In *2016 IEEE 13th international symposium on biomedical imaging (ISBI)*, pp. 514–517. IEEE, 2016.

Zhiwen Wang, Bowen Li, Hui Yu, Zhongzhou Zhang, Maosong Ran, Wenjun Xia, Ziyuan Yang, Jingfeng Lu, Hu Chen, Jiliu Zhou, et al. Promoting fast mr imaging pipeline by full-stack ai. *Iscience*, 27(1), 2024.

Linshan Wu, Jiaxin Zhuang, and Hao Chen. Voco: A simple-yet-effective volume contrastive learning framework for 3d medical image analysis. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 22873–22882. IEEE, 2024.

Linfang Xiao, Yilong Liu, Zheyuan Yi, Yujiao Zhao, Peiheng Zeng, TL Leong, and Ed X Wu. Deep learning pathology detection from extremely sparse k-space data. In *ISMRM (International Society of Magnetic Resonance Imaging) Virtual Conference & Exhibition, 2021*. Internationala Society of Magnetic Resonance Imaging (ISMRM)., 2021.

Yutong Xie, Jianpeng Zhang, Yong Xia, and Qi Wu. Unimiss: Universal medical self-supervised learning via breaking dimensionality barrier. In *European Conference on Computer Vision*, pp. 558–575. Springer, 2022.

Chen-Yu Yen, Raghav Singhal, Umang Sharma, Rajesh Ranganath, Sumit Chopra, and Lerrel Pinto. Adaptive sampling of k-space in magnetic resonance for rapid pathology prediction, 2024.

Yundi Zhang, Nil Stolt-Ansó, Jiazhen Pan, Wenqi Huang, Kerstin Hammernik, and Daniel Rueckert. Direct cardiac segmentation from undersampled k-space using transformers. In *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1–4. IEEE, 2024.

Yundi Zhang, Nil Stolt-Ansó, Jiazhen Pan, Wenqi Huang, Kerstin Hammernik, and Daniel Rueckert. Reconstruction-free segmentation from undersampled k-space using transformers. *arXiv preprint arXiv:2511.03762*, 2025.

Zongwei Zhou, Vatsal Sodha, Jiaxuan Pang, Michael B Gotway, and Jianming Liang. Models genesis. *Medical image analysis*, 67:101840, 2021.