# PATHOFIELD: CAPACITY-CONTROLLED FEATURE-SPACE RESIDUAL ADAPTATION FOR FEW-SHOT HEALTHY-TO-PATHOLOGICAL 3T-TO-7T MRI SYNTHESIS

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Paired pathological 3T–7T MRI is scarce, making adaptation of high-capacity synthesis models difficult in the few-shot regime. We introduce *PathoField*, a capacity-controlled transfer strategy that freezes a healthy-pretrained 3T-to-7T generator and learns four residual bottleneck adapters in decoder feature space. PathoField exposes only 45,696 trainable parameters (0.35% of the backbone), approximately $290\times$ fewer than unrestricted full fine-tuning, while preserving the pretrained mapping exactly at initialization. We further characterize the asymmetric optimization induced by zero initialization and derive a local capacity-dependent generalization analysis. On paired Penn epilepsy data, all reported test checkpoints are re-evaluated with a common all-slice spatial-SSIM/PSNR protocol. At $k = 16$, PathoField reaches 0.7941 SSIM / 21.15 dB with 45.7k trainable parameters, compared with 0.8015 / 21.12 dB for Full-FT with 13.25M parameters and 0.7117 / 21.10 dB for rank-3 Conv-LoRA with 54.9k parameters. Thus, on this fixed split, PathoField comes within 0.0074 SSIM of unrestricted fine-tuning while improving SSIM by 0.0824 over the comparable-budget weight-space baseline. A separate 23-subject frozen-transfer audit yields 0.4731 SSIM / 15.11 dB. These results reveal a data-dependent fidelity–capacity trade-off: strong capacity restriction incurs a clear cost in the lowest-data regimes, but at $k = 16$ PathoField recovers most of Full-FT structural fidelity with approximately $290\times$ fewer trainable parameters.

## 1 INTRODUCTION

Ultra-high-field 7T MRI can improve visualization of subtle structural abnormalities, including focal cortical dysplasia in epilepsy, but 7T scanners remain substantially less accessible than routine 3T systems (Colon et al., 2016; Veersema et al., 2017). This motivates 3T-to-7T synthesis: predict a 7T-like image from routinely acquired 3T MRI while preserving subject anatomy.

Prior field-strength synthesis work studies paired adversarial reconstruction (Nie et al., 2018), spatial–wavelet models (Qu et al., 2020), unpaired translation (Diniz et al., 2025), and lightweight state-space architectures (Zhang et al., 2026). Our setting is different: a useful paired healthy 3T–7T mapping is available, but adaptation must be performed on a much smaller pathological cohort. Unrestricted fine-tuning exposes millions of parameters to only a few pathological subjects, whereas freezing the source model prevents target-specific correction.

PEFT offers an intermediate regime. Bottleneck adapters modify features while freezing the backbone (Houlsby et al., 2019); LoRA constrains weight updates to low-rank factors (Hu et al., 2022). We study healthy-to-pathological transfer as a *capacity-allocation problem*: not only how many parameters to adapt, but where a small target-domain update should act. Given a 3T image $x$,

$$f_P(x) = G_{\theta_0}(x) + \Delta_P(x), \tag{1}$$

where $G_{\theta_0}$ is a healthy-pretrained mapping and $\Delta_P$ is the pathological-domain correction. PathoField freezes $G_{\theta_0}$ and restricts adaptation to a residual feature-space function class.



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.

**Our main contributions are:**

- **Capacity–data characterization:** we show that aggressive parameter restriction is not uniformly beneficial: PathoField underperforms unrestricted adaptation at the smallest verified target-data regimes, but approaches Full-FT at $k = 16$ with only 45,696 trainable parameters.
- **Feature- versus weight-space PEFT:** at $k = 16$, PathoField is compared with rank-3 Conv-LoRA at a comparable trainable budget, alongside Frozen, Scratch, and Full-FT endpoints.
- **Mechanistic analysis:** identity-preserving initialization, asymmetric first-step optimization, and a local capacity-dependent generalization bound.
- **Metric-harmonized evidence:** all main test checkpoints are rescored with one all-slice spatial-SSIM/PSNR protocol; at $k = 16$, PathoField attains 0.7941 SSIM versus 0.8015 for Full-FT and 0.7117 for the comparable-budget Conv-LoRA baseline.

## 2 Related Work

**MRI synthesis.** Medical image translation includes paired adversarial, transformer, diffusion, and hybrid approaches (Wang et al., 2021; Dayarathna et al., 2024; Dalmaz et al., 2022; Ozbey et al., 2023). For 3T-to-7T synthesis, prior studies learn the field-strength mapping using paired or unpaired data (Nie et al., 2018; Qu et al., 2020; Diniz et al., 2025; Zhang et al., 2026). We instead focus on adapting an existing healthy field-strength prior to pathology under severe paired-data scarcity.

**Parameter-efficient adaptation.** Adapters and LoRA avoid updating an entire pretrained model by restricting trainable capacity to task-specific modules or low-rank weight updates (Houlsby et al., 2019; Hu et al., 2022). PEFT is increasingly studied in low-data medical imaging (Dutt et al., 2024; Jang et al., 2024). PathoField uses the standard bottleneck-adapter primitive; the contribution is the healthy-to-pathological transfer formulation, comparable-budget feature- versus weight-space comparison, and capacity-focused analysis.

## 3 PathoField: Capacity-Controlled Adaptation

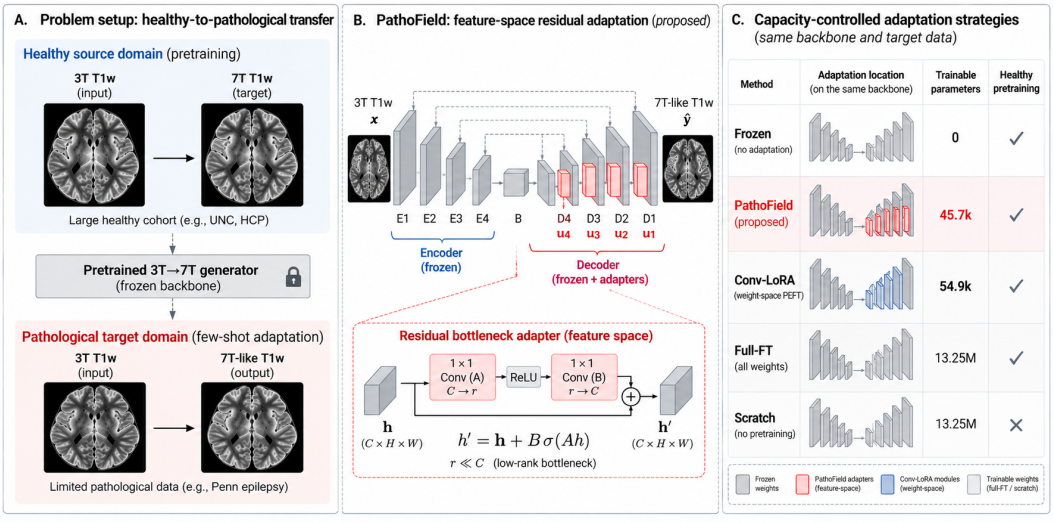



Figure 1: **PathoField and the capacity-controlled adaptation setting. (A)** A healthy paired 3T→7T T1w generator supplies the transfer prior. **(B)** PathoField freezes the backbone and inserts four residual bottleneck adapters after decoder blocks $u_4$–$u_1$, with $h' = h + B\sigma(Ah)$ and 45.7k trainable parameters. **(C)** The controlled study contrasts Frozen, PathoField feature-space PEFT, rank-3 Conv-LoRA weight-space PEFT, Full-FT, and Scratch.

### 3.1 HEALTHY PRIOR AND FEATURE-SPACE RESIDUAL ADAPTATION

Let $\mathcal{D}_H$ denote paired healthy UNC 3T–7T T1w data (Chen et al., 2023) and $\mathcal{D}_P$ paired Penn epilepsy data (Lucas et al., 2026). The source generator is a 2D residual-attention U-Net with four encoder/decoder levels, skip connections, residual blocks, a self-attention bottleneck, InstanceNorm, and Tanh output. It contains 13,201,668 parameters and is pretrained on UNC using a paired one-way GAN; the selected EMA generator (decay 0.999) defines $G_{\theta_0}$. Exact source-stage losses are given in Appendix E.

PathoField freezes $G_{\theta_0}$ and inserts residual bottleneck adapters after decoder blocks $u_4, u_3, u_2, u_1$. At level $l$,

$$A_l(h_l) = h_l + B_l\sigma(A_l^{\text{in}}h_l), \qquad A_l^{\text{in}} : C_l \to r_l,\ B_l : r_l \to C_l, \tag{2}$$

using $1 \times 1$ Conv2d projections and ReLU. For channels $(256, 128, 64, 64)$, $r_l = C_l/4$ gives ranks $(64, 32, 16, 16)$ and 45,696 trainable parameters. The input projection uses the default PyTorch initialization; the output projection weight and bias are zero-initialized, so each adapter is an exact identity at insertion.

During Penn adaptation, only adapter parameters $\phi$ are optimized with paired $\ell_1$ loss,

$$\mathcal{L}_{\text{target}}(\phi) = \|G_{\theta_0,\phi}(x^P) - y^P\|_1. \tag{3}$$

The discriminator and Stage-1 auxiliary losses are discarded, keeping the target-stage comparison controlled.

**Why adapt decoder features?** The adapter placement is deliberately confined to the decoder. The encoder, skip pathways, bottleneck, and pretrained convolutional weights remain unchanged, so the source representation and most of the healthy 3T-to-7T mapping are shared across target domains. The four trainable residual branches instead act on progressively decoded feature maps. This design makes PathoField a direct test of whether a low-dimensional correction in feature space is sufficient to specialize a strong source mapping without reopening the full network. Because the residual branch is zero at initialization, the first forward pass is exactly the healthy-pretrained generator rather than a randomly perturbed variant of it.

### 3.2 CONTROLLED BASELINES AND COMPARABLE-BUDGET PEFT

Frozen applies $G_{\theta_0}$ without Penn updates. Scratch trains the augmented generator from random initialization. Full-FT initializes from the UNC EMA model and updates all 13,247,364 parameters. Conv-LoRA provides the comparable-budget weight-space reference by freezing pretrained convolutions and learning low-rank updates $\Delta W = BA$. Rank 3 exposes 54,912 trainable parameters, close to PathoField's 45,696. The $k = 16$ experiment therefore compares feature-space and weight-space adaptation at similar budgets.

**What each baseline isolates.** The five regimes answer different transfer questions. Frozen measures how far the healthy prior transfers without target optimization. Scratch measures how much the Penn subset can support without source pretraining. Full-FT tests the upper-capacity transfer endpoint by allowing every Stage-2 parameter to move. Conv-LoRA keeps the pretrained backbone frozen but places a similar parameter budget in low-rank decoder-weight updates, whereas PathoField places that budget in residual decoder activations. The controlled $k = 16$ comparison therefore separates three factors that are otherwise confounded: access to the healthy initialization, total trainable capacity, and the location of the target-domain update.

## 4 LOCAL CAPACITY ANALYSIS

The analysis is a local lens on the class optimized by PathoField; it does not assume the full domain shift is globally low-rank.

**Proposition 1** (Initialization dynamics)**.** *If $B_l^{(0)} = 0$ for every adapter, then $G_{\theta_0,\phi_0}(x) = G_{\theta_0}(x)$. For $h' = h + B\sigma(Ah)$,*

$$\left.\frac{\partial \mathcal{L}}{\partial A}\right|_{B=0} = 0, \qquad \left.\frac{\partial \mathcal{L}}{\partial B}\right|_{B=0} = g\,\sigma(Ah)^\top,$$

Table 1: **Datasets and evaluation roles.** The frozen audit and fixed-split adaptation study answer different questions and are not pooled statistically.

| Cohort | Role | Evaluation scope |
|---|---|---|
| UNC healthy | Stage-1 pretraining | Paired 3T/7T T1w; supplies the source prior |
| Penn fixed split | Few-shot adaptation | 16 train-pool / 3 val / 4 held-out test subjects |
| Penn eligible cohort | Frozen transfer audit | 23 subjects; subject-level mean and 95% CI |

*so the first update acts only through the output projection unless the second term vanishes.*

Proofs are in Appendix F. For adapter level $l$,

$$p_l = 2C_l r_l + C_l + r_l, \qquad p_{\text{eff}}(\mathbf{r}) = \sum_l p_l. \tag{4}$$

Under the local boundedness and Lipschitz assumptions in Appendix J, with probability at least $1-\delta$,

$$\mathcal{R}_P(\phi) \le \widehat{\mathcal{R}}_P(\phi) + \widetilde{\mathcal{O}}\left( L_\ell \Lambda S \sqrt{\frac{p_{\text{eff}}(\mathbf{r})}{n}} \right) + 3\sqrt{\frac{\log(2/\delta)}{2n}}. \tag{5}$$

The bound highlights trainable adaptation capacity rather than frozen-backbone size. It motivates an approximation–estimation view, not a universal optimal-rank law. We test this capacity interpretation directly with the rank-by-data ablation in Section 6.7; because that sweep was archived under the legacy similarity evaluator, we use it only for within-protocol capacity trends and never mix its absolute values with the harmonized spatial-SSIM results.

## 5 Experimental Setup

**Datasets and transfer setting.** We use the paired healthy UNC 3T–7T T1w dataset (Chen et al., 2023) for Stage 1 pretraining and the paired Penn 3T–7T epilepsy dataset (Lucas et al., 2026) for target-domain adaptation and evaluation. The transfer setting is deliberately asymmetric: the source stage provides a reusable healthy field-strength prior, whereas the target stage exposes only a small number of paired pathological subjects. Penn volumes undergo N4 bias correction and ANTs-based skull stripping; 7T is rigidly registered and then nonlinearly refined to the 3T reference, followed by robust min–max normalization and 2D slice extraction. The model loader center-crops or pads slices to $256 \times 256$.

**Few-shot protocol.** The fixed-split study uses 23 Penn subjects: 16 in the full training pool, 3 validation subjects, and 4 held-out test subjects. Nested training subsets $k \in \{1, 2, 4, 8, 16\}$ are drawn from the same training pool (seed 2026). The main paper reports the fully verified corrected operating points $k \in \{1, 8, 16\}$, spanning one-shot, intermediate, and largest-data regimes. Scratch, Full-FT, and PathoField use the same held-out subjects. Rank-3 Conv-LoRA is evaluated on that same held-out split at $k = 16$, providing the comparable-budget weight-space PEFT reference. Separately, the frozen UNC generator is evaluated without Penn updates on all 23 eligible Penn subjects.

**Metric harmonization.** Images are evaluated in $[0, 1]$ using standard spatial SSIM (`skimage.metrics.structural_similarity`, `data_range=1.0`) and PSNR. For the fixed split, all 300 slices per held-out subject are scored, slice scores are averaged within subject, and method-level means are computed across the four test subjects. The archived checkpoints were selected during training with a legacy flattened/global validation score on a fixed validation subset; an audit confirmed that the stored validation values exactly reproduce that evaluator. We therefore use the legacy score only for historical checkpoint selection and rescore every main-test checkpoint with the harmonized spatial metric. Brain- and lesion-restricted metrics are not reported because exact annotation-to-evaluation-grid correspondence could not be verified.

**Statistical scope.** The 23-subject frozen audit is summarized with subject-level 95% confidence intervals. The fixed-split adaptation experiment uses only four held-out subjects and one subject split; its numbers are therefore descriptive operating-point estimates rather than cross-validated population estimates or equivalence tests. This distinction is maintained throughout the results and discussion.

**Training and checkpoint selection.** Stage 1 and Stage 2 serve different purposes. UNC pretraining learns the reusable field-strength mapping with adversarial, reconstruction, structural, perceptual, and frequency terms, while Penn adaptation uses only paired $\ell_1$ loss so that Scratch, Full-FT, PathoField, and Conv-LoRA are compared under the same target objective. The archived Penn checkpoints were selected during training with the original validation criterion; final paper metrics are obtained by loading those fixed checkpoints and rescoring every held-out slice with the common spatial evaluator. This distinction is important: model selection is historical, whereas the reported test metric is harmonized across all methods.

**Subject-level aggregation.** Each held-out Penn subject contributes equally to the final fixed-split mean. We first average SSIM/PSNR over that subject's 300 evaluated slices and only then average across the four test subjects. This prevents subjects with more evaluated slices from receiving greater statistical weight. For the 23-subject frozen audit, confidence intervals are likewise computed from one aggregated value per subject.

# 6 Results

## 6.1 Frozen healthy-to-pathological transfer

Without Penn adaptation, the frozen healthy generator obtains 0.4731 SSIM and 15.11 dB across 23 eligible Penn subjects. Table 2 summarizes subject-level uncertainty. This broader-cohort audit establishes a substantial source-to-target transfer gap, while remaining statistically separate from the four-subject fixed-split adaptation study.

Table 2: **Frozen UNC → Penn zero-shot transfer.** The 23 eligible subjects are partitioned only for reporting; the overall estimate uses one subject-level score per unique subject.

| Metric | Fold 0 | Fold 1 | Fold 2 | Fold 3 | Fold 4 | Overall (95% CI) |
|---|---|---|---|---|---|---|
| SSIM↑ | 0.4609 | 0.4600 | 0.4757 | 0.4719 | 0.5026 | 0.4731 [0.4608, 0.4855] |
| PSNR↑ (dB) | 15.31 | 16.58 | 14.97 | 13.37 | 14.95 | 15.11 [14.14, 16.09] |

## 6.2 Few-shot capacity–fidelity profile

Table 3 shows that PathoField is not uniformly superior. At $k = 1$, Full-FT reaches 0.7098 SSIM, Scratch 0.6868, and PathoField 0.6384; at $k = 8$, Full-FT and Scratch remain around 0.80 while PathoField reaches 0.7017. These gaps expose a real approximation cost when the target-domain correction is heavily constrained.

At $k = 16$, the picture changes: PathoField reaches 0.7941 SSIM, only 0.0074 below Full-FT at 0.8015 and close to Scratch at 0.7977. The feature-space adapter therefore recovers most of the unrestricted model's structural fidelity at the largest tested target-data setting while updating approximately $290\times$ fewer parameters. We interpret this as a data-dependent operating point, not as a claim that capacity restriction is beneficial in every few-shot regime.

**Reading the few-shot curve.** The corrected spatial metric changes the interpretation of the archived experiments. PathoField is clearly capacity-limited at the smallest and intermediate verified regimes: relative to Full-FT, the SSIM gaps are $0.0714$ at $k = 1$ and $0.1038$ at $k = 8$. By $k = 16$, that gap contracts to $0.0074$. Thus the result is not a monotonic "PEFT is better when data are scarce" story. Instead, the useful operating point emerges only when the target sample is large enough for the restricted residual class to estimate an effective correction. The three reported $k$ values should therefore be read as a capacity–data interaction rather than as isolated leaderboard entries.

Table 3: **Harmonized fixed-split spatial SSIM.** All methods share the same held-out subjects and spatial evaluator. Bold marks the best value within each verified $k$ column.

| Method | $k = 1$ | $k = 8$ | $k = 16$ |
|---|---|---|---|
| Scratch | 0.6868 | 0.8027 | 0.7977 |
| Full-FT | **0.7098** | **0.8055** | **0.8015** |
| **PathoField** | 0.6384 | 0.7017 | 0.7941 |

Table 4: **Controlled Penn comparison at $k = 16$.** All methods use the same held-out subjects and harmonized all-slice spatial-SSIM/PSNR evaluation.

| Method | Trainable params | SSIM↑ | PSNR↑ |
|---|---|---|---|
| Scratch | 13.25M | 0.7977 | 20.78 |
| Full-FT | 13.25M | **0.8015** | 21.12 |
| Conv-LoRA ($r = 3$) | 54.9k | 0.7117 | 21.10 |
| **PathoField** | **45.7k** | 0.7941 | **21.15** |

### 6.3 Comparable-budget $k = 16$ PEFT comparison and efficiency

Table 4 isolates feature-space versus weight-space PEFT at a similar trainable budget. PathoField reaches 0.7941 SSIM with 45.7k trainable parameters, compared with 0.7117 for rank-3 Conv-LoRA with 54.9k, a 0.0824 SSIM difference on this split. Full-FT remains slightly higher in SSIM (0.8015) but updates 13.25M parameters. PathoField, Full-FT, and Conv-LoRA have nearly identical PSNR (21.15, 21.12, and 21.10 dB), so the larger separation appears in structural similarity rather than average pixelwise error.

The controlled comparison exposes two complementary effects. First, PathoField comes within 0.0074 SSIM of unrestricted Full-FT while reducing the trainable state by approximately $290\times$. Second, at a comparable parameter budget, PathoField is 0.0824 SSIM above Conv-LoRA while using 16.8% fewer trainable parameters. Because the controlled adaptation study contains only four held-out subjects on one fixed split, these differences should be interpreted as operating-point differences rather than statistically established population effects or a universal ranking of PEFT mechanisms.

Table 5: **Capacity and adaptation efficiency.** Exact trainable parameter counts and raw FP32 trainable-state payload. Runtime and peak-memory measurements are omitted because no matched-hardware benchmark is available.

| Method | Trainable params | Trainable % | FP32 payload |
|---|---|---|---|
| Full-FT | 13,247,364 | 100% | 53.0 MB |
| Conv-LoRA ($r = 3$) | 54,912 | 0.41% | 0.22 MB |
| **PathoField** | **45,696** | **0.35%** | **0.18 MB** |

The payload in Table 5 counts only trainable FP32 parameters and excludes optimizer state, activations, frozen backbone parameters, and framework overhead. PathoField therefore reduces trainable capacity by 99.65% relative to Full-FT and by 16.8% relative to rank-3 Conv-LoRA; these analytic payloads are not presented as measured GPU memory.

### 6.4 Qualitative comparison

Figure 2 shows representative held-out $k = 16$ examples. The images illustrate synthesis behavior and residual structure, but do not establish lesion preservation or diagnostic validity.

### 6.5 Cross-study context

Published 3T-to-7T studies use heterogeneous datasets, preprocessing pipelines, pairing assumptions, pathology composition, and metric implementations. Table 6 is therefore contextual rather than a

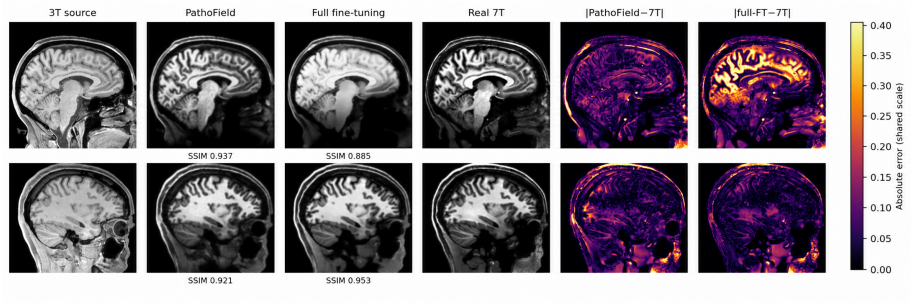



Figure 2: **Qualitative fixed-split comparison at $k = 16$.** Two held-out Penn subjects are shown using the 3T source, PathoField prediction, Full-FT prediction, paired real 7T target, and shared-scale absolute error maps.

common leaderboard. The four Penn rows are the only rows intended for direct comparison because they use the same held-out $k = 16$ evaluation split and harmonized spatial-SSIM/PSNR protocol.

Table 6: **Cross-study context for 3T→7T MRI synthesis.** External studies use heterogeneous protocols and are not directly comparable to the controlled Penn rows. NR: not reported.

| Method / study | Regime | Data | Pathol. | Trainable | SSIM↑ | PSNR↑ |
|---|---|---|---|---|---|---|
| CycleGAN (Zhu et al., 2017)† | Unpaired | UNC T1w | No | 28.29M | 0.679 | 19.98 |
| Nie et al. (Nie et al., 2018) | Paired GAN | 3T→7T T1w | NR | NR | NR | $27.93 \pm 1.18$ |
| Qu et al. (Qu et al., 2020) | Paired | 3T→7T T1w | NR | NR | 0.8782 | 28.27 |
| Diniz et al. (Diniz et al., 2025) | Unpaired | 3T→7T T1w | No | NR | NR | NR |
| LiteMamba-Synth (Zhang et al., 2026) | Paired | UNC T1w | No | 2.15M | 0.711 | 20.82 |
| Scratch (ours) | Scratch | Penn | Yes | 13.25M | 0.7977 | 20.78 |
| Full-FT (ours) | Transfer | UNC→Penn | Yes | 13.25M | **0.8015** | 21.12 |
| Conv-LoRA ($r = 3$) | PEFT | UNC→Penn | Yes | 54.9k | 0.7117 | 21.10 |
| **PathoField (ours)** | Feature PEFT | UNC→Penn | Yes | **45.7k** | 0.7941 | **21.15** |

†CycleGAN values are included only as published benchmark context. External rows must not be interpreted as a common leaderboard; direct conclusions in this paper rely on the controlled Penn rows.

### 6.6 Reproducibility and Evidence Scope

Tables 7 and 8 summarize the verified configuration and the evidence supporting the main claims.

Table 7: **Verified training configuration.**

| Setting | Configuration |
|---|---|
| Backbone | 2D residual-attention U-Net; 13.20M params |
| UNC objective | $L_{\text{adv}} + 10L_1 + 5(1 - \text{SSIM}) + 5L_{\text{VGG}} + L_{\text{freq}}$ |
| UNC optimization | Adam, lr $2 \times 10^{-4}$; EMA 0.999 |
| Stage-2 model | 13,247,364 total parameters |
| PathoField / Conv-LoRA | 45,696 / 54,912 trainable parameters |
| Adapter placement | after $u_4, u_3, u_2, u_1$ |
| Channels / ranks | $(256, 128, 64, 64)$ / $(64, 32, 16, 16)$ |
| Adapter init. | input projection: default; output projection: zero |
| Penn adaptation | $L_1$, Adam, lr $2 \times 10^{-4}$; fixed split |
| Few-shot sizes | $k \in \{1, 2, 4, 8, 16\}$ |

### 6.7 Exploratory Rank–Data Interaction

Table 9 moves the archived rank-by-data ablation into the main paper because it directly explains the capacity-control hypothesis. The table uses the original flattened/global similarity evaluator employed by the archived sweep, so its absolute values are *not* comparable with Tables 3 and 4. Its role is narrower: all entries inside Table 9 were produced under one common legacy protocol, allowing us to examine how the preferred adapter bottleneck changes with target-data availability.

Table 8: **Evidence scope for the main claims.** Fixed-split comparisons are descriptive operating-point estimates, not cross-validated population estimates.

| Experiment | Role | Reporting scope |
|---|---|---|
| Frozen UNC zero-shot | Transfer-gap baseline | 23 Penn subjects; spatial SSIM/PSNR with subject-level 95% CI. |
| Scratch / Full-FT / PathoField | Main adaptation study | Common fixed split; nested $k \in \{1, 2, 4, 8, 16\}$; all-slice spatial test rescoring. |
| Conv-LoRA ($r = 3$) | Comparable-budget weight-space PEFT | Same held-out Penn subjects at $k = 16$; 54,912 trainable parameters. |
| Adapter rank sweep | Main capacity ablation | Table 9; legacy-score within-protocol analysis at $k \in \{1, 4, 16\}$; absolute values not mixed with spatial SSIM. |

Table 9: **Exploratory fixed-split rank-by-data analysis (legacy evaluator).** The best value in each column is bold. These values are used only for within-table capacity trends and are not compared numerically with the harmonized spatial-SSIM results.

| Rank $r$ | $k = 1$ | $k = 4$ | $k = 16$ |
|---|---|---|---|
| 1 | 0.7455 | 0.8537 | 0.8645 |
| 2 | 0.7624 | 0.8685 | 0.8756 |
| 4 | **0.7773** | 0.8745 | 0.8836 |
| 8 | 0.7726 | 0.8826 | **0.8907** |
| 16 | 0.7729 | 0.8896 | 0.8897 |
| 32 | 0.7744 | **0.8958** | 0.8897 |

The empirical optimum is non-monotonic: $r = 4$ at $k = 1$, $r = 32$ at $k = 4$, and $r = 8$ at $k = 16$. At one subject, increasing rank beyond four does not improve the archived score, consistent with a regime in which additional capacity is not reliably supported by the available target evidence. At four subjects, the score continues to rise through the largest tested rank. At sixteen subjects, the curve improves through $r = 8$ and then becomes essentially flat. These patterns support the modest conclusion that useful adaptation capacity depends on the target-data regime, while contradicting any simple claim that optimal rank must increase monotonically with sample size.

This exploratory ablation is qualitatively consistent with the capacity trade-off motivating the theory, but it is not a validation of the bound or of a specific rank-scaling law. The empirical optima are non-monotonic across target-data regimes, indicating that useful adapter capacity depends on both available supervision and the optimization setting. We therefore interpret the theoretical analysis as a capacity lens rather than as a quantitative predictor of the optimal rank.

## 7 Discussion

The corrected evaluation reveals a data-dependent fidelity–capacity trade-off. PathoField trails unrestricted adaptation at $k = 1$ and $k = 8$, showing that aggressive capacity restriction can incur approximation error. At $k = 16$, however, it reaches 0.7941 SSIM, only 0.0074 below Full-FT, while updating approximately $290\times$ fewer parameters. The main result is therefore a compact operating point rather than universal superiority.

On this fixed split, feature-space residual adaptation achieves higher spatial SSIM than the comparable-budget rank-3 Conv-LoRA weight-space baseline, despite using fewer trainable parameters; this comparison remains specific to the tested backbone, LoRA placement, and subject split.

The frozen audit provides complementary evidence: without Penn updates, the healthy source model obtains 0.4731 SSIM / 15.11 dB across 23 eligible subjects. Healthy pretraining is therefore useful as a prior but insufficient by itself, and the $k = 1$ results show that the best adaptation mechanism also depends on target-data availability. The theoretical analysis should be read in the same limited sense: identity preservation and first-step asymmetry are exact architectural properties, while the generalization bound is a local capacity lens rather than a claim that pathology is globally low-rank.

**Feature-space versus weight-space adaptation.** The comparable-budget PEFT result is informative because the two compact methods have similar parameter counts but place those parameters in different objects. Conv-LoRA modifies decoder convolutional weights through low-rank updates,

whereas PathoField leaves every pretrained convolution fixed and adds a residual correction to decoder activations. The observed SSIM separation therefore cannot be explained simply by PathoField having a larger trainable budget; it actually uses fewer parameters. A plausible interpretation is that, for this particular source backbone and Penn split, direct feature correction provides a more suitable restricted hypothesis class than the chosen decoder-weight factorization. This interpretation is intentionally local to the tested configuration and does not imply that feature-space PEFT universally dominates LoRA.

**Why SSIM and PSNR tell different stories.** At $k = 16$, PathoField, Full-FT, and Conv-LoRA have similar PSNR but substantially different SSIM. PSNR is driven by mean squared intensity error, whereas SSIM is more sensitive to local structural relationships. The combination therefore suggests that the compared methods can achieve similar average pixelwise error while differing in how well local structure is reproduced. Because no lesion-aligned evaluation is available, we do not translate this observation into a pathology-preservation claim; it simply motivates reporting both metrics rather than relying on PSNR alone.

**Practical implication of the trainable budget.** PathoField requires only 45,696 task-specific parameters, corresponding to approximately 0.18 MB of raw FP32 trainable state, compared with about 53.0 MB for a complete fine-tuned Stage-2 state. This does not measure end-to-end GPU memory or wall-clock speed, but it matters when a common source backbone must support multiple target cohorts or sites: the shared 13.20M-parameter generator can remain fixed while each adaptation stores only a small residual state.

**Limitations.** The Penn cohort is small and the adaptation study uses one subject split with four held-out test subjects, so we do not make equivalence or population-level ranking claims. Archived checkpoints were selected with a legacy validation metric, although all main test scores are rescored with harmonized spatial SSIM. Conv-LoRA is evaluated only at $k = 16$, matched-hardware runtime/peak-memory measurements are unavailable, and verified lesion-level correspondence and expert-reader assessment remain future work. Whole-image SSIM/PSNR therefore support synthesis-fidelity claims only, not pathology preservation or clinical utility.

## 8 Conclusion

In this paper, we introduce PathoField, a capacity-controlled feature-space strategy for healthy-to-pathological 3T-to-7T MRI adaptation. With 45,696 trainable parameters, it preserves the source mapping at insertion and reaches 0.7941 SSIM / 21.15 dB at $k = 16$. Full-FT reaches 0.8015 / 21.12 dB while updating approximately $290\times$ more parameters; comparable-budget Conv-LoRA reaches 0.7117 / 21.10 dB. The results expose a data-dependent fidelity–capacity trade-off: strong restriction is costly at the lowest-data settings, but at $k = 16$ PathoField recovers most of Full-FT structural fidelity with a far smaller trainable state. Repeated-split and clinical validation remain necessary.

## AI USE STATEMENT

Generative AI tools assisted writing refinement, LaTeX formatting, code scaffolding, and literature search; all experimental results and scientific claims were verified by the authors.

## ETHICS STATEMENT

This work uses previously released neuroimaging datasets and introduces no new human-subject data collection. Synthetic images are not intended to establish diagnostic validity.

## REPRODUCIBILITY STATEMENT

The main paper specifies the method, parameter counts, subject split, harmonized evaluator, cross-study context, verified training configuration, and evidence scope (Tables 6–8). The rank-by-data capacity ablation is reported in the main paper (Table 9); detailed preprocessing, extended loss definitions, proofs, algorithmic pseudocode, detailed architecture, and additional qualitative results are provided in the appendix.

## REFERENCES

X. Chen, L. Qu, Y. Xie, S. Ahmad, and P.-T. Yap. UNC Paired 3T–7T Dataset (Version 2). Figshare, 2023. doi:10.6084/m9.figshare.23706033.v2.

A. J. Colon, M. J. P. van Osch, M. Buijs, J. van der Grond, P. Boon, M. A. van Buchem, and P. A. M. Hofman. Detection superiority of 7 t mri protocol in patients with epilepsy and suspected focal cortical dysplasia. *Acta Neurologica Belgica*, 116(3):259–269, 2016. doi: 10.1007/s13760-016-0662-x.

O. Dalmaz, M. Yurt, and T. Cukur. ResViT: Residual vision transformers for multimodal medical image synthesis. *IEEE Transactions on Medical Imaging*, 41(10):2598–2614, 2022. doi: 10.1109/TMI.2022.3167808.

S. Dayarathna, K. T. Islam, S. Uribe, G. Yang, M. Hayat, and Z. Chen. Deep learning based synthesis of mri, ct and pet: Review and analysis. *Medical Image Analysis*, 92:103046, 2024. doi: 10.1016/j.media.2023.103046.

E. Diniz, T. Santini, H. Karim, H. J. Aizenstein, and T. S. Ibrahim. Cross-modality image translation of 3 tesla magnetic resonance imaging to 7 tesla using generative adversarial networks. *Human Brain Mapping*, 46(9):e70246, 2025. doi: 10.1002/hbm.70246.

R. Dutt, L. Ericsson, P. Sanchez, S. A. Tsaftaris, and T. Hospedales. Parameter-efficient fine-tuning for medical image analysis: The missed opportunity. In *Proceedings of Medical Imaging with Deep Learning*, volume 250 of *PMLR*, pp. 406–425, 2024.

N. Houlsby, A. Giurgiu, S. Jastrzebski, B. Morrone, Q. de Laroussilhe, A. Gesmundo, M. Attariyan, and S. Gelly. Parameter-efficient transfer learning for nlp. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 2790–2799, 2019.

E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022.

U. Jang, J. D. Lee, and E. K. Ryu. LoRA training in the NTK regime has no spurious local minima. In *Proceedings of the 41st International Conference on Machine Learning*, 2024.

A. Lucas, M. Josyula, N. Petillo, J. B. Wagenaar, J. M. Stein, S. R. Das, N. Sinha, K. A. Davis, and T. Imtiaz. Penn 3T–7T Paired Epilepsy Imaging (Version 1). Pennsieve Discover, 2026. doi:10.26275/AOSZ-HJ9U.

D. Nie, R. Trullo, J. Lian, L. Wang, C. Petitjean, S. Ruan, Q. Wang, and D. Shen. Medical image synthesis with deep convolutional adversarial networks. *IEEE Transactions on Biomedical Engineering*, 65(12):2720–2730, 2018. doi: 10.1109/TBME.2018.2814538.

M. Ozbey, O. Dalmaz, S. U. H. Dar, H. A. Bedel, S. Ozturk, A. Gungor, and T. Cukur. Unsupervised medical image translation with adversarial diffusion models. *IEEE Transactions on Medical Imaging*, 42(12):3524–3539, 2023. doi: 10.1109/TMI.2023.3290149.

L. Qu, Y. Zhang, S. Wang, P.-T. Yap, and D. Shen. Synthesized 7t mri from 3t mri via deep learning in spatial and wavelet domains. *Medical Image Analysis*, 62:101663, 2020. doi: 10.1016/j.media.2020.101663.

T. J. Veersema, C. H. Ferrier, P. van Eijsden, P. H. Gosselaar, E. Aronica, F. Visser, J. M. Zwanenburg, G. A. P. de Kort, J. Hendrikse, P. R. Luijten, and K. P. J. Braun. Seven tesla mri improves detection of focal cortical dysplasia in patients with refractory focal epilepsy. *Epilepsia Open*, 2(2):162–171, 2017. doi: 10.1002/epi4.12041.

T. Wang, Y. Lei, Y. Fu, J. F. Wynne, W. J. Curran, T. Liu, and X. Yang. A review on medical imaging synthesis using deep learning and its clinical applications. *Journal of Applied Clinical Medical Physics*, 22(1):11–36, 2021. doi: 10.1002/acm2.13121.

Z. Zhang, J. Dong, H. Yang, and Z. Chen. LiteMamba-Synth: Lightweight state space models for efficient 3t-to-7t mri translation. *Frontiers in Neuroanatomy*, 20:1886476, 2026. doi: 10.3389/fnana.2026.1886476.

J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2223–2232, 2017.

# A Additional main-result details

## A.1 Residual adapter schematic

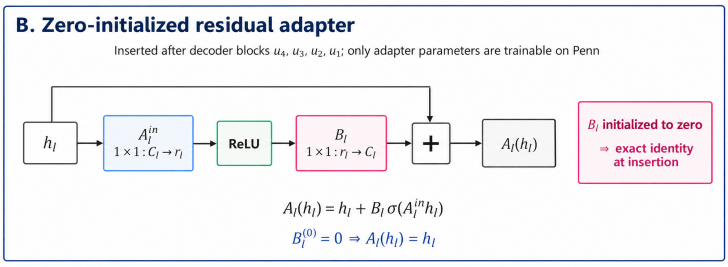



Figure 3: **Zero-initialized residual adapter.** A $1 \times 1$ down-projection, ReLU, and $1 \times 1$ up-projection define a residual feature correction. Zero-initializing the second projection makes the module an exact identity at insertion.

## A.2 Parameter-state accounting

Exact trainable parameter counts and raw FP32 payload are reported in main-paper Table 5. We do not duplicate that table here; runtime and peak-memory measurements remain unavailable on matched hardware.

# B Cross-study context

The complete cross-study context is now reported in the main paper (Table 6) because it is part of the requested conference-page evidence. External rows remain contextual only.

## C Detailed PathoField architecture

The four decoder outputs receiving adapters have channel dimensions $(256, 128, 64, 64)$ for $(u_4, u_3, u_2, u_1)$, respectively. With the default $r_l = C_l/4$, the corresponding adapter ranks are $(64, 32, 16, 16)$ and the exact trainable count is 45,696. Figure 4 provides a detailed overview of the complete PathoField architecture and its two-stage training strategy.

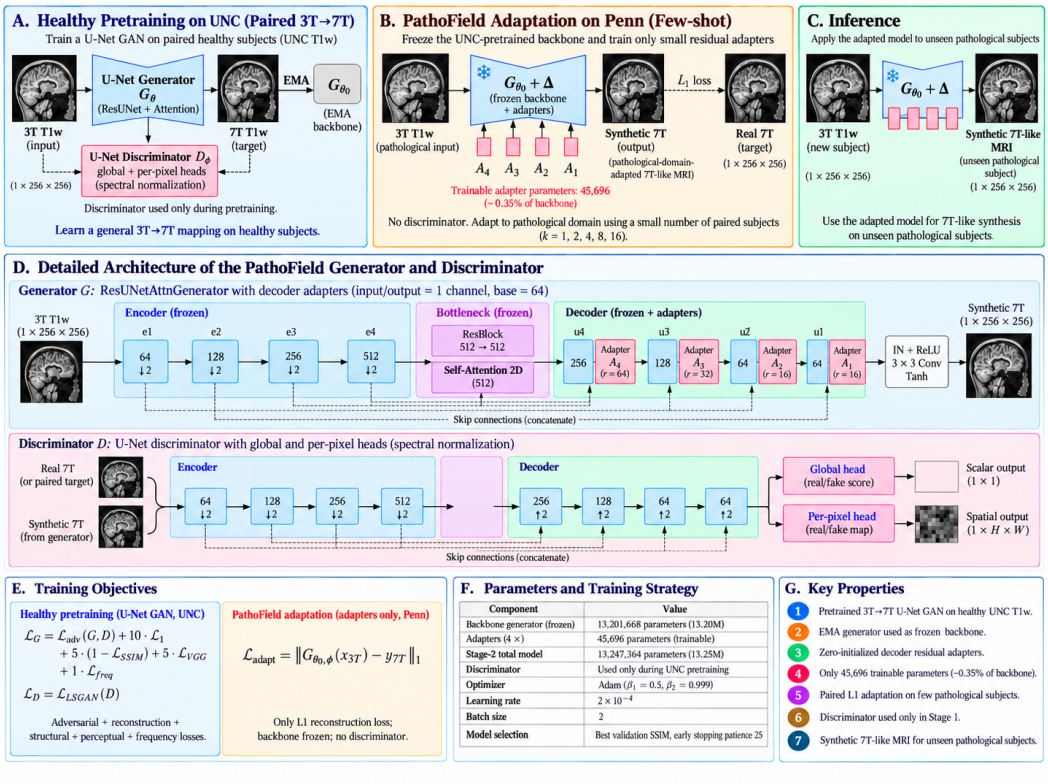



Figure 4: **PathoField architecture and two-stage training strategy.** The generator is a residual-attention U-Net with four encoder levels, a residual/self-attention bottleneck, and four decoder levels. PathoField inserts zero-initialized residual bottleneck adapters after decoder blocks $u_4, u_3, u_2, u_1$. During Stage-1 healthy-domain pretraining, the generator is trained with the U-Net discriminator; the discriminator is discarded before pathological-domain adaptation. During Stage 2, the pretrained generator backbone remains frozen and only the four adapters are optimized on the paired Penn pathological cohort. The frozen backbone contains 13,201,668 parameters (13.20M), while the four adapters introduce 45,696 trainable parameters, corresponding to approximately 0.35% of the backbone parameter count.

## D Datasets and preprocessing

**UNC cohort.** We use the public paired healthy UNC 3T–7T dataset (Chen et al., 2023) for Stage 1 pretraining. Only paired T1w images are used, with one-channel 3T T1w input and one-channel 7T T1w target.

**Penn cohort.** We use the public paired Penn 3T–7T epilepsy dataset (Lucas et al., 2026) for pathological-domain adaptation and evaluation. Both input and target are one-channel T1w images. The eligibility pipeline yields 23 subjects for the frozen-UNC zero-shot analysis. The completed few-shot experiments use a fixed subject split and are reported as fixed-split estimates rather than cross-validated results.

**Preprocessing.** For the Penn paired pipeline, 3T and 7T T1w images undergo N4 bias correction and ANTs-based skull stripping; 7T is rigidly registered and then nonlinearly refined to the 3T reference, followed by robust min–max normalization and 2D slice extraction. The model loader center-crops or pads slices to $256 \times 256$ and maps intensities to the network range. The stored slice indices are taken from this processed 3T-reference volume. Brain-masked metrics are not reported because exact correspondence between the released anatomical masks and the final NPZ evaluation grid could not be verified.

**Few-shot protocol.** The archived fixed-split study contains nested target-domain training subsets $k \in \{1, 2, 4, 8, 16\}$ for Scratch, Full-FT, and PathoField. The main paper reports representative corrected operating points at $k \in \{1, 8, 16\}$, recomputed from saved checkpoints using standard spatial SSIM over all held-out slices; the original flattened/global validation score is retained only as the historical checkpoint-selection criterion. The rank sweep remains explicitly labeled as a legacy-score diagnostic. Rank-3 Conv-LoRA is evaluated on the same held-out Penn split at $k = 16$ as the comparable-budget PEFT baseline.

## E Supplementary training and implementation details

This section records the implementation details needed to reproduce both stages, including adapter dimensions and initialization, the Penn fixed-split setup, and the exact Stage-1 auxiliary losses.

**Stage-specific objectives.** UNC pretraining uses

$$\mathcal{L}_G = \mathcal{L}_{\text{adv}} + 10\|G(x) - y\|_1 + 5\big(1 - \text{SSIM}(G(x), y)\big) + 5\mathcal{L}_{\text{VGG}} + \mathcal{L}_{\text{freq}}, \tag{6}$$

whereas Penn adaptation uses paired $\ell_1$ only. The selected UNC EMA generator supplies the frozen field-strength prior for PathoField; the Stage-2 correction is interpreted as data-driven domain adaptation rather than a guaranteed pathology-specific component.

**Controlled training protocol.** Within the archived fixed-split study, Scratch, Full-FT, and PathoField use the same stored subject split, target-stage objective, image resolution, and legacy validation-based model selection. Rank-3 Conv-LoRA uses the same held-out $k = 16$ evaluation split. For reporting, all fixed-split checkpoints are rescored with the same spatial-SSIM/PSNR implementation over all held-out slices.

For Stage 1, the perceptual term is $\mathcal{L}_{\text{VGG}} = \sum_{q \in \{3,8,15,22\}} \|\varphi_q(G(x)) - \varphi_q(y)\|_1$, where $\varphi_q$ denotes frozen VGG16 features; single-channel MRI inputs are replicated to three channels and normalized with ImageNet statistics. The frequency term is $\ell_1$ distance between orthonormal real-valued rFFT2 representations. LSGAN is applied to both discriminator heads with MSE targets 1 (real) and 0 (fake), without a sigmoid. Specifically, $\mathcal{L}^G_{\text{adv}} = \ell_{\text{LS}}(E(G(x)), 1) + \ell_{\text{LS}}(P(G(x)), 1)$ and $\mathcal{L}_D = \frac{1}{2}[\ell_{\text{LS}}(E(y), 1) + \ell_{\text{LS}}(P(y), 1) + \ell_{\text{LS}}(E(G(x)), 0) + \ell_{\text{LS}}(P(G(x)), 0)]$.

**Configuration table.** The verified training configuration is reported in main-paper Table 7; this appendix retains only the extended loss definitions and implementation notes.

## F Proofs of exact initialization properties

**Proposition 1.** For every layer, Eq. (2) gives $A_l(h_l) = h_l + B_l^{(0)}\sigma(A_l^{\text{in}} h_l) = h_l$. Hence all inserted modules are exact identity maps and the composed generator function is unchanged.

**Proposition 2.** By the chain rule,

$$\frac{\partial \mathcal{L}}{\partial A} = \big(B^\top g \odot \sigma'(Ah)\big)\, h^\top, \tag{7}$$

which vanishes at $B = 0$. Meanwhile, $\partial\mathcal{L}/\partial B = g\sigma(Ah)^\top$ does not contain $B$ and is generally nonzero.

## G PATHOFIELD ALGORITHM

Algorithm 1 summarizes the full two-stage procedure: healthy residual-attention U-Net GAN pretraining (Stage 1) followed by frozen-backbone few-shot adapter adaptation (Stage 2). Stage 1 produces the EMA generator $G_{\theta_0}$ that is reused, unchanged, as the frozen backbone in Stage 2.

**Algorithm 1** PathoField: capacity-controlled residual adaptation

**Require:** Healthy paired data $\mathcal{D}_H$; pathological paired data $\mathcal{D}_P$ with $k$ training subjects; adapter ranks $\{r_l\}_{l=1}^{L}$; loss weights $\lambda_{L1}, \lambda_{\text{SSIM}}, \lambda_{\text{perc}}, \lambda_{\text{freq}}$; EMA decay $\mu$; learning rates $\eta_G, \eta_D, \eta_\phi$.

***Stage 1 — Healthy residual-attention U-Net GAN pretraining (on $\mathcal{D}_H$)***

Initialize generator $G_\theta$, discriminator $D$, EMA copy $\theta_0 \leftarrow \theta$
**for** each pretraining step **do**
  Sample paired batch $(x^H, y^H) \sim \mathcal{D}_H$; $\hat{y} \leftarrow G_\theta(x^H)$
  $\mathcal{L}_G \leftarrow \mathcal{L}_{\text{adv}}(D, \hat{y}) + \lambda_{L1}\|\hat{y} - y^H\|_1 + \lambda_{\text{SSIM}}(1 - \text{SSIM}) + \lambda_{\text{perc}}\mathcal{L}_{\text{VGG}} + \lambda_{\text{freq}}\mathcal{L}_{\text{freq}}$
  $\theta \leftarrow \theta - \eta_G \nabla_\theta \mathcal{L}_G$; update $D$ with $\eta_D$
  $\theta_0 \leftarrow \mu\,\theta_0 + (1-\mu)\,\theta$ ▷ EMA generator
  Track best $\theta_0$ by validation SSIM (early stopping)
**end for**
**return** frozen EMA backbone $G_{\theta_0}$

***Stage 2 — Few-shot adapter adaptation (on $\mathcal{D}_P$)***

Freeze $\theta_0$: set `requires_grad`=False for all backbone params
**for** each decoder block $l \in \{u_4, u_3, u_2, u_1\}$ **do**
  Insert adapter $A_l(h) = h + B_l\,\sigma(A_l^{\text{in}} h)$, $A_l^{\text{in}}: C_l \to r_l$, $B_l: r_l \to C_l$
  Zero-initialize $B_l$ ▷ $A_l$ is identity at init (Prop. 1)
**end for**
Collect trainable adapter parameters $\phi = \{A_l^{\text{in}}, B_l\}$
**for** each adaptation epoch **do**
  Sample paired batch $(x^P, y^P)$ from the $k$-subject subset
  $\hat{y} \leftarrow G_{\theta_0,\phi}(x^P)$ ▷ backbone frozen, adapters active
  $\mathcal{L}_{\text{target}} \leftarrow \|\hat{y} - y^P\|_1$ ▷ $\ell_1$ only, no discriminator
  $\phi \leftarrow \phi - \eta_\phi \nabla_\phi \mathcal{L}_{\text{target}}$
  Select best $\phi$ by archived validation similarity criterion (early stopping)
**end for**
**return** adapted model $G_{\theta_0,\phi}$

## H THEORY SCOPE AND REPRODUCIBILITY NOTES

**Scope of the theory.** Theorem 1 is local and assumption-dependent: ReLU gates can change during training, the frozen downstream map is nonlinear, and subject-level medical data are not i.i.d. at the slice level. We therefore interpret the result as a capacity analysis of the local adapter class rather than a global guarantee for the complete optimization trajectory.

**Reproducibility.** The paper reports the UNC generator/discriminator configuration, composite generator loss and weights, Adam parameters, learning rates, batch size, EMA decay, adapter equations, parameter counts, Penn objective, the archived fixed-split few-shot study, and the separate 23-subject zero-shot evaluation. Scratch, Full-FT, and PathoField use one common archived split; rank-3 Conv-LoRA uses the same held-out split at $k = 16$. The archived checkpoints were selected with the original flattened/global validation score, while all main fixed-split test metrics are harmonized by rescoring the saved checkpoints with standard spatial SSIM over all held-out slices. The rank-by-data grid is reported in main-paper Table 9 as an explicitly labeled legacy-score ablation and is not mixed numerically with the harmonized spatial-SSIM results.

## I IDEALIZED APPROXIMATION–ESTIMATION INTERPRETATION

The main text uses Theorem 1 as a local capacity lens and does not claim a monotonic empirical rank law. For completeness, an additional idealized approximation-error model gives the following heuris-

tic. Let $\epsilon_{\text{app}}(r)$ denote approximation error for a common rank scale $r$. Combining approximation and estimation terms gives

$$\mathcal{R}(\hat{\phi}_r) - \mathcal{R}(f_P) \lesssim \epsilon_{\text{app}}(r) + K\sqrt{\frac{r}{n}}, \tag{8}$$

when the rank-dependent part of $p_{\text{eff}}$ is approximately linear in $r$.

If one further assumes a power law $\epsilon_{\text{app}}(r) = ar^{-\alpha}$ with $a, \alpha > 0$, minimizing the right-hand side yields

$$r^*(n) \propto n^{\frac{1}{2\alpha+1}}. \tag{9}$$

This monotonic relation depends on the additional power-law assumption and is *not* validated by the current rank sweep, whose empirical optima are non-monotonic. We retain it only as an idealized approximation–estimation interpretation.

## J Formal assumptions for the capacity bound

**Assumption 1** (Local fixed-gating regime)**.** *During the analyzed neighborhood of adaptation, the ReLU activation pattern of each adapter is fixed for every training example and spatial location. Gates may differ across examples and locations. The shared trainable objects remain the adapter parameters* $\phi = \{A_l^{\text{in}}, B_l\}_{l=1}^{L}$.

**Assumption 2** (Bounded frozen features and downstream sensitivity)**.** *For each adapter level* $l$*, frozen features satisfy* $\|h_l(x)\|_2 \le H_l$*. The frozen downstream mapping from the adapter output to the synthesized image is locally* $J_l$*-Lipschitz in the analyzed neighborhood (equivalently, its local Jacobian spectral norm is bounded by* $J_l$ *wherever the Jacobian exists).*

**Assumption 3** (Bounded local parameter set)**.** *Let* $\phi_0$ *denote the adapter initialization. The analyzed adapter parameters lie in a bounded neighborhood* $\Phi = \{\phi : \|\phi - \phi_0\|_2 \le S\} \subset \mathbb{R}^{p_{\text{eff}}(\mathbf{r})}$.

**Assumption 4** (Local parameter sensitivity)**.** *For all relevant inputs* $x$ *and all* $\phi, \phi' \in \Phi$*, the prediction map satisfies*

$$\|G_{\theta_0,\phi}(x) - G_{\theta_0,\phi'}(x)\|_2 \le \Lambda\|\phi - \phi'\|_2.$$

*The constant* $\Lambda$ *summarizes the bounded feature magnitudes, adapter-factor norms, and frozen downstream Jacobians in the analyzed neighborhood.*

**Assumption 5** (Lipschitz bounded loss)**.** *The loss is* $L_\ell$*-Lipschitz in the prediction and bounded in* $[0, 1]$ *after normalization.*

## K Proof Details for Theorem 1

We cover the shared adapter parameter space rather than a single data-independent residual matrix. For ranks $\mathbf{r}$, the concatenated trainable adapter vector has dimension

$$p_{\text{eff}}(\mathbf{r}) = \sum_{l=1}^{L} \left(2C_l r_l + C_l + r_l\right),$$

including both $1 \times 1$ weights and biases. By Assumption 3, $\Phi$ lies in an Euclidean ball of radius $S$ in $\mathbb{R}^{p_{\text{eff}}}$. Its covering number therefore satisfies, up to universal constants,

$$\log N(\epsilon, \Phi, \|\cdot\|_2) \lesssim p_{\text{eff}}(\mathbf{r}) \log\left(\frac{cS}{\epsilon}\right). \tag{10}$$

Assumption 4 maps this parameter cover to an output-space cover with radius multiplied by at most $\Lambda$. Applying Dudley's entropy integral to the resulting local function class gives empirical Rademacher complexity

$$\widetilde{\mathcal{O}}\left(\Lambda S\sqrt{\frac{p_{\text{eff}}(\mathbf{r})}{n}}\right). \tag{11}$$

Lipschitz contraction by $L_\ell$ and the standard Rademacher generalization inequality for bounded losses then yield Eq. (5). Because $p_{\text{eff}}(\mathbf{r})$ is linear in the ranks to leading order when $r_l \ll C_l$, the bound retains the $\widetilde{\mathcal{O}}(\sqrt{r/n})$ dependence used in the sample-size-dependent capacity hypothesis. The result remains local and assumption-dependent; it does not describe gate changes or the complete nonlinear optimization trajectory.

## L RESULTS EXTENSION

**Qualitative few-shot progression (appendix only).** We keep this visualization outside the nine-page main paper because Table 3 already provides the quantitative few-shot progression and the main qualitative figure directly compares PathoField with Full-FT. Figure 5 provides an exploratory visualization from the fixed-split few-shot runs, showing how PathoField predictions evolve as the number of pathological training subjects increases from $k = 1$ to $k = 16$. For the same held-out Penn subjects, the overall anatomical structure and 7T-like contrast remain qualitatively stable across the tested few-shot regimes, while increasing $k$ can modify local contrast and intensity patterns. Individual visual changes need not be monotonic. These examples are descriptive and are not intended to establish lesion preservation or diagnostic validity.

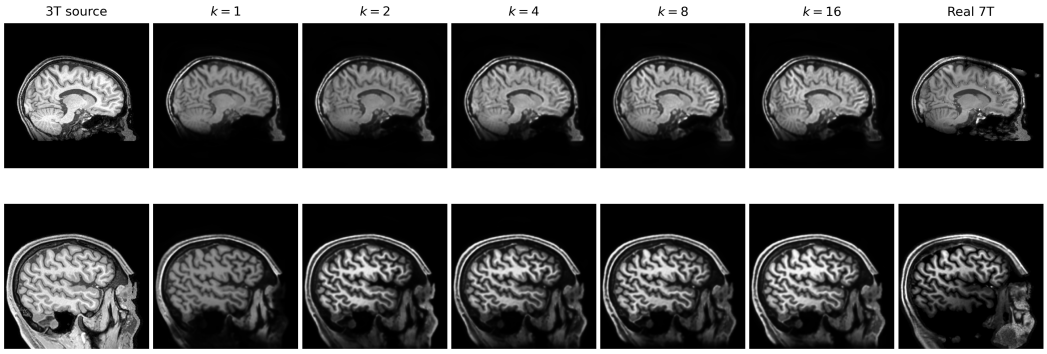



Figure 5: **Qualitative progression across few-shot adaptation sizes.** Representative pathological-domain synthesis results obtained with PathoField using $k \in \{1, 2, 4, 8, 16\}$ training subjects. As the number of pathological adaptation subjects increases, the synthesized 7T images progressively reflect the target-domain appearance while retaining the structural content of the corresponding 3T input. The examples illustrate the qualitative effect of adaptation capacity across few-shot regimes; they are intended as representative visual comparisons rather than evidence of monotonic improvement with $k$.