MRI-CARE: COMPOSE, ASSESS, AND REFINE WITH EXPERTS FOR 3D MRI RESTORATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Mixed-artifact MRI restoration requires effective correction without compromising anatomy. Specialized experts offer useful priors, but incomplete correction leaves residual artifacts and mismatched inputs for subsequent experts. We propose MRI-CARE, Compose, Assess, and Refine with Experts, for multi-artifact 3D MRI restoration. An agent combines adaptive expert selection and dose control with verification and learned stopping. Its policy is trained through supervised initialization and trajectory-level reinforcement learning. We further introduce Ruler-guided Residual Refinement (R^3): a 3D Ruler compares the endpoint with acceptable MRI exemplars, and a class-agnostic residual map guides a single refinement when needed. HCP experiments show improved restoration fidelity and cortical reconstruction, with R^3 providing up to 1.09 dB over expert composition. External evaluation on IXI and ABIDE examines transfer to simulated and real artifacts. Code will be released upon acceptance.

1 INTRODUCTION

Magnetic resonance imaging (MRI) is widely used to characterize brain anatomy, yet patient motion, noise, and intensity inhomogeneity can compromise image quality. These artifacts affect more than visual appearance: they obscure tissue boundaries and bias anatomical measurements, even in scans that pass conventional quality control (Andre et al., 2015; Reuter et al., 2015; Elyounssi et al., 2025). Multiple artifacts often coexist, and their composition is unknown at inference. Restoring such volumes therefore requires both effective artifact removal and preservation of anatomical detail.

Existing restoration methods broadly follow two paradigms (Fig. 1a). Task-specific methods specialize in a particular degradation, providing useful priors for bias correction, denoising, or motion correction (Tustison et al., 2010; Maggioni et al., 2013; Adame-Gonzalez et al., 2024; Singh et al., 2024). However, their selection and execution order must be determined for each input. All-in-one methods instead learn a unified mapping across degradations (Potlapalli et al., 2023; Conde et al., 2024; Tian et al., 2025; Yang et al., 2025). Covering MRI artifact types, severities, and their mixtures with one model is difficult, particularly when paired GT images are scarce.

Agentic restoration offers an alternative by composing specialized models according to the current image (Chen et al., 2024; Zhu et al., 2025; Jiang et al., 2025). Its central challenge is *reliable composition under incomplete correction*. An expert can leave residual artifacts, so its output differs from the remaining-artifact distribution expected by the next expert. For example, after motion correction, a volume containing motion and noise need not become a noise-only volume. The same denoiser can consequently behave differently on the actual intermediate volume and its simulated noise-only counterpart (Fig. 1a). In MRI, this mismatch matters because further correction must distinguish residual artifacts from fine anatomical structures. A sequence of individually useful experts does not by itself establish an acceptable endpoint.

We address this challenge with MRI-CARE: *Compose, Assess, and Refine with Experts*. MRI-CARE first composes a fixed model zoo of artifact-selective restorers (Fig. 1b). A learned agent selects the next expert from the current volume and restoration history, while a pairwise verifier checks each candidate before updating the image. A separate STOP policy determines when to end expert composition. Supervised fine-tuning initializes these decisions, followed by trajectory-level reinforcement learning.

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.

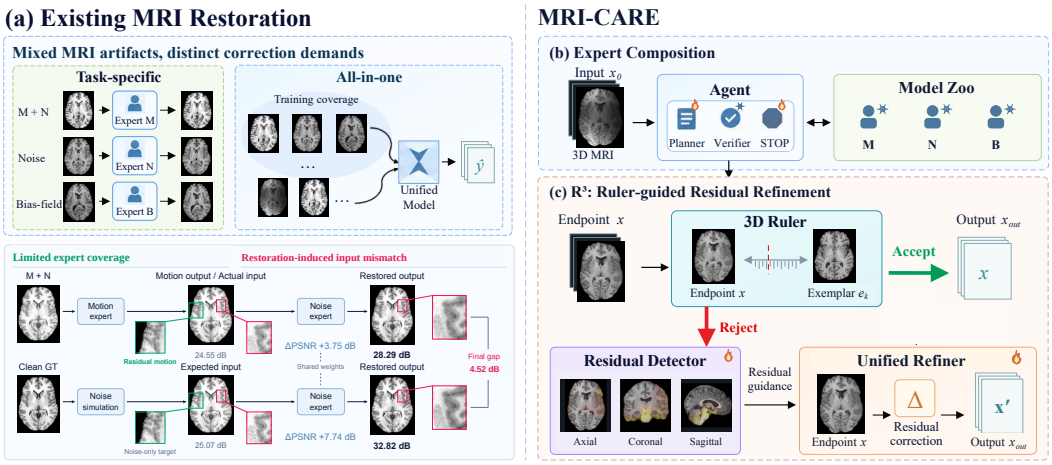


Figure 1: **Motivation and overview of MRI-CARE.** (a) Residual artifacts and input mismatch limit sequential restoration. (b) The agent selects, verifies, and stops expert execution. (c) R^3 assesses endpoint x and applies one residual-guided refinement when needed, returning x or $x\prime$.

We further introduce Ruler-guided Residual Refinement (R^3) to address artifacts remaining after expert composition (Fig. 1c). The 3D Ruler compares the endpoint with independent acceptable MRI exemplars. If the endpoint is rejected, a class-agnostic Residual Detector localizes remaining reconstruction discrepancies, and a unified refiner applies a single terminal correction. This design separates stepwise improvement from endpoint assessment and directs refinement to residual errors.

Our contributions are threefold:

- We introduce MRI-CARE, an agentic framework for multi-artifact 3D MRI restoration that combines adaptive expert selection, transition verification, and learned stopping.
- We propose R^3, which couples 3D acceptability assessment with residual detection and guided refinement to correct endpoint discrepancies left by incomplete expert correction and intermediate input mismatch.
- We evaluate external generalization on IXI (IXI Project, n.d.) across artifact severities and on ABIDE (Di Martino et al., 2014) with real mixed artifacts, and assess downstream cortical reconstruction to quantify benefits beyond image restoration quality.

2 Method

Given an input volume x_0, MRI-CARE performs expert composition followed by Ruler-guided Residual Refinement (R^3; Fig. 2). Expert composition updates the current volume x_t and returns an endpoint x. The 3D Ruler scores x and triggers a single residual-guided refinement to produce x' when the score is below the acceptance threshold. The final output x_{out} is x or x', respectively.

2.1 Expert Composition

To adapt restoration to each input, the agent composes specialized experts through sequential selection and verification (Fig. 2a). Starting from x_0, the Planner selects an expert from the model zoo, the Verifier accepts or rejects its correction, and the STOP policy determines when to return endpoint x.

Model zoo. The model zoo provides specialized correction for bias-field, motion, and noise artifacts under a shared anatomical prior. It contains three experts, $\mathcal{E} = \{E_B, E_M, E_N\}$. Given a volume v, a frozen 3D MRI backbone (Qiu et al., 2026) produces a reconstruction $\mathcal{B}(v)$ that guides an artifact-specific restoration branch G_a:

$$E_a(v) = G_a\big(v, \mathcal{B}(v)\big), \qquad a \in \{B, M, N\}. \tag{1}$$

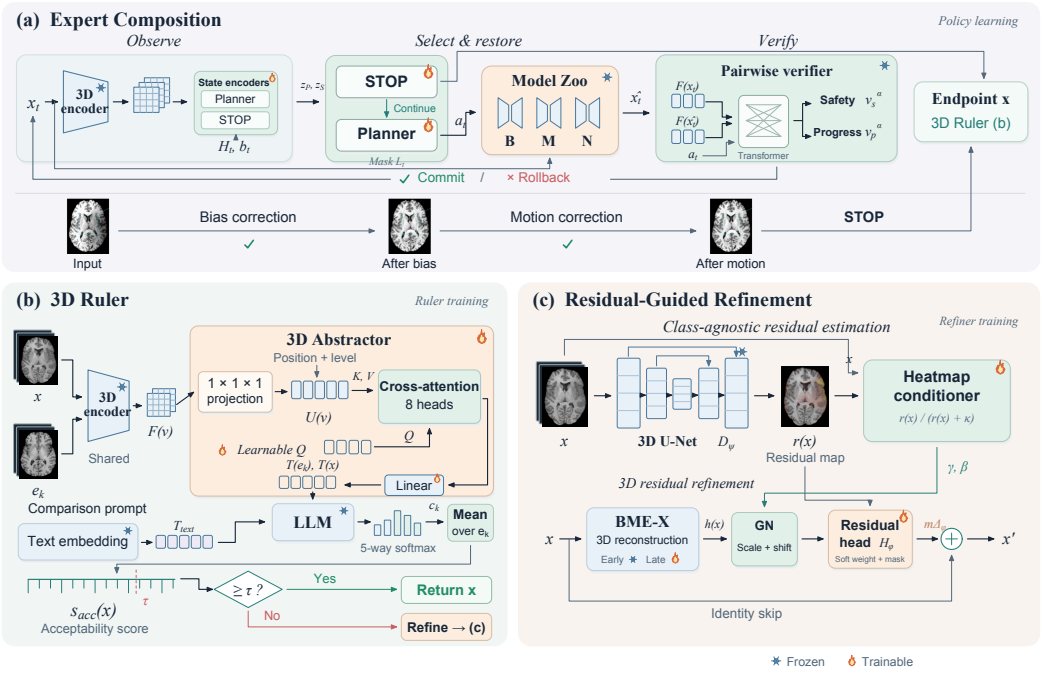


Figure 2: **MRI-CARE architecture.** (a) Expert composition with pairwise verification and learned stopping. (b) Exemplar-based 3D acceptability assessment. (c) Class-agnostic residual detection and BME-X refinement. Snowflakes/flames indicate frozen/trainable parameters during policy, Ruler, and refiner training, respectively; all parameters are frozen at inference.

Each branch uses the observed volume and structural prior to correct its designated artifact. Training targets retain the other corruptions, and target-absent examples use the input itself as the target to encourage identity. The backbone is frozen during expert training, and the entire model zoo is frozen during agent training.

Planner and adaptive stopping. The agent determines both which expert to apply and how long to restore each input (Fig. 2a). A learned budget allocator reads the initial image descriptor and predicts an input-dependent call budget b_0. At step t, the state $s_t = (x_t, H_t, b_t, L_t)$ consists of the current volume, accepted-action history, remaining budget, and available experts. The Planner and STOP policy each combine pooled MRI features $F(x_t)$, a history GRU, and a budget embedding through an independent state encoder. The Planner predicts an expert distribution q_t, while the STOP policy predicts the probability p_t of ending restoration:

$$\pi(\textsc{Stop} \mid s_t) = p_t, \qquad \pi(a \mid s_t) = (1 - p_t)q_t(a), \quad a \in L_t. \tag{2}$$

During inference, the STOP policy returns the current volume as endpoint x when $p_t \geq 0.5$. Otherwise, the Planner selects the highest-scoring available expert. The agent re-evaluates this decision after each attempted correction, allowing restoration to end before the allocated budget is used. Previously attempted experts are excluded from subsequent selection. Budget prediction and its supervision are detailed in Appendix A.

Expert execution and verification. To limit overcorrection, we control the correction strength of every selected expert $a_t \in \{B, M, N\}$. The expert is evaluated once, and its output is blended with the current volume to form three candidates:

$$\widehat{x}_t^{\alpha} = x_t + \alpha\big(E_{a_t}(x_t) - x_t\big), \quad \alpha \in \{0.5, 0.75, 1\}. \tag{3}$$

For each candidate $\widehat{x}_t^{\alpha}$, the Verifier receives the current-volume features $F(x_t)$, candidate features $F(\widehat{x}_t^{\alpha})$, and selected action a_t (Fig. 2a):

$$(v_s^{\alpha}, v_p^{\alpha}) = \mathcal{V}\big(F(x_t), F(\widehat{x}_t^{\alpha}), a_t\big). \tag{4}$$

Aligned feature comparisons form pair tokens, which are combined with the action embedding and aggregated by a Transformer. Two sigmoid heads output the safety probability v_s^α and progress probability v_p^α. Candidates with $v_s^\alpha < \tau_s$ are rejected. Among the remaining candidates, we select the one with the highest v_p^α, choosing the weaker correction when scores differ by at most ϵ_p. Writing the selected candidate as $\widehat{x}_t$, the Commit/Rollback update is

$$x_{t+1} = \begin{cases} \widehat{x}_t, & \text{a candidate is accepted,} \\ x_t, & \text{all candidates are rejected.} \end{cases} \tag{5}$$

Accepted actions are added to H_t, and each attempted correction consumes one call from b_t. The agent then observes the updated volume x_{t+1} to select the next action or stop.

Verifier training. We train the Verifier to assess anatomical preservation and restoration improvement without a clean reference at inference. Training uses expert-generated before–after pairs with a shared clean reference, which supplies safety labels from anatomical segmentation agreement and progress labels from restoration gain. Binary cross-entropy is applied to each head's labeled pairs, excluding uncertain labels. The feature-comparison modules and prediction heads are optimized with the MRI encoder frozen. The trained Verifier is then frozen during policy learning (Fig. 2a).

Policy training. The policy is trained to improve restoration quality while avoiding unnecessary expert calls. We use supervised fine-tuning (SFT) for initialization, followed by trajectory-level reinforcement learning (RL):

$$\begin{aligned} \text{SFT:} \quad & \mathcal{L}_{\text{SFT}} = \mathcal{L}_{\text{expert}} + \mathcal{L}_{\text{stop}}, \\ \text{RL:} \quad & \mathcal{L}_{\text{RL}} = \mathcal{L}_{\text{GRPO}} + \lambda_{\text{pref}} \mathcal{L}_{\text{pref}} + \beta D_{\text{KL}}(\pi_\theta \| \pi_{\text{SFT}}). \end{aligned} \tag{6}$$

Here, $\mathcal{L}_{\text{expert}}$ and $\mathcal{L}_{\text{stop}}$ supervise expert selection and stopping using cross-entropy and binary cross-entropy. $\mathcal{L}_{\text{GRPO}}$ favors trajectories with higher relative rewards, $\mathcal{L}_{\text{pref}}$ favors the better trajectory in each preference pair, and the KL term limits deviation from the SFT policy. The weights λ_{pref} and β balance the latter two terms. A trajectory with n expert calls receives

$$R = \begin{cases} \Delta q - cn, & \text{anatomy is preserved,} \\ \min(\Delta q, 0) - \lambda_{\text{anat}} - cn, & \text{otherwise,} \end{cases} \tag{7}$$

where Δq is the reduction in normalized log spectral error relative to the input, c penalizes calls, and λ_{anat} penalizes anatomical degradation. Experts, the MRI encoder, and Verifier remain frozen. Appendix A provides the targets, budget supervision, reward measurements, and rollout and optimization settings.

2.2 3D RULER

The 3D Ruler assesses endpoint x against a fixed bank of acceptable MRI volumes, $\mathcal{A} = \{e_k\}_{k=1}^{K_e}$ (Fig. 2b). Here, e_k is the k-th exemplar and K_e is the number of exemplars. Exemplars are independent quality references, not paired GT.

Volumetric abstraction. The abstractor converts each MRI volume into a compact sequence of visual tokens for the comparison decoder. For $v \in \{x, e_k\}$, the shared frozen encoder extracts a feature pyramid $F(v)$. Level-specific $1\times1\times1$ projections, normalized 3D-position embeddings, and level embeddings form the spatial memory $U(v)$. A shared learnable query bank Q aggregates this memory through multi-head cross-attention:

$$T_\theta(v) = W_o \,\text{MHA}\big(Q, U(v), U(v)\big). \tag{8}$$

Support masking excludes background positions. The output projection W_o maps the attended queries to the LLM embedding width. The same abstractor processes the endpoint and each exemplar independently, without cross-subject spatial correspondence.

Comparison and supervision. We estimate acceptability by aggregating quality comparisons with the exemplars. The decoder's tokenizer and frozen token-embedding layer convert a fixed MRI comparison prompt into text tokens T_{text}. These tokens identify the exemplar and endpoint

and instruct the decoder to compare their quality. The visual tokens $T_\theta(e_k)$ and $T_\theta(x)$ replace the prompt's two image placeholders, forming a joint text–visual input. The pretrained comparison decoder (Zhu et al., 2024) returns five response logits: inferior, worse, similar, better, and superior. A softmax over these responses gives $\pi_{k,j}$, the probability of response j for exemplar e_k and endpoint x. The pairwise acceptability score $c_k(x)$ and aggregate score $s_{\mathrm{acc}}(x)$ are

$$c_k(x) = \sum_{j\in\{\mathrm{similar,better,superior}\}} \pi_{k,j}(x), \qquad s_{\mathrm{acc}}(x) = \frac{1}{K_e}\sum_{k=1}^{K_e} c_k(x). \tag{9}$$

Each exemplar–query pair is supervised by the query's human acceptability label through binary cross-entropy on $c_k(x)$. Only the abstractor, including Q, is trained. The encoder, text embeddings, and decoder remain frozen. The resulting acceptability score determines whether to retain endpoint x or apply residual-guided refinement. The exemplar bank and threshold τ are fixed before evaluation.

2.3 Residual-Guided Refinement

This stage corrects errors left by expert composition through residual localization and a single heatmap-guided update to x' (Fig. 2c), without reassessment.

Residual estimation. The detector estimates local error magnitude from endpoint x using a 3D U-Net D_ψ. Its nonnegative map $r(x) = D_\psi(x)$ is supervised by local RMS error between volume v and clean reference $x^\star$:

$$r_v^\star(p) = \sqrt{\frac{\sum_{q\in\mathcal{N}(p)} m(q)[v(q) - x^\star(q)]^2}{\sum_{q\in\mathcal{N}(p)} m(q)}}, \quad m(p) = 1, \tag{10}$$

where p, q index voxels, $\mathcal{N}(p)$ is a 7^3 neighborhood, and m is the binary brain-support mask. Paired pretraining supervises residual magnitude, correction-induced changes, and zero residuals for clean volumes, followed by adaptation to expert outputs.

Heatmap-conditioned 3D refinement. The refiner concentrates correction on regions with high predicted residuals. BME-X (Sun et al., 2025) supplies full-resolution 3D reconstruction features $h(x)$. The heatmap conditioner uses x and normalized residuals $\bar{r} = r(x)/(r(x) + \kappa)$ to predict scale and shift fields γ, β. The residual head H_ϕ and soft weighting produce:

$$\begin{aligned} \widetilde{h} &= (1+\gamma) \odot \mathrm{GN}(h(x)) + \beta, \\ \Delta_\phi &= (\rho + (1-\rho)\bar{r}) \odot \tanh\big(H_\phi(\widetilde{h})\big), \\ x' &= x + m \odot \Delta_\phi. \end{aligned} \tag{11}$$

Here, GN denotes group normalization, $\odot$ is elementwise multiplication, and ϕ denotes the refiner's trainable parameters. The constants $\kappa > 0$ and $0 < \rho < 1$ control map normalization and minimum correction weight. Soft weighting scales the correction by $\rho + (1-\rho)\bar{r}$. The identity skip adds the masked correction $m \odot \Delta_\phi$ to x, preserving unsupported voxels. Zero initialization gives an identity mapping. Averaging overlapping 3D patch predictions yields x'.

Refiner training and output. The refiner learns to reduce residual errors while preserving low-residual regions. Intensity, three-axis gradient, and low-residual identity losses supervise voxel fidelity, boundary detail, and preservation, respectively. With the detector frozen, training updates only late BME-X reconstruction layers, the conditioner, and residual head. BatchNorm statistics remain fixed. The Ruler selects the final output:

$$x_{\mathrm{out}} = \begin{cases} x, & s_{\mathrm{acc}}(x) \geq \tau, \\ x', & s_{\mathrm{acc}}(x) < \tau, \end{cases} \tag{12}$$

where τ is the acceptance threshold. Appendix B details the architecture, losses, and training settings.

Table 1: **HCP restoration and cortical reconstruction.** A/B/C: one/two/three artifact types (18/54/54 volumes). PSNR is in dB; SSIM is 3D. SHN (surface holes) and CT-MAE (thickness error, mm) use nine matched volumes, three per group. †: GT artifact labels guide prompting or routing.

	Group A			Group B			Group C			Cortical (subset)	
Method	PSNR↑	SSIM↑	NRMSE↓	PSNR↑	SSIM↑	NRMSE↓	PSNR↑	SSIM↑	NRMSE↓	SHN↓	CT-MAE↓
General all-in-one restoration											
PromptIR	30.16	0.8390	0.2668	29.51	0.7990	0.2952	28.30	0.7679	0.3168	180.3	0.3056
InstructIR	31.09	0.8616	0.2413	30.04	0.8086	0.2777	28.38	0.7586	0.3166	154.2	0.2571
DFPIR†	34.25	0.9085	0.1752	31.74	0.8591	0.2334	29.27	0.7844	0.2846	128.7	0.2053
Task-specific restoration											
Restormer-ST ×3 (oracle)†	35.87	0.9332	0.1460	30.80	0.7809	0.2744	26.84	0.6194	0.3867	109.8	0.1859
Restormer-ST ×3 (fixed)	29.63	0.8102	0.2882	29.43	0.7379	0.3146	26.84	0.6194	0.3867	220.7	0.3441
MRI-specific restoration											
TAT-MRI (retrained)	34.39	0.9269	0.1692	32.69	0.8903	0.2062	31.17	0.8703	0.2291	68.7	0.1283
MRI-CARE	**36.46**	**0.9425**	**0.1355**	**33.52**	**0.9090**	**0.1841**	**31.75**	**0.8862**	**0.2118**	**26.0**	**0.0877**

3 EXPERIMENTS

3.1 EXPERIMENTAL SETUP

Datasets. We construct a counterfactual artifact bank from HCP MRI (Van Essen et al., 2013), with subject-disjoint training, model-selection, and validation splits. The bank covers bias field (B), motion (M), noise (N), their mixtures, and clean controls. We simulate smooth multiplicative intensity fields, motion-corrupted k-space acquisitions, and stochastic noise at multiple strengths; mixtures follow a fixed corruption order. For each corrupted input, a selective target omits the designated artifact while preserving the parameters, random realizations, and order of the remaining corruptions. For example, $BMN \rightarrow MN$ supervises bias correction without requiring removal of motion or noise. Target-absent pairs supervise identity. These partial targets train selective experts, whereas the paired clean volume provides GT for final restoration evaluation. Simulation and selective-pair construction are detailed in Appendix E. The validation comparison uses 126 corrupted volumes from 49 subjects: Groups A, B, and C contain one, two, and three artifact types (18/54/54 inputs). External evaluation uses simulated artifacts on IXI (IXI Project, n.d.) and real artifacts on ABIDE (Di Martino et al., 2014).

Baselines and metrics. We compare all-in-one restorers PromptIR, InstructIR, and DFPIR (Potlapalli et al., 2023; Conde et al., 2024; Tian et al., 2025), task-specific Restormer cascades (Zamir et al., 2022), and TAT-MRI (Yang et al., 2025). Fixed-order and oracle-routed cascades are reported separately. All methods receive matched inputs; volumetric PSNR, 3D SSIM, and NRMSE use common foreground masks and paired ground truth (GT).

Implementation details. The policy is initialized by SFT and optimized with trajectory-level GRPO using eight rollouts per input and a learning rate of 10^{-5}. Experts, MRI features, and the Verifier remain frozen during policy optimization. All parameters are fixed at inference. Architectures, component training, and baseline settings appear in Appendices A, A, and F.

3.2 COMPARISON WITH RESTORATION METHODS

Comparison with all-in-one and expert-based methods. Table 1 compares MRI-CARE with unified restorers and expert cascades on matched HCP inputs. TAT-MRI exceeds the fixed Restormer cascade by 3.26/4.33 dB on Groups B/C. Oracle routing improves Restormer on Groups A/B but leaves Group C unchanged, highlighting the difficulty of correcting mixtures even with known artifact types.

Figure 3 compares intermediate expert outputs and final baseline results under a shared display window. The displayed trajectories use M, N→M, and B→M→N, illustrating input-dependent restoration rather than a common expert order. Matched enlargements and GT-error maps in Fig. 4 examine local residuals alongside tissue detail, complementing the volume-averaged fidelity metrics.

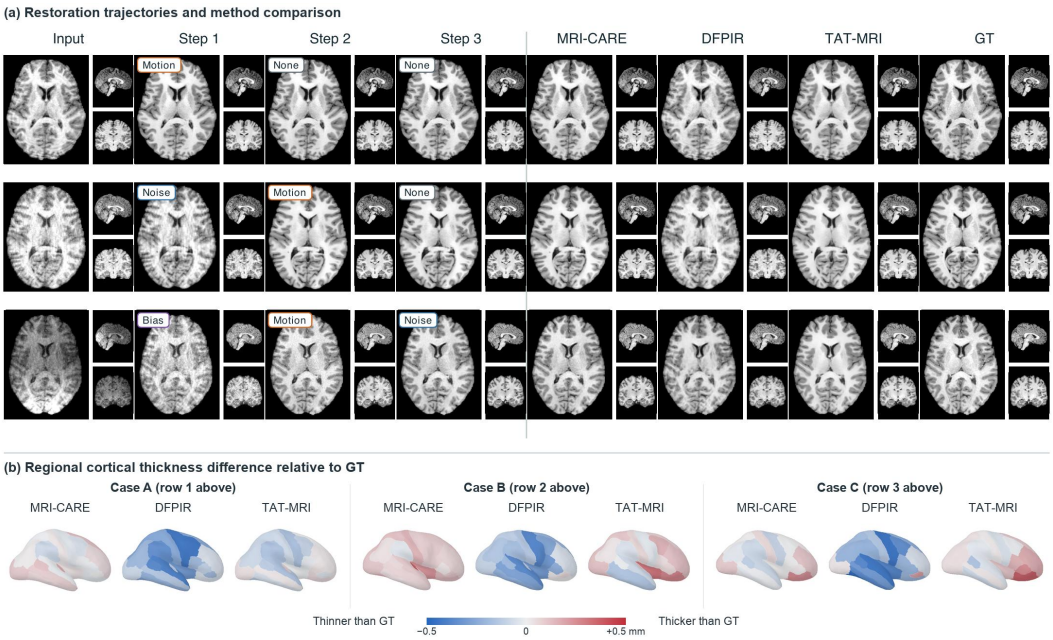


Figure 3: **Restoration trajectories and cortical fidelity.** (a) Intermediate outputs and method comparisons on three HCP validation cases; None denotes no further expert call. DFPIR receives artifact labels. (b) Regional cortical-thickness differences from GT on the corresponding left-hemisphere surfaces: blue/red indicates thinner/thicker estimates (± 0.5 mm).

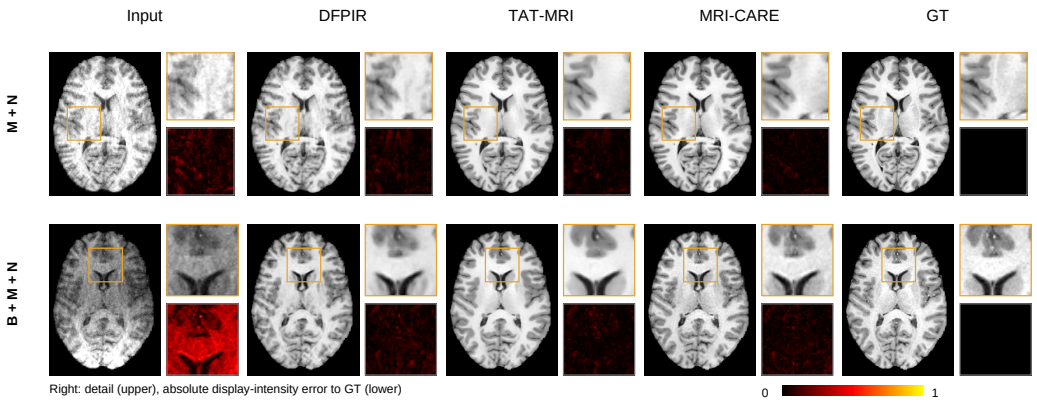


Figure 4: **Mixed-artifact restoration.** Motion–noise (top) and bias–motion–noise (bottom). Insets show local detail and absolute display-intensity error to GT, using a shared 0–1 scale.

Cortical reconstruction. We assess anatomical utility with a fixed FreeSurfer pipeline on nine HCP subjects, three per artifact group, shared across methods. Table 1 reports surface hole number (SHN) before topology correction and regional cortical-thickness error (CT-MAE) against the paired clean reconstruction (Elyounssi et al., 2025). Figure 3 includes corresponding thickness-error maps. SHN measures initial surface defects, while CT-MAE assesses the accuracy of the resulting anatomical measurements; neither is replaced by image-level PSNR. Completion and aggregation are defined in Appendix F.1.

Table 2: **Agent and refinement ablations.** A/B/C follow Table 1. Dose: full or adaptive correction strength.

Configuration						Group A		Group B		Group C	
Variant	Selection	Dose	Training	Verify	Stop	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑
SFT only	Adaptive	Full	SFT	✓	Learned	34.05	0.8974	30.10	0.8326	28.35	0.7992
Fixed order	Fixed	Full	SFT	✓	Learned	34.38	0.9044	28.92	0.7669	27.68	0.7056
Expert Dose Control	Adaptive	Adaptive	SFT	✓	Learned	34.92	0.9185	32.06	0.8786	30.87	0.8474
w/o verifier	Adaptive	Adaptive	SFT	×	Learned	34.92	0.9185	31.72	0.8726	30.45	0.8438
w/o learned stop	Adaptive	Adaptive	SFT	✓	Budget	34.76	0.9175	31.31	0.8572	29.47	0.8207
MRI-CARE (w/o R^3)	Adaptive	Adaptive	SFT+GRPO	✓	Learned	35.37	0.9272	32.64	0.8869	31.12	0.8636
MRI-CARE	Adaptive	Adaptive	SFT+GRPO	✓	Learned	**36.46**	**0.9425**	**33.52**	**0.9090**	**31.75**	**0.8862**

3.3 Ablation Study

Table 2 evaluates adaptive expert selection, dose control, verification, stopping, and policy learning. Comparing agent endpoints with the complete system further assesses the benefit of terminal refinement.

Effect of expert selection and dose control. Adaptive expert selection improves mixed-artifact restoration, while dose control adjusts correction strength to the current volume. With SFT and full-strength execution, adaptive selection gains 1.18/0.67 dB over fixed ordering on Groups B/C, despite a 0.33 dB decrease on Group A. Adding adaptive dose control further improves PSNR by 0.87/1.96/2.52 dB across Groups A/B/C. The larger gains on mixtures support adapting both expert choice and correction strength.

Effect of verification and stopping. Verification and learned stopping improve endpoint quality by controlling whether to accept a correction and continue restoration. Relative to Expert Dose Control, removing verification lowers PSNR by 0.34/0.42 dB on Groups B/C, with Group A unchanged. Replacing learned STOP with budget exhaustion reduces PSNR by 0.16/0.75/1.40 dB across Groups A/B/C. These results support checking proposed updates and stopping before further calls degrade the endpoint.

Effect of policy learning. Trajectory-level RL improves expert composition beyond SFT initialization. Compared with Expert Dose Control, MRI-CARE without R^3 adds GRPO while retaining adaptive selection, dose control, verification, and stopping. PSNR increases by 0.45/0.58/0.25 dB and SSIM by 0.0087/0.0083/0.0162 across Groups A/B/C, supporting optimization of final trajectory quality.

Effect of residual refinement. Terminal refinement improves restoration beyond the agent endpoint. Adding R^3 increases PSNR by 1.09/0.88/0.63 dB and SSIM by 0.0153/0.0221/0.0226 across Groups A/B/C. Figure 5 compares predicted residual maps with GT-derived errors and refined outputs, complementing these gains with spatial evidence.

3.4 External Generalization

We evaluate transfer to simulated IXI artifacts, real ABIDE artifacts, and unseen Ghosting. All models and thresholds remain fixed without external fine-tuning; IXI supports paired-reference evaluation, whereas ABIDE requires no-reference assessment.

Cross-dataset restoration. MRI-CARE records PSNR/3D SSIM of 31.05/0.8473, 26.10/0.7335, and 23.54/0.6391 across IXI Groups A/B/C, respectively (Table 3). It also improves single-artifact restoration across bias, motion, and noise (Table 5), achieving 28.88/32.40/31.87 dB and exceeding DFPIR by 2.55/3.85/0.85 dB, with corresponding SSIM gains of 0.0243/0.0218/0.0223. On real-artifact ABIDE, MRI-CARE achieves 38.82/0.3361/0.5680 in MUSIQ/MANIQA/CLIP-IQA, the highest values among the compared methods (Table 4) (Ke et al., 2021; Yang et al., 2022; Wang et al., 2023).

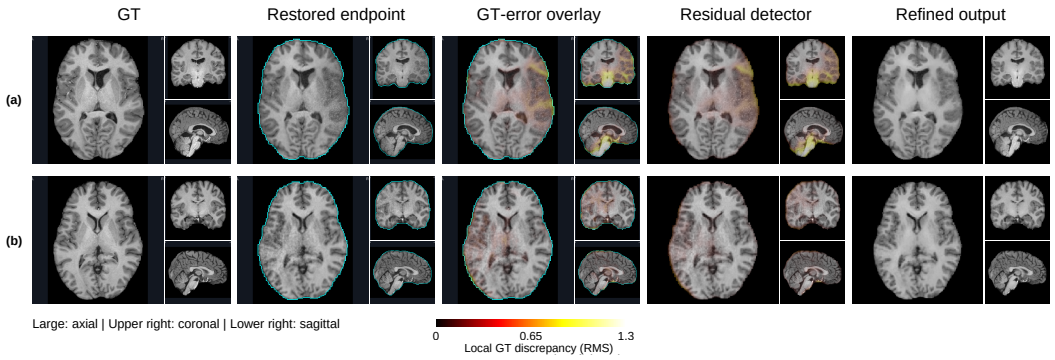


Figure 5: **Residual detection and refinement on two HCP validation cases.** GT-derived local RMS errors and predicted residual maps share the same scale. Axial views include coronal and sagittal insets.

Table 3: **IXI restoration.** PSNR/3D SSIM; A/B/C: one/two/three artifact types. O/F: oracle/fixed cascade. †: GT artifact labels guide prompting or routing.

Method	Group A	Group B	Group C
PromptIR	24.69/0.7067	23.77/0.6670	22.90/0.6257
InstructIR	24.04/0.6801	22.83/0.6193	22.01/0.5773
DFPIR†	28.63/0.8245	25.50/0.7346	24.13/0.6778
Restormer-ST (O)†	28.67/0.8285	24.52/0.6743	22.42/0.5809
Restormer-ST (F)	24.35/0.6747	23.40/0.6361	22.42/0.5809
TAT-MRI	24.24/0.6864	23.50/0.6584	22.91/0.6321
MRI-CARE	**31.05/0.8473**	**26.10/0.7335**	**23.54/0.6391**

Table 4: **ABIDE restoration.** No-reference quality scores (↑). O/F: oracle/fixed cascade. †: frozen artifact labels guide prompting or routing.

Method	MUSIQ↑	MANIQA↑	CLIP-IQA↑
PromptIR	37.25	0.2789	0.5258
InstructIR	38.23	0.2997	0.5437
DFPIR†	37.43	0.2921	0.5386
Restormer-ST (O)†	38.47	0.2930	0.5458
Restormer-ST (F)	38.29	0.2976	0.5411
TAT-MRI	38.09	0.3146	0.5442
MRI-CARE	**38.82**	**0.3361**	**0.5680**

Table 5: **IXI single-artifact restoration.** PSNR/3D SSIM over 47 subjects. B/M/N: bias/motion/noise. †: artifact-aware prompting.

Method	B	M	N
PromptIR	24.10/0.6791	24.69/0.6998	25.27/0.7412
InstructIR	23.67/0.6606	24.23/0.6819	24.22/0.6978
DFPIR†	26.33/0.7400	28.55/0.8559	31.02/0.8776
Restormer-ST (fixed)	24.42/0.6701	24.16/0.6578	24.49/0.6961
TAT-MRI	24.62/0.6961	23.92/0.6740	24.18/0.6891
MRI-CARE	**28.88/0.7643**	**32.40/0.8777**	**31.87/0.8999**

Table 6: **Zero-shot Ghosting restoration.** DFPIR uses a null prompt.

Method	PSNR↑	SSIM↑	NRMSE↓
PromptIR	29.45	0.7800	0.2870
InstructIR	30.38	0.7864	0.2707
DFPIR	34.64	0.7378	0.2564
Restormer-ST (F)	28.34	0.7813	0.3204
TAT-MRI	33.00	0.8530	0.2310
MRI-CARE	**34.70**	**0.8722**	**0.2285**

Unseen-artifact restoration. MRI-CARE also generalizes to Ghosting without a corresponding expert action (Table 6). Its PSNR is close to DFPIR (34.70 versus 34.64 dB), while SSIM increases from 0.7378 to 0.8722, indicating better structural fidelity under this unseen artifact.

4 Conclusion

We presented MRI-CARE for 3D MRI restoration under incomplete expert correction and input mismatch. Adaptive expert composition integrates dose control, verification, and learned stopping. Ruler-guided Residual Refinement couples endpoint assessment with a single residual-guided correction. HCP experiments show improved restoration fidelity and cortical reconstruction; IXI, ABIDE, and unseen-artifact evaluations examine external generalization.

Reproducibility Statement

The appendix specifies the state and action representation, reward computation, policy supervision, endpoint assessment, and evaluation protocols. Comparisons use matched input/GT pairs, fixed artifact realizations, and subject-level statistical units.

Ethics Statement

MRI restoration can remove or alter anatomical structures alongside artifacts. Evaluation therefore considers anatomical fidelity and residual artifacts in addition to image quality. Data use follows the access conditions of each dataset; clinical deployment requires independent validation.

Use of AI Tools

We used AI tools for literature review, research design, manuscript drafting and editing. The authors take full responsibility for the paper's content.

References

Walter Adame-Gonzalez, Aliza Brzezinski-Rittner, Yashar Zeighami, M. Mallar Chakravarty, Reza Farivar, and Mahsa Dadar. FONDUE: Robust resolution-invariant denoising of MR images using nested UNets. *Imaging Neuroscience*, 2:imag–2–00374, 2024. doi: 10.1162/imag_a_00374.

Jalal B. Andre, Brian W. Bresnahan, Mahmud Mossa-Basha, Michael N. Hoff, C. Patrick Smith, Yoshimi Anzai, and Wendy A. Cohen. Toward quantifying the prevalence, severity, and cost associated with patient motion during clinical MR examinations. *Journal of the American College of Radiology*, 12(7):689–695, 2015. doi: 10.1016/j.jacr.2015.03.007.

Haoyu Chen, Wenbo Li, Jinjin Gu, Jingjing Ren, Sixiang Chen, Tian Ye, Renjing Pei, Kaiwen Zhou, Fenglong Song, and Lei Zhu. RestoreAgent: Autonomous image restoration agent via multimodal large language models. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-3512.

I-Hsiang Chen, Isma Hadji, Enrique Sanchez, Adrian Bulat, Sy-Yen Kuo, Radu Timofte, Georgios Tzimiropoulos, and Brais Martinez. Restore, assess, repeat: A unified framework for iterative image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 15506–15515, 2026.

Marcos V. Conde, Gregor Geigle, and Radu Timofte. InstructIR: High-quality image restoration following human instructions. In *European Conference on Computer Vision*, 2024.

Adriana Di Martino, Chao-Gan Yan, Qingyang Li, Erin Denio, F. Xavier Castellanos, Kaat Alaerts, Jeffrey S. Anderson, Michal Assaf, Susan Y. Bookheimer, Mirella Dapretto, et al. The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry*, 19(6):659–667, 2014. doi: 10.1038/mp.2013.78.

Safia Elyounssi, Keiko Kunitoki, Jacqueline A. Clauss, et al. Addressing artifactual bias in large, automated MRI analyses of brain development. *Nature Neuroscience*, 28:1787–1796, 2025. doi: 10.1038/s41593-025-01990-7.

Hákon Gudbjartsson and Samuel Patz. The rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, 34(6):910–914, 1995. doi: 10.1002/mrm.1910340618.

IXI Project. IXI Dataset. `https://brain-development.org/ixi-dataset/`, n.d. Accessed September 26, 2026.

Xu Jiang, Gehui Li, Bin Chen, and Jian Zhang. Multi-agent image restoration. *arXiv preprint arXiv:2503.09403*, 2025.

Yitong Jiang, Zhaoyang Zhang, Tianfan Xue, and Jinwei Gu. AutoDIR: Automatic all-in-one image restoration with latent diffusion. In *European Conference on Computer Vision*, 2024.

Junjie Ke, Qifei Wang, Yilin Wang, Peyman Milanfar, and Feng Yang. Musiq: Multi-scale image quality transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 5148–5157, 2021.

Matteo Maggioni, Vladimir Katkovnik, Karen Egiazarian, and Alessandro Foi. Nonlocal transform-domain filter for volumetric data denoising and reconstruction. *IEEE Transactions on Image Processing*, 22(1):119–133, 2013. doi: 10.1109/TIP.2012.2210725.

Vaishnav Potlapalli, Syed Waqas Zamir, Salman Khan, and Fahad Shahbaz Khan. PromptIR: Prompting for all-in-one blind image restoration. In *Advances in Neural Information Processing Systems*, 2023.

Zelin Qiu, Xi Wang, Zhuoyao Xie, et al. Large-scale multi-sequence pretraining for generalizable MRI analysis in versatile clinical applications. *Nature Biomedical Engineering*, 2026. doi: 10.1038/s41551-026-01740-5.

Martin Reuter, M. Dylan Tisdall, Abid Qureshi, Randy L. Buckner, André J. W. van der Kouwe, and Bruce Fischl. Head motion during MRI acquisition reduces gray matter volume and thickness estimates. *NeuroImage*, 107:107–115, 2015. doi: 10.1016/j.neuroimage.2014.12.006.

Shihao Shan, Hongying Liu, Fanhua Shang, Liang Wan, and Jingjing Deng. Beyond the static-world: Lifelong learning for all-in-one medical image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 13702–13711, 2026. URL https://openaccess.thecvf.com/content/CVPR2026/papers/Shan_Beyond_the_Static-World_Lifelong_Learning_for_All-in-One_Medical_Image_Restoration_CVPR_2026_paper.pdf.

Nalini M. Singh, Neel Dey, Malte Hoffmann, Bruce Fischl, Elfar Adalsteinsson, Robert Frost, Adrian V. Dalca, and Polina Golland. Data consistent deep rigid MRI motion correction. In *Medical Imaging with Deep Learning*, volume 227, pp. 368–381, 2024. URL https://proceedings.mlr.press/v227/singh24a.html.

Yue Sun, Limei Wang, Gang Li, Weili Lin, and Li Wang. A foundation model for enhancing magnetic resonance images and downstream segmentation, registration and diagnostic tasks. *Nature Biomedical Engineering*, 9:521–538, 2025. doi: 10.1038/s41551-024-01283-7.

Xiangpeng Tian, Xiangyu Liao, Xiao Liu, Meng Li, and Chao Ren. Degradation-aware feature perturbation for all-in-one image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 28165–28175, 2025.

Nicholas J. Tustison, Brian B. Avants, Philip A. Cook, Yuanjie Zheng, Alexander Egan, Paul A. Yushkevich, and James C. Gee. N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6):1310–1320, 2010. doi: 10.1109/TMI.2010.2046908.

David C. Van Essen, Stephen M. Smith, Deanna M. Barch, Timothy E. J. Behrens, Essa Yacoub, Kamil Ugurbil, and WU-Minn HCP Consortium. The WU-Minn human connectome project: An overview. *NeuroImage*, 80:62–79, 2013. doi: 10.1016/j.neuroimage.2013.05.041.

Jianyi Wang, Kelvin C. K. Chan, and Chen Change Loy. Exploring clip for assessing the look and feel of images. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 2555–2563, 2023. doi: 10.1609/aaai.v37i2.25353.

Sidi Yang, Tianhe Wu, Shuwei Shi, Shanshan Lao, Yuan Gong, Mingdeng Cao, Jiahao Wang, and Yujiu Yang. Maniqa: Multi-dimension attention network for no-reference image quality assessment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1191–1200, 2022.

Zhiwen Yang, Jiaju Zhang, Yang Yi, Jian Liang, Bingzheng Wei, and Yan Xu. TAT: Task-adaptive transformer for all-in-one medical image restoration. In *Medical Image Computing and Computer Assisted Intervention*, volume 15975, pp. 565–575, 2025. doi: 10.1007/978-3-032-05325-1_54.

Maxim Zaitsev, Julian Maclaren, and Michael Herbst. Motion artifacts in MRI: A complex problem with many partial solutions. *Journal of Magnetic Resonance Imaging*, 42(4):887–901, 2015. doi: 10.1002/jmri.24850.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5728–5739, 2022.

Hanwei Zhu, Haoning Wu, Yixuan Li, Zicheng Zhang, Baoliang Chen, Lingyu Zhu, Yuming Fang, Guangtao Zhai, Weisi Lin, and Shiqi Wang. Adaptive image quality assessment via teaching large multimodal model to compare. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-1024.

Kaiwen Zhu, Jinjin Gu, Zhiyuan You, Yu Qiao, and Chao Dong. An intelligent agentic system for complex image restoration problems. In *International Conference on Learning Representations*, 2025.

A Agent Architecture and Training

State representation and parameter sharing. The frozen MRI encoder supplies three spatial feature levels. Each policy branch forms a 256-dimensional image descriptor, a 64-dimensional GRU encoding of accepted actions, and a 32-dimensional encoding of the remaining-budget fraction. Planner and STOP branches have independent state, history, budget, and query encoders. The STOP branch uses a frozen copy of the planner's action embeddings for history encoding. Rejected actions remain in the execution log but are excluded from this history; their consequences enter through the reduced budget and legal-action mask. Agent feature pooling uses all valid volume positions, without a clean-reference brain mask. The Ruler uses its own shared exemplar–query abstractor over frozen encoder features.

Input-dependent call budget. A budget allocator maps the initial 256-dimensional image descriptor to an ordinal distribution over useful trajectory lengths $\ell \in \{0, \dots, K\}$, where $K = |\mathcal{E}|$. A layer-normalized two-layer MLP with hidden width 128 and GELU predicts a scalar burden score. Subtracting learned ordered thresholds and applying sigmoid yields $P(\ell \geq j \mid x_0)$. Differences between adjacent survival probabilities give the length distribution, whose calibrated upper quantile determines b_0. The allocator is supervised by the shortest anatomically safe trajectory whose endpoint utility lies within an equivalence margin of the best safe endpoint. Binary cross-entropy trains each ordinal target $\mathbb{1}[\ell^\star \geq j]$. The allocated budget remains fixed for that input, while the STOP policy may terminate earlier. Each attempted expert reduces the remaining budget, including rejected corrections. Exhausting the allocation terminates composition; the expert count bounds the allocation but does not prescribe a fixed trajectory length. Fixed-cap experiments and adaptive-budget experiments retain their respective runtime configurations.

Candidate selection. Each selected Bias, Motion, or Noise expert executes once and constructs three residual-scaled candidates using Eq. 3. The verifier produces safety probability v_s and progress probability v_p for each candidate, using $\tau_s = 0.05$ and $\epsilon_p = 0.02$. Candidates with $v_s < 0.05$ are excluded. Among surviving candidates, the smallest α whose progress score is within 0.02 of the maximum is selected. If no candidate survives, the parent is retained. The primary configuration does not apply a hard progress threshold; the hard-progress comparison is specified in Appendix F.3. All candidate checks belong to one attempted expert call.

Verifier architecture and targets. Shared level-specific projections compare aligned parent and candidate maps. Signed differences, absolute differences, and elementwise products provide spatial correspondence features. Spatial, moment, and gradient pooling yield three tokens per level. A classification token and a projected action embedding accompany these nine pair tokens in a small Transformer; two linear heads output safety and progress logits. Safety labels distinguish anatomically noninferior transitions from harmful ones using clean-referenced evidence. Progress labels mark safe transitions whose lower-bound utility exceeds the meaningful-gain margin. Conflicting evidence is excluded from the affected loss. For binary targets y_s, y_p and validity masks w_s, w_p, the objective is

$$\mathcal{L}_V = \frac{\sum_i w_{s,i}\,\mathrm{BCE}(v_{s,i}, y_{s,i})}{\sum_i w_{s,i}} + \lambda_p \frac{\sum_i w_{p,i}\,\mathrm{BCE}(v_{p,i}, y_{p,i})}{\sum_i w_{p,i}}, \tag{13}$$

with any empty term omitted. Anatomically unsafe pairs supervise safety but not progress; unresolved conflicts supervise neither head. Safe pairs without robust progress supervise a negative progress label. STOP identity pairs are excluded. The separately trained verifier is frozen throughout agent optimization.

Supervised policy initialization. Bounded nonrepeating continuation graphs provide full-reference action utilities. For utility replicate b, a temperature-scaled softmax over evaluated legal actions gives a soft target; averaging replicates preserves uncertainty rather than assigning a hard best-action label. Conditional expert targets are renormalized after removing STOP. Confidence-weighted cross-entropy trains expert ranking. A separate soft binary target compares STOP with the best supervised expert and trains the stopping policy. Unsupervised actions and unresolved stopping targets are masked. The two losses update their respective branches; GT-based utilities and uncertainty enter the targets, not the observation vector.

Terminal reward. Let $e_{\mathrm{HF}}(v, x^\star)$ denote mean absolute log-amplitude spectral error after foreground masking, over 3D Fourier frequencies with radial magnitude at least 0.25 cycles per voxel. Define

$$q(v) = -\log\bigl(1 + e_{\mathrm{HF}}(v, x^\star)/s_{\mathrm{HF}}\bigr), \quad \Delta q = q(x_T) - q(x_0). \tag{14}$$

The positive scale s_{HF} is fitted from training-subject error summaries and frozen. For a trajectory with n attempted calls,

$$R = \begin{cases} \Delta q - cn, & \text{anatomically noninferior,} \\ \min(\Delta q, 0) - 8 - cn, & \text{otherwise,} \end{cases} \tag{15}$$

where $c = 0.0154154$. A frozen SynthSeg model supplies the anatomical proxy: compared with the parent or root, allowed clean-referenced Dice drops are 0.02 for the macro average and 0.05 for the minimum-label score. Rejected attempts still incur call cost. Verifier probabilities are not reward inputs. This spectral reward is distinct from the LoG-based HFEN evaluation metric.

Group-relative optimization and preferences. Eight fresh trajectories per input are sampled from $\mu_\theta = 0.8\pi_\theta + 0.2\,\mathrm{Uniform}(\mathcal{A}_t)$ over legal actions including STOP. The clipping radius is $\epsilon_c = 0.2$. Within each group, terminal rewards are standardized to $A_i = (R_i - \bar{R})/(\sigma_R + \epsilon)$ and assigned to all decisions of trajectory i. The clipped objective uses ratios $r_{i,t} = \mu_\theta(a_{i,t} \mid s_{i,t})/\mu_{\mathrm{old}}(a_{i,t} \mid s_{i,t})$:

$$\mathcal{L}_{\mathrm{GRPO}} = -\mathbb{E}_i \frac{1}{T_i} \sum_t \min\bigl(r_{i,t}A_i, \mathrm{clip}(r_{i,t}, 1-\epsilon_c, 1+\epsilon_c)A_i\bigr). \tag{16}$$

A separate memory retains unique trajectories from the same input and frozen environment. Reward differences exceeding their uncertainty spreads define preferred pairs (ζ^+, ζ^-). With $\ell_\theta(\zeta)$ the summed policy log-probability of its decisions, the auxiliary loss is

$$\mathcal{L}_{\mathrm{pref}} = \mathbb{E}\ \mathrm{softplus}\bigl[-\ell_\theta(\zeta^+) + \ell_\theta(\zeta^-)\bigr], \qquad \mathcal{L}_{\mathrm{RL}} = \mathcal{L}_{\mathrm{GRPO}} + \mathcal{L}_{\mathrm{pref}} + 0.05\, D_{\mathrm{KL}}(\pi_\theta \| \pi_{\mathrm{SFT}}). \tag{17}$$

Historical and supplemental trajectories contribute only to preference supervision, never to fresh-group normalization. The supervised budget allocator is fitted separately using the ordinal objective above.

Policy training settings. The evaluated policy-training configuration starts from SFT with fresh AdamW and uses the settings below. Experts, the MRI encoder, Verifier, and the STOP branch's copied action bank remain frozen.

Setting	Value
Fresh rollouts per input	8
Behavior exploration	$0.8\pi_\theta + 0.2\,\mathrm{Uniform}(\mathcal{A}_t)$
GRPO clipping radius	0.2
Preference weight λ_{pref}	1
SFT KL weight β	0.05
Call cost c	0.0154154
Anatomical penalty λ_{anat}	8
Learning rate	10^{-5}
Training inputs / optimizer updates	256 / 64
Advantage	Terminal reward; turn-level weight 0

B Endpoint Assessment and Residual Refinement

3D abstraction and comparison. Three feature levels are projected to width 256. Normalized 3D coordinates are linearly embedded and added with learned level embeddings. A bank of 64 query vectors, initialized once from $\mathcal{N}(0, 0.02^2)$, attends to the supported feature memory through one eight-head cross-attention block. A linear projection maps the outputs to the comparison decoder's embedding width. The abstractor is shared between exemplar and endpoint. The frozen prompt embedding table supplies text tokens; visual tokens replace the two image placeholders in exemplar–query order. The final sequence position supplies the five response-token logits, normalized within that five-word set.

Ruler supervision and reference bank. Human review assigns binary acceptable/unacceptable labels to MRI states; uncertain labels are resolved or excluded. Each training pair contains an independently accepted exemplar and a labeled query. The query label y supervises its grouped comparison probability:

$$\mathcal{L}_{\mathrm{Ruler}} = -\mathbb{E}_{(e_k,x,y)}\big[y \log c_k(x) + (1-y)\log(1-c_k(x))\big]. \tag{18}$$

Thus supervision applies to each pair, whereas inference averages comparisons over the fixed reference set. The reviewed bank contains eight exemplars; four fixed exemplars are used for tuning, with the full bank also used for sensitivity analysis. Exemplars provide quality references, not paired anatomical targets. The abstractor is optimized with AdamW at 10^{-4}, five warmup updates, and eight accumulated pairs; encoder and decoder weights remain frozen. Threshold selection is separate from fitting the pairwise score and uses development labels only. The scalar s_{acc} is an acceptability score; passing the gate returns the existing endpoint, while a refined output is returned without a second acceptance decision.

Detector architecture and training stages. The 3D U-Net has encoder widths 8/16/32, two 3^3 convolutions with group normalization and SiLU per block, average-pooling downsampling, and trilinear upsampling with skip concatenation. A 1^3 output convolution and softplus produce a nonnegative map. Let ℓ_m denote supported smooth-L_1 with transition 0.1. Paired pretraining uses a corrupted input u, its expert endpoint v, and clean reference $x^\star$:

$$\begin{aligned}\mathcal{L}_D = \tfrac{1}{2}\sum_{z\in\{u,v\}} \ell_m(D_\psi(z), r_z^\star) \\ + 0.25\ell_m(D_\psi(v) - D_\psi(u), r_v^\star - r_u^\star) + 0.25\ell_m(D_\psi(x^\star), 0).\end{aligned} \tag{19}$$

Endpoint adaptation subsequently optimizes $\frac{1}{2}\ell_m(D_\psi(v), r_v^\star)$ alone. Neighborhood targets use shared image coordinates and intensity units, with loss evaluated only at supported voxels. The fixed detector checkpoint provides predicted maps for refiner training and inference.

BME-X adaptation. We initialize the 3D reconstruction backbone from pretrained BME-X (Sun et al., 2025). A fixed mapping $100(x+5)$ converts standardized intensities in $[-5, 5]$ to its input scale. Physical geometry is preserved using orientation permutations and flips; no repeated bias correction, histogram matching, or spatial resampling is applied. Full-resolution features have 112 channels. The conditioner takes $[x/5, \bar{r}]$ through a 3^3 convolution of width 16, SiLU, and a 1^3 projection to 224 scale/shift channels. Equation 11 uses group normalization with eight groups, $\kappa = 0.25$, and $\rho = 0.1$. A 3^3 residual head maps the modulated features to one channel. Conditioner outputs and the residual head are initialized to zero. The late reconstruction convolutions, conditioner, and residual head are learned; earlier layers, the segmentation pathway, and BatchNorm statistics remain fixed.

Refinement loss and patch aggregation. For support Ω, let $\langle\cdot\rangle_\Omega$ denote a voxel mean. The loss is

$$\begin{aligned}\mathcal{L}_R = \tfrac{1}{5}\langle|x' - x^\star|\rangle_\Omega + \tfrac{0.5}{15}\sum_{d=1}^{3}\langle|\nabla_d x' - \nabla_d x^\star|\rangle_{\Omega_d} \\ + \tfrac{0.01}{5}\langle(1-\bar{r})|x' - x|\rangle_\Omega,\end{aligned} \tag{20}$$

where ∇_d divides finite differences by physical voxel spacing and Ω_d requires both adjacent voxels to be supported. AdamW uses learning rate 10^{-5}, weight decay 10^{-4}, and effective batch size eight. Training uses 40^3 patches with central 30^3 supervision and stride 30. Inference uses stride 18 and averages overlapping central predictions; unsupported voxels retain the endpoint. Patch aggregation constitutes one volumetric refinement pass.

Image support and geometry. All states of a paired trajectory retain their original prepared grid, support, and intensity transform. For simulated paired data, the clean-derived preprocessing mask and normalization are fixed before corruption and shared across methods; they are not recomputed from a restored candidate. This retrospective paired preprocessing is distinct from an input-only preprocessing pipeline. Agent feature pooling uses the valid volume grid, while Ruler attention, detector supervision, and refiner masking use the prepared support. The detector itself consumes only an image. Paired clean images provide targets and evaluation measurements, not an exemplar matched to the query subject. Dataset-specific external preprocessing is specified with the evaluation protocol.

C External Evaluation

IXI uses matched simulated artifact realizations grouped by the number of artifact types: Group A comprises bias field (B), motion (M), and noise (N); Group B comprises BM, BN, and MN; Group C comprises BMN. All methods share the corrupted inputs, GT images, and evaluation masks. Metrics are computed per volume and averaged within each group; artifact severity is distinct from this grouping. GT images support PSNR, 3D SSIM, and NRMSE evaluation. ABIDE uses naturally corrupted scans and frozen no-reference perceptual-quality scorers. Models and decision thresholds are fixed before external evaluation. ABIDE cortical reconstruction uses the same software and configuration for all methods; failed reconstructions are retained in completion accounting. No paired-reference metric is computed on ABIDE.

Supplementary high-frequency fidelity. Normalized three-dimensional high-frequency error norm (HFEN) complements the full-volume fidelity metrics:

$$\mathrm{HFEN}(\widehat{y}, y) = \frac{\|\mathcal{L}_\sigma(\widehat{y}) - \mathcal{L}_\sigma(y)\|_{2,\Omega_{\mathrm{int}}}}{\|\mathcal{L}_\sigma(y)\|_{2,\Omega_{\mathrm{int}}}}, \tag{21}$$

where $\mathcal{L}_\sigma$ is a three-dimensional Laplacian-of-Gaussian operator with physical standard deviation $\sigma = 1.5$ mm, truncated at four standard deviations. Derivatives account for voxel spacing without resampling the image. The common evaluation mask Ω_{int} excludes voxels whose filter footprint extends outside the foreground support or volume. Empty masks and zero reference denominators are reported as undefined with completion counts. Predictions and GT images share the fixed intensity convention; outputs are not independently rescaled or clipped for scoring. HFEN is aggregated by subject and artifact group and measures high-frequency discrepancy, not diagnostic accuracy.

Independent reader assessment. The IXI reader protocol specifies 24 distinct subjects with one corrupted input each, eight per artifact group. Selection uses metadata and fixed random seeds before method outputs are inspected, with coverage of artifact combinations and severities. Two readers with brain-MRI expertise independently review the corrupted input and all six system outputs under anonymized method identities and randomized presentation order. Readers use synchronized three-plane volume browsing, matched display settings, and the corresponding GT image. Separate anchored five-point scales assess residual-artifact burden and tissue-boundary preservation; a separate absent/present/uncertain judgment records suspected newly introduced structural changes. These assessments concern restoration quality and do not establish diagnostic performance. Development QC labels and Ruler-training labels are excluded from the independent reader outcomes.

ABIDE no-reference evaluation. We use MUSIQ, MANIQA, and CLIP-IQA with fixed pretrained weights and implementation versions (Ke et al., 2021; Yang et al., 2022; Wang et al., 2023). We use PyIQA 0.1.13, with the KonIQ-10k MUSIQ and MANIQA checkpoints and the RN50 CLIP-IQA implementation. The latter averages five positive/negative prompt pairs and is a prompt-ensemble variant of the original CLIP-IQA. These natural-image scorers are used without MRI-specific calibration or ABIDE-specific tuning. Their scores are exploratory perceptual-quality measurements and do not establish anatomical correctness or clinical acceptability; they are not converted to acceptance or damage rates.

For each subject, slice locations, the brain-region crop, and intensity-display bounds are determined from the original input and frozen across all methods. Nine evenly spaced locations between the 10th and 90th percentiles of the input brain-mask extent are sampled in each of the sagittal, coronal, and axial orientations. Grayscale slices are repeated into three identical channels. The input brain-foreground 1st and 99th intensity percentiles define a shared linear mapping to $[0, 1]$ for IQA only; clipping fractions are recorded. Restoration outputs are never independently contrast-normalized. Deterministic crops, resizing and seeds required by each scorer are fixed and shared across methods; annotated figure panels are not scored.

Slice scores are averaged within each orientation, then equally across the three orientations to give one score per subject. Cohort results report subject means, paired changes from the original input, and 95% paired subject-bootstrap intervals with 10,000 resamples and seed 23. The same resampled subjects are used for all methods; slices are not treated as independent subjects. Missing outputs, undefined scores and method-specific completion counts are reported explicitly, without replacing

failed cases or substituting zero scores. These scorers are reserved for evaluation and are not used to select subjects, checkpoints, trajectories or restoration endpoints. Any SHN results remain separate downstream measurements and do not establish anatomical correctness.

D Additional Related Work

Unified and agentic restoration. Instruction-conditioned and degradation-aware restoration combine multiple tasks in a shared model (Conde et al., 2024; Jiang et al., 2024; Tian et al., 2025). Agentic approaches select specialized tools from image feedback (Chen et al., 2024; Zhu et al., 2025; Jiang et al., 2025). RAR integrates restoration and assessment in an iterative framework (Chen et al., 2026). MRI-CARE focuses on volumetric expert composition and endpoint acceptability, with explicit residual guidance after expert execution.

MRI restoration and anatomical fidelity. Bias fields, noise, and motion require different correction mechanisms (Tustison et al., 2010; Maggioni et al., 2013; Singh et al., 2024). Their effects extend to quantitative anatomy (Reuter et al., 2015; Elyounssi et al., 2025). We therefore complement image fidelity with cortical reconstruction measurements and independent acceptance assessment.

E Artifact-Selective Expert Training

Simulation and pairing. Bias-field simulation applies a smooth spatially varying multiplicative field, following the intensity-nonuniformity model used in MRI bias correction (Tustison et al., 2010). Motion simulation combines k-space data from different head poses to reproduce acquisition inconsistencies (Zaitsev et al., 2015). Magnitude-image noise is modeled with Rician statistics (Gudbjartsson & Patz, 1995), with a fixed realization for each corruption recipe. Mixed inputs use an ordered recipe rather than an unordered sum of artifacts. Each recipe records the active artifacts, strengths, random seeds, and operation order. A counterfactual target is generated from the same clean volume by omitting only the designated operation and replaying all retained operations with their recorded parameters and randomness. Consequently, a partial target retains the same nuisance corruptions as its paired input; it is not generated by independently resampling them. Corrupted inputs, partial targets, and clean references share spatial geometry and preprocessing support.

Selective supervision. For clean image $x^\star$, a fixed ordered corruption recipe ω defines the input $\mathcal{C}_\omega(x^\star)$. The target removes the designated artifact while retaining the other corruptions and their random realizations:

$$(x_a, y_a) = \big(\mathcal{C}_\omega(x^\star), \mathcal{C}_{\omega\setminus a}(x^\star)\big). \tag{22}$$

The order of the remaining operations is preserved; the notation does not assume that MRI corruptions commute. Target-absent pairs supervise identity. Paired branches share geometry, normalization, and preprocessing support. Motion simulation precedes skull stripping. Subject roles are assigned before training-pair selection.

Backbone-guided expert interface. Equation 1 expresses the unified model-zoo design: an artifact-specific branch receives the observed volume and a frozen backbone reconstruction as structural context. The branch learns the designated correction from selective pairs, and the backbone remains frozen. Expert parameters are subsequently fixed for agent optimization.

Expert versions in recorded comparisons. The original policy-training bank uses a frozen-backbone bias-field predictor, a multiband noise restorer, and a Motion action combining a volumetric predictor with tri-axis full-plane restoration. In these recorded versions, Bias accesses multiscale backbone features directly, Noise uses the backbone reconstruction and its difference from the input, and Motion uses standalone restoration networks. These versions are distinguished from the common reconstruction-conditioned interface in Eq. 1. The Motion wrapper produces one full-volume correction before the candidate scaling in Eq. 3. All experts expose the same volume-to-volume interface and remain fixed during policy optimization.

Evaluation identity. Each comparison fixes the expert bank, candidate construction, preprocessing, and policy checkpoint. A full-dose Motion control differs from the residual-candidate configuration only in its allowed correction strengths; candidate scaling does not introduce additional expert forwards. Historical matched-path illustrations retain their original expert identities.

F Baseline Protocol and Artifact-Wise Results

Method families and applicability. Table 1 distinguishes general all-in-one restorers (PromptIR, InstructIR, and DFPIR), three task-specific Restormer-ST models, and the MRI-specific TAT comparison. TAT is retrained for MRI artifact correction rather than evaluated with its super-resolution weights. DFPIR receives artifact-aware prompts and is marked with †. Fixed-order and verifier/R^3 variants isolate composition decisions within MRI-CARE.

Cohort and aggregation. The reported PromptIR, InstructIR, DFPIR, and TAT-MRI results use the same 128 full-volume validation inputs from 49 HCP subjects. Two clean controls are excluded from the artifact-correction groups. Group A comprises B, M, and N (six inputs each); Group B comprises BM, BN, and MN (18 each); Group C comprises 54 BMN inputs. Here B, M, and N denote bias field, motion, and noise. The groups contain 16, 33, and 35 subjects, respectively; subjects can occur in more than one group. Each volume is scored separately, then metrics are averaged over inputs within a group. Equal counts per artifact combination make this equivalent to averaging the combination means. Repeated corruptions are not independent subjects.

Full-volume fidelity metrics. PSNR and SSIM are shared fidelity measures in recent all-in-one restoration evaluations; medical restoration additionally reports RMSE (Shan et al., 2026), while photographic restoration also considers learned perceptual measures (Chen et al., 2026). We use PSNR, 3D SSIM, and Euclidean NRMSE for MRI. For foreground mask Ω, prediction $\widehat{y}$, and GT image y, the definitions are

$$\mathrm{PSNR} = 10\log_{10}\frac{(\max_\Omega y - \min_\Omega y)^2}{|\Omega|^{-1}\sum_{v\in\Omega}(\widehat{y}_v - y_v)^2}, \qquad \mathrm{NRMSE} = \frac{\|\widehat{y} - y\|_{2,\Omega}}{\|y\|_{2,\Omega}}. \tag{23}$$

The 3D SSIM uses a $7\times7\times7$ uniform window, unit stride, zero padding, and constants $(0.01L)^2$ and $(0.03L)^2$, with $L = \max_\Omega y - \min_\Omega y$. Prediction and GT are zeroed outside the foreground before local statistics are computed; the resulting SSIM map is averaged over foreground voxels. All methods use the same GT images, masks, and intensity convention. The table reports the mean of volume-wise PSNR values, not PSNR computed from a pooled voxel error.

F.1 Downstream Cortical Reconstruction

HCP reconstruction subset. The cortical comparison in Table 1 uses nine validation inputs from nine distinct subjects: one each from B, M, N, BM, BN, and MN, and three from BMN. A fixed seeded hash ordering selects inputs within artifact strata, excluding already selected subjects; selection uses manifest metadata without restoration scores or reconstruction outcomes. All methods share the selected inputs and paired GT images. This small subset provides descriptive downstream evidence; group-wise estimates contain only three inputs and do not establish population-level differences. Full-volume fidelity metrics retain the complete validation cohort.

Common reconstruction pipeline. Corrupted inputs, restored outputs, and paired GT images undergo the same FreeSurfer release, configuration, and preprocessing, preserving physical voxel geometry when assembling the three-dimensional volumes. Intensity conversion and reconstruction settings are fixed across methods; no method-specific manual surface edits are used. The clean reconstruction is a reference for measurement consistency, not anatomical ground truth.

Surface defects and cortical thickness. Following the use of surface defects for automated MRI quality assessment (Elyounssi et al., 2025), SHN is the sum of the left- and right-hemisphere surface hole counts recorded before topology correction. It is not computed from the final topology-corrected mesh. On HCP, let $t_r(\widehat{x})$ and $t_r(x_{\mathrm{clean}})$ denote mean cortical thickness in region r of the restored and

paired clean reconstructions. We use the 68 bilateral Desikan–Killiany cortical regions and compute

$$\text{CT-MAE}(\widehat{x}, x_{\text{clean}}) = \frac{1}{68} \sum_{r=1}^{68} |t_r(\widehat{x}) - t_r(x_{\text{clean}})| \,. \tag{24}$$

CT-MAE is in millimeters and gives equal weight to each region. Lower SHN indicates fewer initial surface defects; lower CT-MAE indicates closer agreement with the clean-reference measurements. Neither alone establishes anatomical correctness or clinical acceptability. Without paired clean ABIDE scans, thickness changes cannot be interpreted as thickness errors.

Aggregation and reconstruction completion. Cortical metrics are averaged over a common set of successfully reconstructed volumes within each comparison; CT-MAE additionally requires successful reconstruction of the paired GT image. Task-restricted comparators are evaluated separately on their applicable subsets. The HCP nine-input summary weights eligible inputs equally; with complete reconstruction this gives equal weight to the three groups, rather than estimating the 18:54:54 full-cohort mixture. The small HCP subset is reported descriptively, with paired case values and completion counts; confidence intervals for larger cohorts resample subjects, retaining repeated corruptions together. Completion accounting records successful reconstructions out of all attempted volumes for each method, the common comparison-set size, per-group eligible counts, and exclusions caused by clean-reference failures. Infrastructure failures are distinguished from image-dependent reconstruction failures. Failed cases are not replaced or assigned zero SHN or zero thickness error and remain in the denominator of blinded human acceptance.

F.2 Restoration Implementation and Computation

MRI adaptation and evaluated checkpoints. PromptIR and InstructIR use their 2D networks on axial slices, with a single MRI channel replicated into RGB. The fixed input mapping is $(x + 5)/10$; outputs are averaged over RGB channels and mapped back before reconstruction of the 3D volume and evaluation. No registration, intensity clipping, or spatial resizing is applied during evaluation. InstructIR uses the fixed instruction *Restore image quality while preserving anatomical structures.* Its language encoder is frozen; artifact and severity labels are not supplied at inference.

The reported PromptIR checkpoint is from epoch 24 (6,400 optimizer updates), selected using calibration L1. The reported InstructIR checkpoint is from the fixed epoch 90 (23,296 updates), without validation-based selection. DFPIR and TAT-MRI are each trained for 36,096 optimizer updates; their evaluated checkpoints are selected by calibration L1 at epoch 140. DFPIR uses the true artifact set to select its text prompt and is therefore marked task-aware. TAT-MRI operates on single-channel axial slices with the same fixed intensity mapping and full-volume metric protocol; artifact labels are not supplied at inference. Each of the three Restormer-ST models is trained for 36,096 optimizer updates, with its evaluated checkpoint selected by calibration L1 before validation. The oracle-routing variant uses the ground-truth artifact set to select only the corresponding Restormer-ST models, retaining Bias–Motion–Noise order; it is marked task-aware and does not select outputs by GT-based PSNR. The fixed cascade applies Bias, Motion, and Noise models to every input in that order. Restormer-ST uses the same axial inference, intensity mapping, GT images, masks, and full-volume metrics, without clipping or resizing. These checkpoints have different pretraining and optimization budgets; their comparison does not establish performance under equal training compute. All methods use identical validation inputs, references, and masks.

Computation and qualitative comparisons. Runtime and expert-call accounting are reported separately from fidelity. Attempted expert forwards include rejected candidates; encoding, verification, and other pipeline work contribute to end-to-end cost. A unified-model forward and an expert call are not assumed to have equal cost. Qualitative comparisons use identical input/GT pairs, slice locations, display windows, and regions of interest, spanning single-, two-, and three-artifact inputs.

F.3 Ablation Controls and Decision Analysis

Paired configuration controls. Agent comparisons use matched evaluation inputs, frozen experts, and an identical call cap. Fixed-order selection changes ordering only, retaining the verifier and STOP policy. Inference-time verifier and stopping interventions use the same GRPO policy checkpoint;

SFT-only uses its pre-reinforcement-learning checkpoint. Calls count attempted expert actions, including rejections; Motion dose candidates are not counted as separate actions. The SFT evaluation and GRPO control protocol use deterministic nonrepeating selection, Motion doses $\{0.5, 0.75, 1.0\}$, and safety threshold 0.05. Progress scores are recorded without a hard veto. The two clean controls are excluded, leaving 126 corrupted inputs (54 in Group C). These are development-validation results, not an untouched test set. The evaluated GRPO checkpoint uses 64 updates, a terminal high-frequency-error reward, preference-memory assistance, and zero turn-level reward weight; it evaluates the terminal-reward configuration rather than the full dual-level objective. Table 2 reports the learned agent before and after residual refinement. Fixed one-, two-, and three-call policies provide quality–cost comparisons with learned stopping, with the ordering policy and verifier held fixed.

Ruler assessment and threshold selection. Ruler discrimination is assessed against independent human acceptability labels, with AUROC and the fraction of unacceptable endpoints incorrectly passed through the gate. This false-acceptance rate conditions on human-unacceptable endpoints. Threshold selection uses development data only; the threshold is frozen before evaluation. Ruler-training and tuning labels are excluded from independent acceptance outcomes. All-endpoint and gated refinement use the same refiner weights, allowing changes in quality, human acceptance, refinement frequency, and runtime to be attributed to gating.

MRI-CARE evaluation protocol. The MRI-CARE row in Table 1 is identical to the full MRI-CARE row in Table 2. At inference, the learned planner selects nonrepeating experts, the verifier commits or rolls back each attempted correction, and the STOP policy returns the agent endpoint. R^3 then applies at most one acceptability-gated refinement pass. Neither the agent nor R^3 observes the paired clean image. GT images are used only to compute PSNR, whole-volume 3D SSIM, and NRMSE, which are averaged over all inputs within each group.