

패치 기반 수의학 X-ray 영상 복원을 위한 확산 기반 경계 정제 기법

조연경¹ · 장효영¹ · 김지훈¹ · 김정년¹ · 김민형² · 한수민² · 박승조² · 이상민^{1*}

¹광운대학교 인공지능응용학과 / ²전남대학교 수의학과

Diffusion Refinement for Boundary-Consistent Veterinary X-ray Reconstruction

Yeongyeong Jo¹ · Hyoyoung Jang¹ · Jihoon Kim¹ · Jeongnyeon Kim¹ · Minhyung Kim² · Sumin Han² ·
Seungjo Park² · Sangmin Lee¹

¹Department of Artificial Intelligence Applications, Kwangwoon University

²Department of Veterinary Medicine, Chonnam National University

In AI-based analysis pipelines, large medical images are often restored in patches for high-resolution processing due to memory limitations. Patch-wise restoration approaches cause brightness and texture stitching artifacts at patch boundaries, degrading structural clarity. In this study, we propose a method to remove artificial lines when generating high-resolution X-ray images from patches. We present a selective refinement pipeline focusing on patch boundaries. Using pre-restored high-resolution patches, we apply a soft boundary mask and perform noise-based reconstruction only in boundary areas. This process helps blend patch edges smoothly without extra post-processing. Experimental results show that the proposed method improves both structural and visual quality. Artificial seams are reduced, and the shapes of bones and tissues appear more natural and continuous. This approach offers a practical way to improve patch-based medical image restoration by making boundaries clearer and more consistent.

Keywords: Computer Vision, Diffusion Model, Medical Image Super-Resolution, Patch-wise Reconstruction, Boundary Refinement

1. 서 론

X-ray 영상은 비침습성과 비용 효율성으로 인해 수의학 분야에서 가장 널리 활용되는 핵심 진단 영상이다. 수의학 X-ray는 해부학적 다양성, 촬영 비표준성, 고해상도 및 국소 병변의 존재로 인해 사람 환자의 X-ray와는 구별된다. 고해상도 X-ray 영상은 이미지 크기와 복잡한 조직 구조로 인해 분석 과정에서 높은 연산 비용을 요구한다(Ahmadian and Reza-Alikhani, 2022). 메모리 및 연산 자원의 한계로 해부학적 구조의 경계 정

보가 중요함에도 X-ray 영상을 패치 단위로 분할하여 처리하는 방식이 사용되고 있다(Ahmadian and Reza-Alikhani, 2022). 패치 단위 접근은 특히 초해상도(super-resolution; SR)나 잡음 제거와 같은 영상 복원 작업에서 윤곽 보존과 계산 효율성을 동시에 확보할 수 있다(Ahmadian and Reza-Alikhani, 2022). 최근의 딥러닝 기반 영상 복원 모델들은 수많은 파라미터와 깊은 네트워크 구조로 인해, 추론 시 GPU 메모리 사용량과 처리 시간이 증가하는 문제가 존재한다(Weese and Lorenz, 2016). 이러한 한계는 고해상도 의료 영상 처리에서 더욱 두드러지며 실시

* 연락저자 : 이상민 교수, 01897, 서울특별시 노원구 광운로 20, 광운대학교 새빛관 607호, Tel : 02-940-5764, Fax : 02-940-5121,
E-mail : smlee5679@kw.ac.kr

2026년 4월 8일 접수; 2026년 6월 16일 수정본 접수; 2026년 7월 8일 게재 확정.

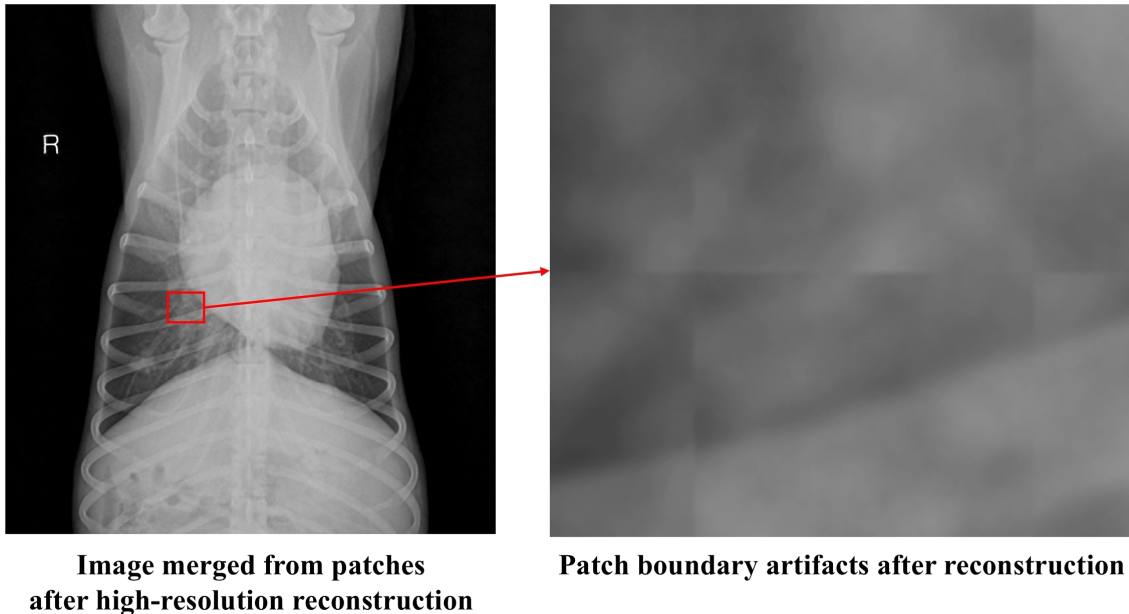


Figure 1. Detected Boundary Artifacts after Patch-wise High-resolution Restoration

간 적용이나 임상 환경으로의 통합을 저해한다. 반면, 고해상도로 복원된 영상은 구조물의 경계와 윤곽을 정밀하게 재현하여 전문가의 평가와 진단에 중요한 정보를 제공한다. 따라서 고해상도 복원은 영상의 진단적 가치와 구조적 신뢰도 확보를 위한 핵심 요인으로 간주된다(Ahmadian and Reza-Alikhani, 2022; Weese and Lorenz, 2016).

패치 단위 영상 복원 방식은 고해상도 영상 처리에서 연산 효율성과 병렬 처리에 강점을 가지므로 영상 복원 분야에서 널리 활용되고 있다. 다만, 인접 패치 간의 통계적 불일치로 밝기 및 질감의 부조화가 발생하며, 경계 불연속(boundary discontinuity)으로 이어진다(Buglakova *et al.*, 2025). 경계 불연속은 각 패치가 독립적으로 복원되는 구조에서 기인하며, 패치를 병합(stitching)하는 과정에서 경계 부근에 밝기 및 질감의 불연속으로 나타난다. 이는 단순한 시각적 결함을 넘어 X-ray 영상에서 중요한 해부학적 경계의 연속성을 훼손하여 구조적 신뢰도를 저해한다. 조직 경계가 진단에 중요한 역할을 하는 수의학 X-ray 영상에서는, 경계 왜곡이 해부학적 정보 손상으로 이어져 임상적 판단의 불확실성으로 직결된다.

기존 연구들은 이러한 문제를 완화하기 위해 결정론적(deterministic) 패치 병합 기법과 학습 기반 전역 복원(global restoration) 모델을 포함한 다양한 접근법을 제안해왔다(Pérez *et al.*, 2003; Pielawski and Wählby, 2020). Gaussian 또는 Hann window 기반의 가중 평균 방식은 대표적인 결정론적 병합 기법으로, 구현이 간단하나 패치 간 통계적 차이를 반영하지 못해 일관된 질감 복원이 어렵다. 반면, convolutional neural network(CNN), U-Net, generative adversarial network(GAN) 기반 전역 복원 모델은 전체 문맥 정보를 학습하여 구조적 연속성을 확보할 수 있으나(Ledig *et al.*, 2017; Ronneberger *et al.*,

2015; You *et al.*, 2019), 고해상도 의료 영상의 대용량 특성상 전체 학습이 어렵고, 패치 단위 병합에서 발생하는 경계 불연속을 근본적으로 해결하지 못한다(Uzunova *et al.*, 2019).

최근에는 영상 생성 및 복원 분야에서 확산확률모델(diffusion probabilistic model) 기반 연구가 활발히 수행되고 있다. 확산확률모델은 데이터에 점진적으로 노이즈를 추가한 뒤 이를 제거하는 반복적 역확산 과정을 통해 복잡한 데이터 분포를 학습하여 고품질 영상을 생성 및 복원하는 생성 모델이다(Ho *et al.*, 2020). 그러나 고해상도 영상을 패치 단위로 분할하여 처리하는 경우 패치 간 정보 단절로 인해 경계 부근에서 불연속성이 발생할 수 있으며, 이를 완화하기 위한 다양한 경계 정제 연구가 수행되고 있다.

기존 확산 기반 영상 복원 연구들은 영상 전체의 전역적 복원에 초점을 두어, 패치 기반 복원에서 발생하는 국소적 경계 불연속 문제나 경계 중심 영역의 선택적 정제(refinement)는 충분히 다루지 못한다(Saharia *et al.*, 2022a). 특히 super-resolution via repeated refinement(SR3)와 같은 확산 기반 모델은 반복적 복원을 통해 고품질 영상을 생성하나, 영상 전체를 대상으로 설계되어 경계 중심의 선택적 복원에는 적합하지 않다(Saharia *et al.*, 2022a). 결과적으로 패치 병합으로 인해 통계적 불확실성과 질감 변동성이 집중되는 경계 영역은 명시적인 복원 대상으로 설정되지 않아 패치 내부의 복원 품질이 우수해도 경계 불연속 문제는 잔존한다.

본 연구는 확산확률모델의 확률적 복원 능력을 패치 기반 영상 복원에 접목하여, 경계 불연속 문제를 해결하고자 한다. 제안 기법은 경계 중심 영역에 점진적으로 변화하는 가중치를 적용하여 해당 부위만 선택적으로 복원함으로써, 불필요한 연산을 줄이는 동시에 경계의 시각적 자연스러움을 효과적으로

확보하도록 설계되었다.

논문의 구성은 다음과 같다. 제2장에서는 패치 단위 영상 복원 과정에서 발생하는 경계 불연속 문제를 중심으로 기존 연구의 접근 방법과 한계점을 고찰한다. 제3장에서는 초해상도 복원 후 병합 과정의 경계 불일치 문제를 해결하기 위한 제안 방법론을 설명한다. 제4장에서는 제안 기법의 성능을 입증하기 위한 정량적 및 정성적 실험 결과를 제시하고, 제5장에서는 본 연구의 주요 성과를 요약 후 향후 연구 방향을 논의한다.

2. 관련 문헌 연구

본 장에서는 제안하는 경계 정제(boundary refinement) 기법과 관련된 연구를 세 가지 관점에서 고찰한다. 첫째, 고해상도 영상 생성을 위한 패치 단위 영상 복원 기법, 둘째, 확률적 복원 메커니즘을 제공하는 확산확률모델의 응용 확장, 셋째, 병합 과정에서 발생하는 경계 불연속을 완화하기 위한 경계 정합 기법 및 블렌딩(blending) 방법론을 정리한다. 특히 패치 기반 복원 과정에서 발생하는 패치 간 통계적 연속성(statistical consistency) 문제를 중심으로 기존 연구의 한계를 분석한다.

2.1 패치 단위 초해상도 기반 복원

패치 단위 초해상도 기법은 제한된 메모리 환경에서도 고해상도 복원을 가능하게 하는 효율적인 접근 방식으로, 의료 영상을 포함한 다양한 영상 복원 분야에서 널리 활용됐다(Zhao *et al.*, 2018). 패치 단위 처리는 전체 영상을 작은 패치로 분할하여 처리함으로써 계산 자원을 효율적으로 사용할 수 있다는 장점이 있다. 그러나 패치 기반 복원에서는 각 패치가 독립적으로 처리되기 때문에 인접 패치 간 통계적 분포가 일치하지 않는 문제가 발생할 수 있으며, 이로 인해 병합 단계에서 경계 불연속 현상이 나타난다(Mac Aodha *et al.*, 2012). 이는 결국 패치 간 통계적 연속성 저하로 이어진다. 특히 X-ray 영상과 같이 조직 경계의 정밀한 표현이 중요한 경우, 경계 불연속은 해부학적 구조의 왜곡과 진단 신뢰도 저하를 유발한다.

패치 기반 복원에 활용되는 대표적인 초해상도 모델로는 enhanced deep super-resolution network(EDSR), residual channel attention network(RCAN), swin transformer for image restoration(SwinIR) 등이 있다(Liang *et al.*, 2021; Lim *et al.*, 2017; Zhang *et al.*, 2018). 이들은 저해상도 입력으로부터 고해상도 영상을 직접 생성하는 결정론적 복원 모델로서 다양한 영상 복원 과제에서 우수한 성능을 보였다. 그러나 각 패치를 독립적으로 복원하기에 인접 패치 간 상호작용을 충분히 고려하지 못하며, 패치 병합 과정에서 발생하는 경계 불연속 문제를 근본적으로 해결하지 못한다.

최근에는 확산확률모델의 도입으로 초해상도 복원 성능이 크게 향상되었다(Li *et al.*, 2024). SR3는 반복적인 역확산(re-

verse diffusion) 과정을 통해 고해상도 영상을 생성하였으며(Saharia *et al.*, 2022a), super-resolution diffusion(SRDiff)는 잔차 학습 방식을 적용하여 복원 정밀도를 향상시켰다(Li *et al.*, 2022). 기존 결정론적 접근보다 복원 품질은 우수하지만, 전역 복원 중심이어서 패치 병합 시 국소 경계 정합 문제에 한계가 있다. 따라서 기존 패치 기반 방식은 효율적이지만, 패치 간 경계 정합을 위한 구조적 설계에는 한계가 있다(Ma *et al.*, 2021).

2.2 Diffusion Model의 응용 확장 연구

확산확률모델은 확률적 복원 능력과 안정적인 학습으로 다양한 영상 복원 및 편집 과제에 적용되어 왔다. inpainting using denoising diffusion probabilistic models(RePaint)는 마스크로 가려진 손상 영역을 반복적 샘플링 과정으로 복원하며(Lugmayr *et al.*, 2022), image-to-image diffusion models(Palette)는 다양한 이미지 복원 작업을 단일 프레임워크에서 수행한다(Saharia *et al.*, 2022b). 또한 stochastic differential editing(SDEdit)은 조건부 입력을 주입하여 원본 구조를 유지하면서도 부분적인 이미지 편집이 가능하다(Meng *et al.*, 2021). 최근에는 복원 대상 영역의 공간적 특성을 반영하여 선택적으로 복원을 수행하는 region-aware diffusion(RAD) 기반 접근법도 제안되었으며, 영역별 특성에 적응적인 복원이 가능함을 보여주었다(Kim *et al.*, 2025).

해당 연구들은 확산확률모델이 기존의 결정론적 복원 방식에 비해 데이터의 불확실성을 효과적으로 반영함을 보여준다(Lugmayr *et al.*, 2022; Saharia *et al.*, 2022b). 그러나 대부분 전체 영상 복원이나 명시적 손상 영역에 초점을 두어, 패치 기반 복원 과정에서 발생하는 경계 불연속 문제는 다루지 못한다. 특히 RePaint의 반복 샘플링 방식은 높은 연산량을 요구하여 고해상도 의료 영상 환경에서는 실용성이 제한된다(Lugmayr *et al.*, 2022).

2.3 경계 정합 기법 및 Blending 방법론

경계 정합 문제는 영상 합성 및 복원 분야에서 오랫동안 연구되어 온 주제이다. 초기 연구는 밝기나 윤곽선 단절 완화에 초점을 두었으며, Poisson image editing과 Laplacian pyramid blending이 대표적이다(Burt and Adelson, 1983; Pérez *et al.*, 2003). 이러한 기법들은 색상 및 밝기 보정 중심으로 설계되어 구조적 질감의 연속성을 보장하지 못한다.

의료 영상 복원에서는 Hann window와 같은 테이퍼 함수(taper function) 기반의 결정론적 블렌딩 기법이 주로 활용되어 왔다(Pielawski and Wählby, 2020). Tukey window 기반 가중치 블렌딩이나 Laplacian pyramid blending 역시 경계 완화에 일정 부분 효과를 보이나, 미세한 질감이나 구조적 패턴의 통계적 분포를 충분히 반영하지 못한다(Burt and Adelson, 1983). <Table 1>은 본 장에서 검토한 대표적인 관련 연구들을 범주

Table 1. Summary of Related Studies

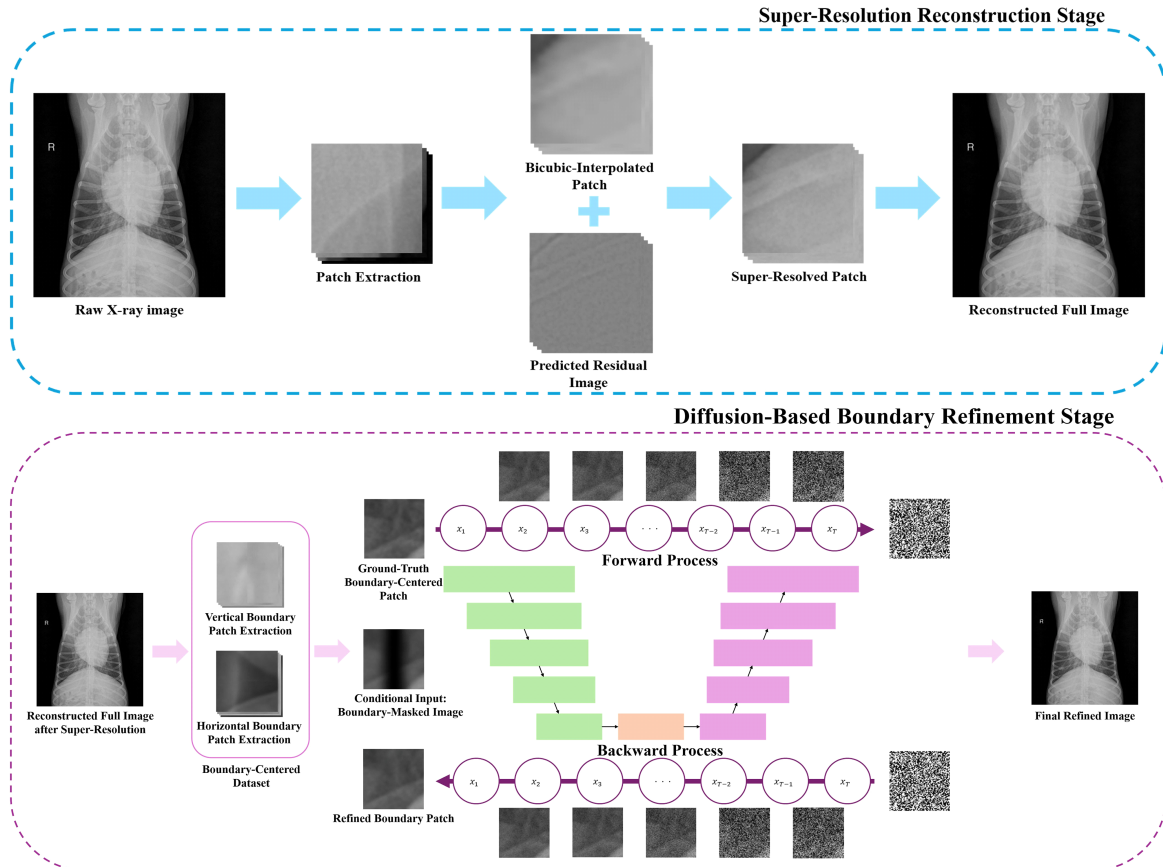
Category	Method	Reference
Patch-based SR	enhanced deep super-resolution network (EDSR)	Lim <i>et al.</i> (2017)
	residual channel attention network (RCAN)	Zhang <i>et al.</i> (2018)
	swin transformer for image restoration (SwinIR)	Liang <i>et al.</i> (2021)
Diffusion-based SR	super-resolution via repeated refinement (SR3)	Saharia <i>et al.</i> (2022a)
	super-resolution diffusion (SRDiff)	Li <i>et al.</i> (2022)
Diffusion-based Restoration	stochastic differential editing (SDEdit)	Meng <i>et al.</i> (2021)
	inpainting using denoising diffusion probabilistic models (RePaint)	Lugmayr <i>et al.</i> (2022)
	image-to-image diffusion models (Palette)	Saharia <i>et al.</i> (2022b)
Region-aware Diffusion	region-aware diffusion (RAD)	Kim <i>et al.</i> (2025)
Boundary Blending	Laplacian Pyramid Blending	Burt and Adelson(1983)
	Poisson Editing	Pérez <i>et al.</i> (2003)
	Hann Window	Pielawski and Wählby(2020)

별로 요약한 것이다(<Table 1>).

결과적으로 기존 연구들은 패치 단위 복원의 연산 효율성, 확산확률모델의 확률적 복원 능력, 그리고 결정론적 블렌딩 기법의 경계 완화 측면에서 각각의 성과를 보여 왔다. 그러나 패치 간 통계적 연속성은 확보되지 않아, 경계 불연속 문제를 해결할 새로운 접근이 필요하다. 이를 해결하기 위해 본 연구는 확산확률모델 기반 경계 정제 기법을 제안한다.

3. 제안 방법론

<Figure 2>는 제안 기법의 전체 파이프라인을 나타내며, 초해상도 복원 단계(super-resolution restoration stage)와 확산 기반 경계 정제 단계(diffusion-based boundary refinement stage)로 구성된다. 첫 번째 단계에서는 입력 X-ray 영상을 패치 단위로 분할 후, SR3 및 SRDiff 기반 확산확률모델로 저해상도 패치를

**Figure 2.** Overall Pipeline for High-resolution Image Restoration and Boundary Refinement

고해상도 패치로 복원한다. 복원 결과는 전체 영상의 초기 고해상도 영상을 구성하며 후속 경계 정제 단계의 입력으로 사용된다. 이후 인접 패치 간 경계 위치를 기준으로 경계 중심 패치를 추출하여 경계 영역 학습 데이터셋을 구성한다.

두 번째 단계에서는 추출된 경계 중심 패치에 확산확률모델 기반 경계 정제를 수행한다. 점진적 가중치 함수로 정의된 soft mask를 적용하여 경계 중심 영역을 선택적으로 강조하고, 해당 영역에 마스크된 확산 샘플링을 적용한다. 학습 시에는 정답 경계 패치에 순방향 확산(forward diffusion)을 적용하여 노이즈를 단계적으로 추가하고 모델이 각 시점의 노이즈를 예측하도록 학습하며, 추론 시에는 완전한 노이즈 상태에서부터 역확산을 반복하여 경계 영역의 세부 구조를 복원한다. 복원된 경계 패치는 원래 경계 위치에 재삽입되며, 패치 내부 구조는 유지한 채 경계 영역의 통계적 분포만 재구성하도록 설계되었다. 제안 기법은 패치 병합 시 경계 불연속을 완화하고, 구조적 연속성 및 시각적 자연성을 향상시키는 것을 목표로 한다.

3.1 초해상도 복원 단계

본 연구의 기반 복원 단계는 확산확률모델을 이용한 초해상도 복원 기법인 SR3와 그 잔차 기반 변형인 SRDiff를 기반으로 구축되었다(Li *et al.*, 2022; Saharia *et al.*, 2022a). 저해상도(LR) X-ray 영상으로부터 고해상도(HR) 복원 영상을 생성하고, 후속 확산확률모델 기반 경계 정제 단계의 입력에 사용할 예측 패치를 생성하는 데 있다.

(1) 패치 단위 복원 구조

전체 X-ray 영상은 연산 효율성과 지역적 세부 복원을 동시에 확보하기 위해 128×128 크기의 패치 단위로 분할되었다. 인접 패치 간에는 stride=64(약 50% 중첩)를 적용하여 중첩 구간의 정보 손실을 최소화하고, 복원 결과의 지역적 연속성을 유지하였다. 패치 기반 복원 방식은 메모리 사용량을 효과적으로 제어할 수 있으나, 경계 부위에서 밝기 및 질감의 불연속이 발생한다. 본 연구에서는 초해상도 복원 결과를 후속 경계 정제 단계의 입력으로 활용함으로써, 확산 기반의 확률적 복원 과정으로 경계 불연속 문제를 보정하도록 설계하였다. 모델은 LR로부터 직접 HR을 생성하는 대신, 조건 입력을 기반으로 세밀한 잔차(residual) 정보를 확률적으로 복원하며, 복원 영상의 질감 표현을 향상시킨다.

(2) 학습 데이터 구성 및 입력 정의

본 연구의 초해상도 복원 단계는 저해상도 영상으로부터 고해상도 영상을 복원하기 위한 조건부 확산확률모델 구조를 기반으로 한다. 학습 데이터는 저해상도 영상, Bicubic 보간을 통해 생성된 초해상도 영상, 그리고 정답 고해상도 영상으로 구성된다. SRDiff 구조에서는 HR 영상을 직접 복원하는 대신, HR 영상과 SR 영상 간의 차이를 잔차로 정의하고 이를 예측하

도록 학습을 수행한다. 모델은 조건 입력으로 제공되는 SR 영상을 기반으로 세부 질감 및 구조 정보를 포함하는 잔차를 복원하며, 최종적으로 고해상도 영상을 생성한다. 입력 및 출력 데이터의 구성은 <Table 2>에 제시하였다.

Table 2. Training Data Setup and Input Specification for High-resolution Restoration

Role	Variable	Description
Input	x_{LR}	Low-resolution image
Conditional Input	x_{SR}	Bicubic-interpolated low-resolution input
Target	$x_r = x_{HR} - x_{SR}$	Residual between GT and SR
Ground Truth	x_{HR}	High-resolution ground truth image
Output	$\hat{x}_{residual}$	Predicted residual output from the model

(3) 확산 기반 학습 원리

학습 과정은 노이즈 예측 기반 확률적 복원(stochastic noise prediction)에 기반한다. 잔차 패치 x_0 에 대해 시점 t 에서 가우시안 노이즈 $\epsilon \sim \mathcal{N}(0, I)$ 를 점진적으로 추가하여 노이즈가 섞인 중간 상태 x_t 를 생성한다. 이 관계는 다음과 같이 정의된다:

$$x_t = \sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \epsilon. \quad (1)$$

여기서 α_t 는 시간 단계 t 에 따른 노이즈 감쇠 계수(noise attenuation factor)이다. 모델 f_θ 는 조건 입력 x_{SR} 과 시점 t 를 함께 받아 추가된 노이즈 ϵ 를 예측한다:

$$\hat{\epsilon} = f_\theta(x_t, x_{SR}, t). \quad (2)$$

손실 함수는 실제 노이즈와 예측 노이즈 간의 절대적 차이를 최소화하는 형태로 정의하였으며, 의료 영상에서 세부 구조와 경계 정보를 안정적으로 보존하는 데 유리하다:

$$\mathcal{L} = \|\epsilon - \hat{\epsilon}\|_1. \quad (3)$$

(4) 역확산 기반 복원 과정

확산확률모델 기반 복원 과정은 학습된 모델을 이용하여 입력 영상의 세부 구조를 점진적으로 복원하는 추론 단계로 구성된다. 학습 과정에서는 잔차에 단계적으로 가우시안 노이즈를 추가하여 노이즈 상태를 생성하고, 모델은 조건 입력으로부터 해당 노이즈를 예측하도록 학습된다. 추론 과정에서는 무작위 노이즈에서 시작하여 학습된 역확산 과정을 반복 수행함으로써 잔차를 점진적으로 복원한다. 최종 고해상도 영상은 복원된 잔차를 조건 입력에 더하여 다음과 같이 계산된다:

$$\hat{x}_{HR} = x_{SR} + \hat{x}_{residual} \quad (4)$$

조건 입력 x_{SR} 에 복원된 잔차를 더하는 방식으로 세밀한 구조 정보를 보완하며, 이 결과는 이후 경계 정제 단계의 입력으로 활용된다.

(5) 요약

기반 복원 단계는 SR3/SRDiff 확산확률모델을 활용하여 128×128 패치 단위의 X-ray 영상을 확률적으로 복원하는 과정으로 구성된다. 모델은 조건 입력을 기반으로 잔차 공간에서 노이즈를 예측하고, 역확산 과정을 통해 고해상도 구조를 복원한다. 생성된 예측 패치는 경계 정제의 입력으로 사용되어, 패치 단위 복원에서 발생하는 경계 불연속 완화에 활용된다.

3.2 Diffusion 기반 Boundary Refinement

본 연구의 확산확률모델 기반 경계 정제 단계는 초해상도 복원 결과를 입력으로 사용하여, 패치 단위 복원 및 병합 과정에서 발생하는 경계 불연속을 완화하도록 설계되었다. 세부 처리 과정은 <Figure 2>의 하단에 제시하였으며, 입력 구성, soft mask 정의, 학습 및 추론 절차는 다음 절에서 상세히 기술한다.

(1) 입력 구성 및 데이터 정의

경계 정제 단계는 기반 복원 단계에서 생성된 고해상도 패치들로부터 경계 중심 64×64 크기의 패치를 추출하여 경계 중심 데이터셋을 구성하였다. 각 경계 부위는 경계 중심에서 외곽으로 갈수록 완만하게 mask를 적용하는 형태로 정의하였으

며, band 폭은 b 로 정의하였다. 점진적 가중치는 이진으로 마스크를 적용하는 것에 비하여 경계의 점진적 변화를 보존하도록 하며 불연속 경계 부근의 분포를 자연스럽게 학습하도록 유도한다(<Table 3>).

Table 3. Training Data Construction and Input Definition for the Diffusion-based Boundary Refinement Stage

Notation	Description
x_0	Ground-truth patch aligned with the boundary center
x_{pred}	Super-resolved patch centered at the boundary
M	Soft mask (Tukey window, band= b)
$x_{cond} = x_{pred} \odot M$	Masked conditional input
$\hat{x}_{refine}$	Diffusion refined output

Diffusion Model의 입력은 시점 t 에서 노이즈가 점진적으로 추가된 영상 x_t 이며, 조건부 입력은 기반 단계의 복원 결과 x_{pred} 에 mask M 을 곱한 형태로 정의된다. 중간 상태는 다음과 같이 구성된다:

$$x_t = \sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \epsilon, x_{cond} = x_{pred} \odot M. \quad (5)$$

여기서 x_0 은 고해상도 정답 패치, α_t 는 cosine β -schedule을 따르는 시간 의존적 감쇠 계수이다. 모델은 x_{cond} 와 x_t 를 입력으로 받아 경계 영역의 세부 구조를 복원하도록 학습된다.

(2) Soft Mask 설계

Soft mask $M(x)$ 는 경계 중심 c 를 기준으로 $\pm b$ 범위에서 0

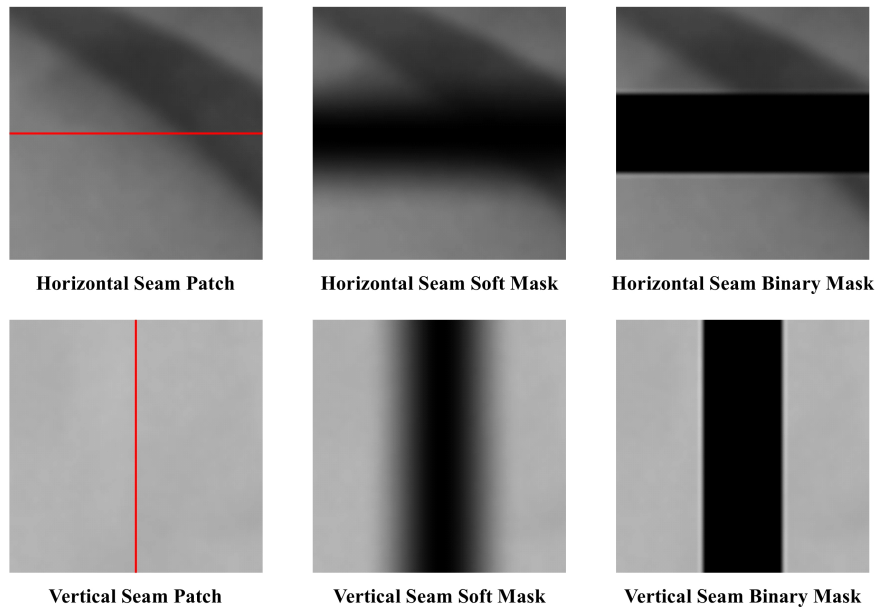


Figure 3. Comparison of the proposed soft mask and binary mask for horizontal and vertical boundary regions

에서 1로 완만히 변화하며, 다음과 같이 정의된다:

$$M(x) = \begin{cases} 1 - \frac{1}{2}(1 + \cos(\pi(\frac{|x-c|}{b}))), & |x-c| \leq b \\ 1, & \text{otherwise} \end{cases} \quad (6)$$

이 설계는 경계 대상 영역을 명확히 구분하고, binary mask 대비 분포 왜곡을 줄여 안정적인 학습을 유도한다. <Figure 3>은 수평 및 수직 경계 영역에 대해 생성된 soft mask와 binary mask의 비교 결과를 보여준다. Binary mask는 경계 영역을 이진으로 구분하여 급격한 전이를 발생시키는 반면, 제안한 soft mask는 경계 중심을 기준으로 점진적으로 변화하는 가중치를 적용하여 자연스러운 분포 전이를 제공한다.

(3) 학습 절차

확산확률모델은 SR3의 노이즈 예측 구조를 따라 각 확산 시점에서 실제 노이즈를 예측하도록 학습된다. 계산 효율성과 경계 보존을 위해 L1 기반 노이즈 예측 손실 함수를 사용하였다:

$$L_{diff}^{L1} = \mathbb{E}_{t, x_0, \epsilon} [\| \epsilon - \hat{\epsilon}_\theta(\sqrt{\alpha_t} x_0 + \sqrt{1-\alpha_t} \epsilon, x_{cond}, t) \|_1]. \quad (7)$$

여기서 $\hat{\epsilon}_\theta(\cdot)$ 는 파라미터 θ 를 갖는 네트워크가 예측한 노이즈이며, 총 1,000단계의 노이즈 스케줄을 cosine β -schedule 기반으로 적용하였다. 학습에는 고해상도 복원 결과에 soft mask M 을 곱한 $x_{cond} = x_{pred} \odot M$ 을 사용하며, 동일한 mask 폭 b 조건하에 수행된다.

(4) 추론 및 Reinsertion 단계

추론 단계에서는 조건 입력 x_{cond} 을 기반으로 학습된 모델이 복원된 패치 $\hat{x}_{refine}$ 를 생성한다. 초기 상태 x_T 는 표준 정규 분포 $\mathcal{N} \sim (0, I)$ 에서 샘플링되며, 모델은 각 시점 $t \in [T, 1]$ 에 대해 다음의 역확산 과정을 반복적으로 수행한다:

$$x_{t-1} = \mu_\theta(x_t, x_{cond}, t) + \sigma_t z, z \sim \mathcal{N}(0, I). \quad (8)$$

해당 구조는 RePaint의 반복 샘플링(resampling) 방식과 유사하나, 본 연구에서는 의료 영상의 구조적 연속성을 고려하여, 정합 대상 영역을 경계부로 확장하였다. 복원된 패치는 다음과 같이 해당 경계 영역에 선택적으로 삽입되어, 원본 영상 내에 통합되며, 별도의 후처리(post-processing) 없이 자연스럽게 일관된 경계 질감 정합을 수행한다.

$$\hat{x}_{final} = x_{merge} \odot M + \hat{x}_{refine} \odot (1 - M). \quad (9)$$

본 연구에서 제안하는 확산 기반 경계 정제 단계는 다음 절차로 구성된다. 먼저 점진적 가중치 함수를 기반으로 경계 중심 영역을 정의하고 해당 부위의 패치를 선택적으로 추출한

다. 이후 조건부 확산확률모델을 활용하여 경계 영역에 대한 확률적 복원을 학습하고, 복원된 경계 패치를 원본 경계선 영상에 병합하여 일관된 결과 영상을 생성한다.

4. 실험

4.1 실험 설정

본 연구에서는 전남대학교 수의과대학에서 수집한 수의학 X-ray 영상 데이터셋을 사용하였다. 데이터셋은 총 98장의 고해상도 영상으로 구성되며, 다양한 해부학적 구조와 명암을 포함한다. 각 영상의 해상도는 1664×1792이다. 실험에서는 원본 고해상도 영상을 정답 영상으로 사용하였으며, 저해상도 영상은 다운샘플링을 통해 생성하였다. 또한 Bicubic 보간을 적용하여 초해상도 입력 영상을 생성하고, 이를 기반으로 SR3 및 SRDiff 모델의 학습과 평가를 수행하였다.

초해상도 복원 단계와 경계 정제 단계로 구성된 2-stage 구조를 사용하며, 두 모델은 end-to-end 방식이 아닌 독립적으로 구성하였다. 먼저 SR3 기반 초해상도 복원 모델을 이용하여 고해상도 예측 영상을 생성한 후, 생성된 결과로부터 경계 중심 데이터셋을 구축하여 경계 정제 모델을 학습하였다. 데이터 분할은 원본 X-ray 영상 단위로 수행하였다. 전체 98장 중 78장은 학습에, 20장은 평가에 사용하였으며, 이후 각 분할에 포함된 영상으로부터 패치를 추출하여 동일 영상의 패치가 학습과 평가에 동시에 포함되지 않도록 구성하였다. 초해상도 복원 단계에서는 52,650개의 학습 패치와 13,500개의 추론 패치를, 경계 정제 단계에서는 22,432개의 학습 패치와 5,608개의 추론 패치를 사용하였다. 두 단계의 모델은 동일한 diffusion 설정을 사용하였으며, 각각 1,000,000 iterations 동안 학습하였다.

실험은 NVIDIA RTX 4090 GPU 환경에서 수행하였으며, cosine 스케줄러(cosine scheduler)를 적용한 1,000 step의 노이즈 주입 과정을 사용하였다. 배치 크기(batch size)는 16으로 설정하였고, SR3 기반의 L1 노이즈 예측 손실 함수를 활용하였다. diffusion 설정, optimizer, learning rate 및 학습 환경은 초해상도 복원 모델과 경계 정제 모델에 동일하게 적용하였다. 모든 실험은 동일한 X-ray 고해상도 정답 데이터를 기준으로 수행하였다. 비교 기준인 Baseline은 SR3 구조에 SRDiff의 잔차 학습 방식을 적용하여 복원한 고해상도 패치를 경계 정제 없이 병합한 결과로 정의하였으며, 제안 방법과 모든 비교 방법은 이 Baseline 결과를 공통 입력으로 사용하였다. 경계 정제 단계에서 사용된 soft mask는 Tukey window 기반의 점진적 가중치 함수로 구성하여 패치 내부 구조는 유지한 채 각 패치의 경계 영역만 재생성하도록 구현하였으며, band 폭 b 는 20 px로 설정하였다. 해당 값은 경계 영역의 연속성을 유지하면서도 복원 대상 영역을 국소적으로 제한하기 위해 경험적으로 결정하였다.

비교 방법으로는 경계 정제 이전의 결과와 함께 전통적인 블렌딩 기반 방법인 Hann window, Laplacian pyramid blending, Poisson image editing을 포함하였다. 최신 확산 기반 국소 복원 기법과의 비교를 위해 RAD-style 및 RePaint-style 방법을 추가하였다. RAD-style 및 RePaint-style 방법은 각각 RAD와 RePaint의 핵심 아이디어를 경계 정제 문제에 적용한 비교 구현으로, 동일 데이터셋과 평가 프로토콜을 사용하여 공정하게 성능을 비교하였다. 이러한 비교 방법은 4.2절 패치 단위 평가, 4.3절 병합 영상 단위 평가, 4.4절 국소 영역 단위 평가에 공통으로 적용하였으며, 각 절에서는 평가 단위와 사용 지표만 실험 목적에 따라 다르게 설정하였다.

정량 평가는 feature similarity index(FSIM), Dice coefficient, 95th-percentile hausdorff distance(HD95), peak signal-to-noise ratio(PSNR), structural similarity index measure(SSIM), learned perceptual image patch similarity(LPIPS)의 6가지 지표로 수행하였으며, 각 지표의 정의는 <Table 4>에 정리하였다. 이 중 Dice coefficient와 HD95는 분할 정답 라벨 없이 정답 영상과 각 방법 결과 영상에 동일한 Sobel 기반 절차로 생성한 이진 경계 마스크를 비교하여, 경계 후보 마스크 간의 중첩 정도와 위치 차이를 측정한다. 경계 마스크는 정규화 영상에 Sobel 필터를 적용해 경사 크기를 구한 뒤 이진화하여 생성하였으며, 패치 및 병합 영상 평가에서는 임계값 0.05와 dilation 반경 3을, 국소 단위 평가에서는 경계 관심 영역 내 Sobel 경사 분포 상위 25%에 해당하는 값과 dilation 반경 1을 적용하였다. dilation은 1 - 2 픽셀 수준의 미세한 위치 차이로 인한 과도한 불일치를 완화하기 위한 허용 범위이다. 정성 평가는 보정 변화 지도

(calibration map), 경계 정합성 지도(edge alignment map), 정답 영상 기반 절대 차이 지도를 이용하였다(<Table 4>).

4.2 패치 단위 평가

본 평가는 경계 중심 패치를 대상으로 제안 방법이 경계 부위의 구조적 정합성과 시각적 품질에 미치는 영향을 분석하기 위해 수행되었다. 평가는 경계 정제 이전 결과(Baseline)와 제안 방법의 차이를 중심으로 하되, Hann window, Laplacian pyramid blending, Poisson image editing, RAD-style 및 RePaint-style과의 비교를 통해 상대적 성능도 함께 분석하였다.

정량 평가는 세 가지 유형의 지표로 구성하였다. 구조적 정합도는 Sobel edge 기반 이진 경계 마스크의 중첩성과 위치 차이를 통해 Dice coefficient와 HD95로 평가하였고, 시각적 유사성은 LPIPS로, 위상 및 구조 기반 유사성은 FSIM으로 측정하였다. 제안 방법과 정제 이전 결과 간의 통계적 유의성은 one-sided Mann - Whitney U 검정을 통해 평가하였다.

(1) 정량적 분석 결과

제안 기법은 패치 단위 평가에서 구조적·시각적 품질 지표를 중심으로 개선을 나타냈다(<Table 5>).

<Table 5>의 결과와 같이, 제안 방법은 FSIM, Dice coefficient, HD95 및 LPIPS에서 가장 우수한 성능을 보였다. FSIM은 0.942에서 0.949로 증가하였으며, 모든 비교 방법 중 가장 높은 값을 나타냈다. 이는 제안 방법이 경계 중심 패치의 구조적 및 위상학적 특성을 효과적으로 복원하였음을 의미한다.

Table 4. Overview of the Evaluation Metrics and Their Key Characteristics

Metric	Description
feature similarity index (FSIM)	Structural and perceptual similarity metric based on phase congruency and gradient information
Dice coefficient	Spatial overlap ratio between Sobel edge-derived binary boundary masks from GT and reconstructed images
95th-percentile hausdorff distance (HD95)	95th percentile Hausdorff distance between Sobel edge-derived binary boundary masks from GT and reconstructed images
peak signal-to-noise ratio (PSNR)	Signal-to-noise ratio based on mean squared error (MSE)
structural similarity index measure (SSIM)	Structural similarity based on luminance, contrast, and structure information
learned perceptual image patch similarity (LPIPS)	Deep feature-based perceptual similarity metric aligned with human visual perception

Table 5. Patch-level quantitative comparison of boundary refinement methods. Statistical significance was evaluated using a one-sided Mann - Whitney U test

Metric	Baseline	Hann	Laplacian	Poisson	RAD-style	RePaint-style	Proposed	p-value
FSIM(↑)	0.942 ± 0.015	0.940 ± 0.016	0.937 ± 0.016	0.935 ± 0.017	0.945 ± 0.012	0.946 ± 0.012	0.949 ± 0.012	1.41 × 10 ⁻⁴ (***)
Dice(↑)	0.104 ± 0.243	0.093 ± 0.234	0.090 ± 0.228	0.094 ± 0.228	0.100 ± 0.243	0.109 ± 0.255	0.144 ± 0.312	2.86 × 10 ⁻¹ (-)
HD95(↓)	4.413 ± 11.870	4.139 ± 11.221	4.531 ± 12.258	5.335 ± 13.548	3.671 ± 10.109	3.601 ± 10.030	2.070 ± 5.950	5.71 × 10 ⁻¹ (-)
LPIPS(↓)	0.214 ± 0.042	0.214 ± 0.042	0.216 ± 0.043	0.221 ± 0.042	0.218 ± 0.042	0.206 ± 0.040	0.189 ± 0.041	3.72 × 10 ⁻⁵ (***)

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Dice coefficient는 정제 이전 0.104에서 정제 이후 0.144로 향상되었으나, 통계적으로 유의한 차이는 확인되지 않았다. 이는 Sobel edge 기반 이진 경계 마스크의 중첩성이 패치 단위에서 다소 개선되는 경향을 보였음을 시사한다.

HD95는 제안 방법이 4.413에서 2.070으로 크게 감소하였으며, Hann, Laplacian, Poisson, RAD-style 및 RePaint-style을 포함한 모든 비교 방법 중 가장 낮은 값을 기록하였다. 그러나 통계적으로 유의한 차이는 확인되지 않았기에 패치 단위에서 Sobel edge 기반 경계 마스크의 위치 오차가 감소하는 경향이 관찰되었다고 해석하는 것이 적절하다.

지각적 유사성을 나타내는 LPIPS 역시 정제 이전 0.214에서 정제 이후 0.189로 감소하였으며, 통계적으로 유의한 향상을 나타냈다. RePaint-style에서도 감소가 확인되었으나, 제안 방법이 모든 비교 방법 중 가장 낮은 값을 보였다. 이는 제안 방법이 경계 부위의 시각적 자연성과 지각적 유사성 향상에 기여했을 가능성을 시사한다.

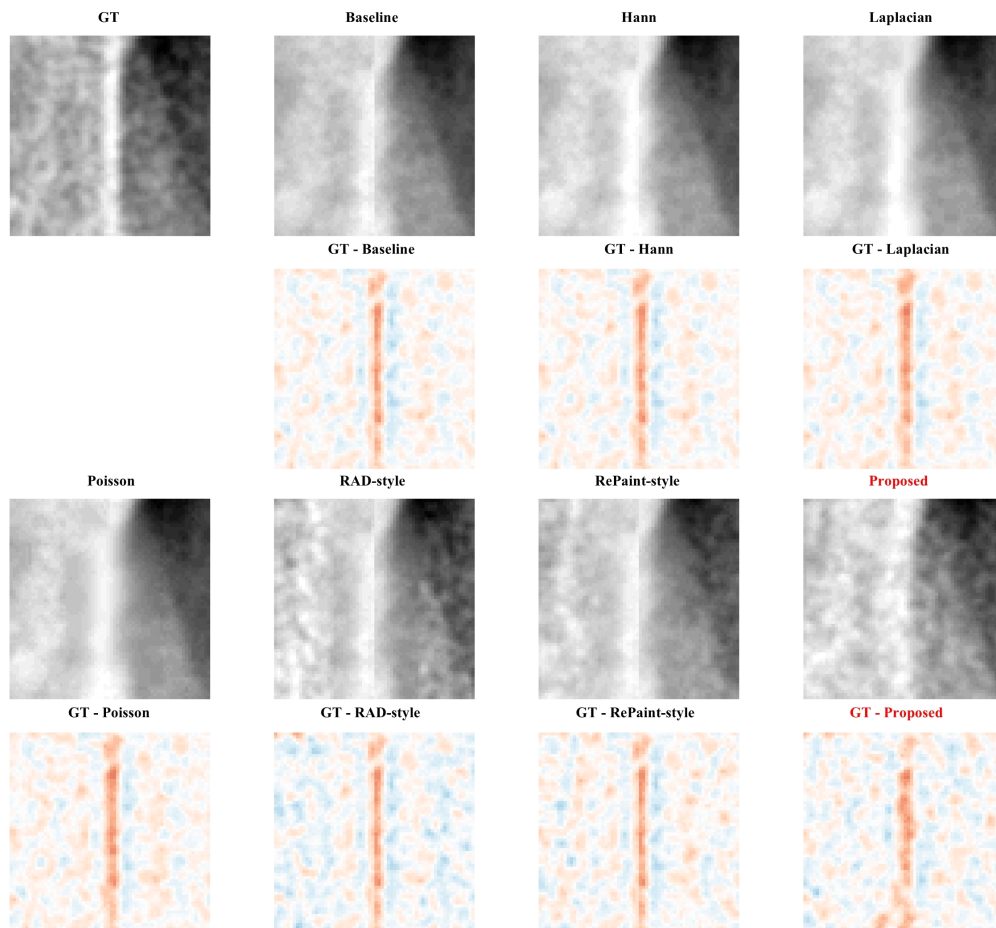
종합적으로, 제안한 확산 기반 경계 정제 기법은 패치 단위 평가에서 FSIM과 LPIPS를 유의하게 개선하였으며, Dice co-

efficient와 HD95 평균값에서도 개선 경향을 보였다. 이러한 결과는 제안 방법이 단순한 경계 이음새 완화가 아니라, 경계 중심 패치의 구조적 품질과 지각적 유사성을 함께 향상시키는 방향으로 작동함을 시사한다.

(2) 정성적 분석 결과

정성적 분석 결과 역시 정량 평가 결과와 유사한 경향을 나타냈다. <Figure 4>는 FSIM 향상이 크게 나타난 대표 패치 사례를 보여준다. 정제 이전 결과에서는 경계 중심 영역에서 밝기 단절과 경계선 흔적이 관찰되었으나, 제안 방법은 정답 영상에 가까운 명암 분포를 복원하였다. 보정 변화 지도에서는 경계 중심 영역에서 발생하는 오차가 비교 방법들 대비 감소하는 양상이 확인되며, 구조적 정합성이 향상되었음을 보여준다.

<Figure 5>는 Dice coefficient 향상이 크게 나타난 패치 사례를 나타낸다. 제안 방법은 정답 영상에서 Sobel 기반으로 추출한 경계와 더 유사한 구조를 복원하였다. 특히 정답 영상과의 오차가 경계 중심부에서 감소하는 양상이 관찰되며, 공간적 정합성이 개선되었음을 확인할 수 있다. 다만 Dice coefficient

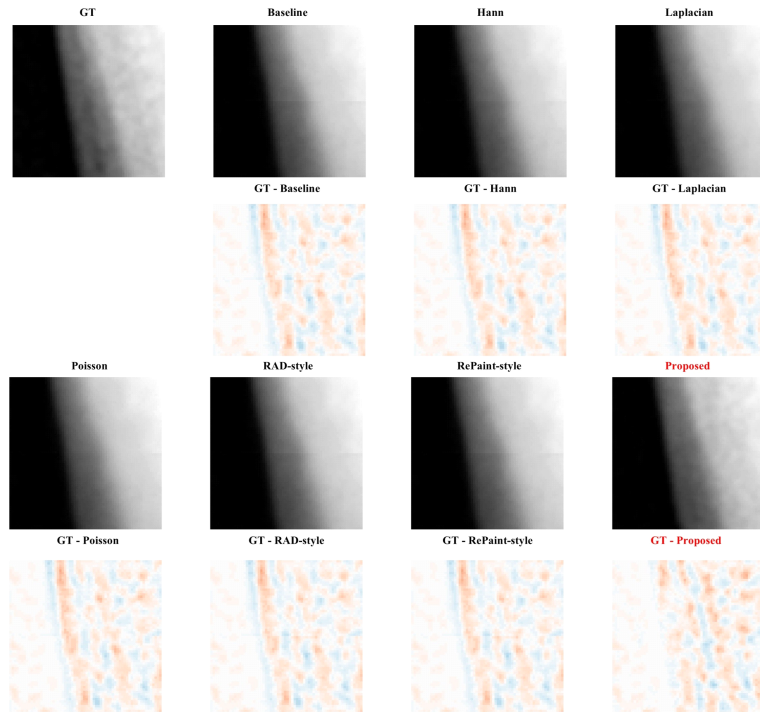


Ground Truth (GT), Baseline, Hann, Laplacian, Poisson, RAD-style, RePaint-style, and Proposed results with corresponding difference maps (GT=Method). Red indicates regions reconstructed darker than GT, Blue indicates regions reconstructed brighter than GT

Figure 4. Representative Patch-level Qualitative Results in Terms of FSIM

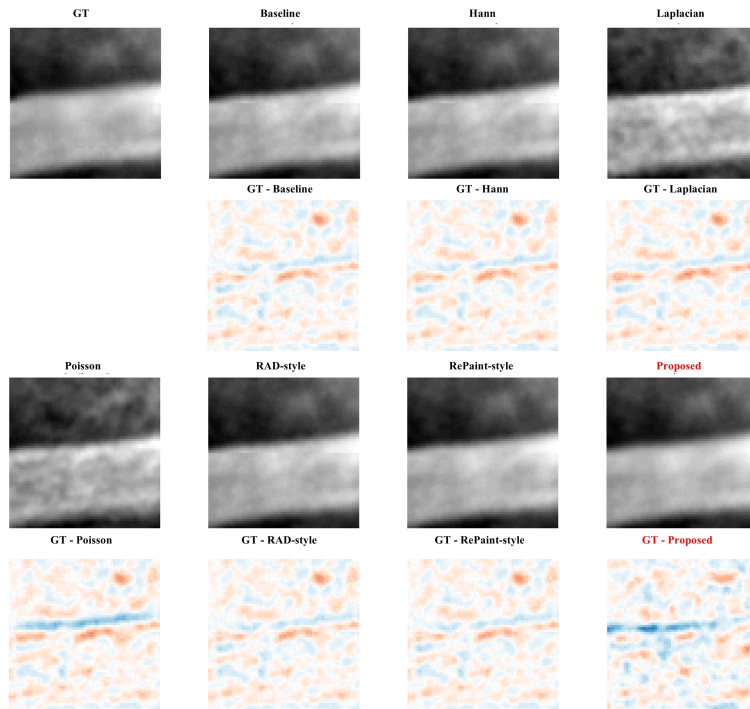
는 이진 경계 마스크 기반 지표이므로, 일부 사례에서는 다른 비교 방법과 유사한 결과가 나타났다.

<Figure 6>의 결과를 보면, 정제 이전 결과 및 전통적 블렌딩 방법에서는 경계선을 따라 잔존하는 구조 왜곡이 관찰되지만



Ground Truth (GT), Baseline, Hann, Laplacian, Poisson, RAD-style, RePaint-style, and Proposed results with corresponding difference maps (GT-Method). Red indicates regions reconstructed darker than GT, Blue indicates regions reconstructed brighter than GT

Figure 5. Representative Patch-level Qualitative Results in Terms of the Dice Coefficient



Ground Truth (GT), Baseline, Hann, Laplacian, Poisson, RAD-style, RePaint-style, and Proposed results with corresponding difference maps (GT-Method). Red indicates regions reconstructed darker than GT, Blue indicates regions reconstructed brighter than GT

Figure 6. Representative Patch-level Qualitative Results in Terms of HD95

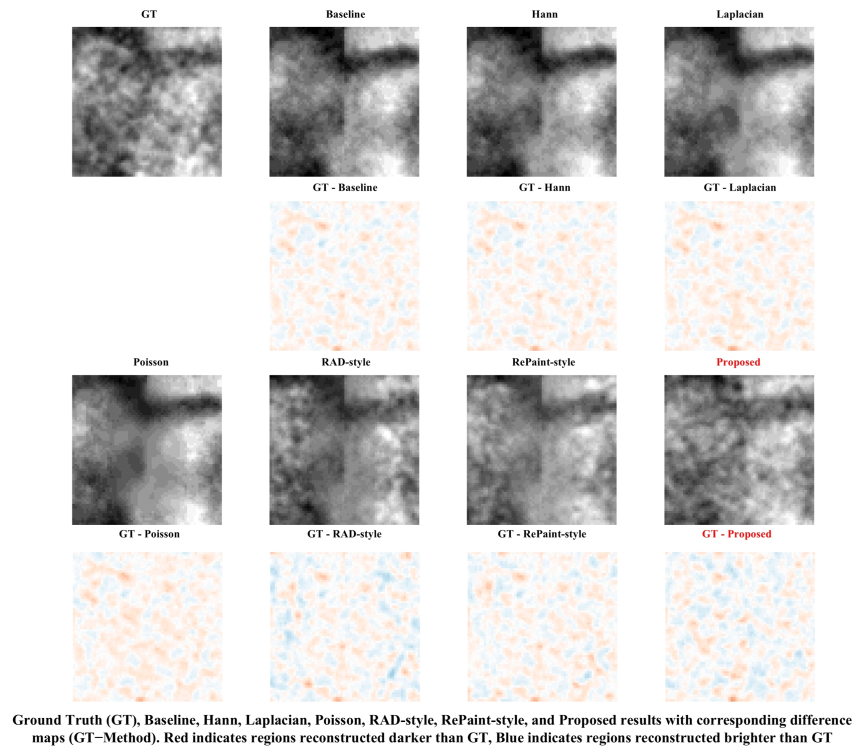


Figure 7. Representative Patch-level Qualitative Results in Terms of LPIPS

제안 방법은 정답 영상에서 Sobel 기반으로 추출한 경계 후보와 더 유사한 구조를 복원하였다. 보정 변화 지도 비교 결과에서도 경계 부근 오차가 감소하는 경향이 확인되며 이는 HD95 감소 결과와 일치한다.

또한 <Figure 7>에서는, LPIPS가 감소한 패치들에서는 정제 이전의 경계선이 배경과 명확히 구분되는 형태로 나타났던 반면, 제안 방법 적용 후에는 경계 부위가 주변 조직과 자연스럽게 연결되면서 시각적 이질감이 감소하는 경향을 확인할 수 있었다. Hann, Laplacian 및 Poisson 방법은 경계선 부근의 밝기 불연속을 완화하는 효과를 보였으나, 미세 질감 복원에는 한계가 존재하였다. 반면 제안 방법은 정답 영상과 유사한 질감 패턴을 복원하며 가장 낮은 LPIPS 값을 달성하였다. RAD-style 및 RePaint-style과 비교하더라도 안정적인 질감 복원이 이루어졌음이 확인되었다. 이러한 결과는 제안 방법이 단순히 경계선을 흐리게 만드는 방식이 아닌 경계 부위의 구조와 질감을 동시에 복원하는 확산 기반 경계 정제를 수행함을 보여준다.

4.3 병합된 이미지 단위 평가

본 평가는 패치 단위 복원 결과를 단일 X-ray 영상으로 병합하는 과정에서 발생하는 경계 불연속이 실제 영상 수준에서 어떻게 개선되는지 검증하기 위해 수행되었다. 이는 제안 기법이 단일 패치 수준의 국소적인 개선에 그치지 않고 전체 영상 단위에서 해부학적 구조의 연속성과 시각적 자연성을 실질적으로 향상시키는지 확인하기 위한 핵심 실험이다. 비교 방

법은 4.2절과 동일하게 구성하였으며, 동일한 입력 영상과 평가 프로토콜을 사용하여 비교하였다.

정량 평가는 세 가지 유형의 지표로 구성하였다. 구조적 경계 정합도는 Sobel edge 기반 이진 경계 마스크의 중첩성과 위치 차이를 통해 Dice coefficient와 HD95로 평가하였고, 전역 화질은 PSNR과 SSIM으로, 시각적 유사성은 LPIPS로 측정하였다. 이를 통해 제안 기법이 병합 과정에서 발생하는 경계 왜곡을 효과적으로 완화하고, 전역 구조적 안정성 및 시각적 품질을 동시에 향상시키는지를 규명하고자 한다.

(1) 정량적 분석 결과

제안한 경계 정제 기법은 병합된 영상 수준에서도 해부학적 경계 기반 지표에서 뚜렷한 성능 향상을 나타냈다(<Table 6>).

<Table 6>의 결과와 같이, Dice coefficient는 정제 이전 0.683에서 정제 이후 0.777로 증가하였으며, HD95는 41.237 px에서 22.150 px로 감소하여, 병합 과정에서 발생하던 구조적 경계 오차가 실질적으로 완화되었음이 확인되었다.

반면, PSNR과 SSIM은 각각 40.328 dB에서 39.053 dB, 0.943에서 0.934로 소폭 감소하였는데, 이는 경계 영역의 고주파 정보와 구조적 연속성을 복원하는 과정에서 발생하는 일반적인 지표 간 절충 현상으로 해석할 수 있다. 평균 기반의 전역 화질 지표는 경계 부근의 국소적 구조 복원을 오차 증가로 반영할 수 있기 때문이다.

LPIPS는 0.307에서 0.303으로 감소하여 제안 방법이 시각적 특성 공간(perceptual feature space)에서 정답 영상에 근접한 결

Table 6. Quantitative comparison of boundary refinement methods on merged images. Statistical significance was assessed using a one-sided Mann - Whitney U test

Metric	Baseline	Hann	Laplacian	Poisson	RAD-style	RePaint-style	Proposed	p-value
Dice(↑)	0.683 ± 0.022	0.683 ± 0.021	0.677 ± 0.022	0.680 ± 0.024	0.663 ± 0.027	0.679 ± 0.022	0.777 ± 0.015	1.43×10^{-2} (*)
HD95(↓)	41.237 ± 5.747	39.476 ± 5.044	39.456 ± 5.284	39.841 ± 6.850	42.757 ± 5.222	40.834 ± 5.905	22.150 ± 4.781	1.43×10^{-2} (*)
PSNR(↑)	40.328 ± 0.443	40.203 ± 0.430	39.894 ± 0.399	39.646 ± 0.391	39.104 ± 0.328	40.218 ± 0.423	39.053 ± 0.615	1.00×10^0 (-)
SSIM(↑)	0.943 ± 0.001	0.942 ± 0.001	0.940 ± 0.001	0.939 ± 0.001	0.924 ± 0.000	0.941 ± 0.001	0.934 ± 0.001	1.00×10^0 (-)
LPIPS(↓)	0.307 ± 0.003	0.311 ± 0.003	0.316 ± 0.003	0.318 ± 0.003	0.335 ± 0.003	0.312 ± 0.003	0.303 ± 0.003	5.71×10^{-2} (-)

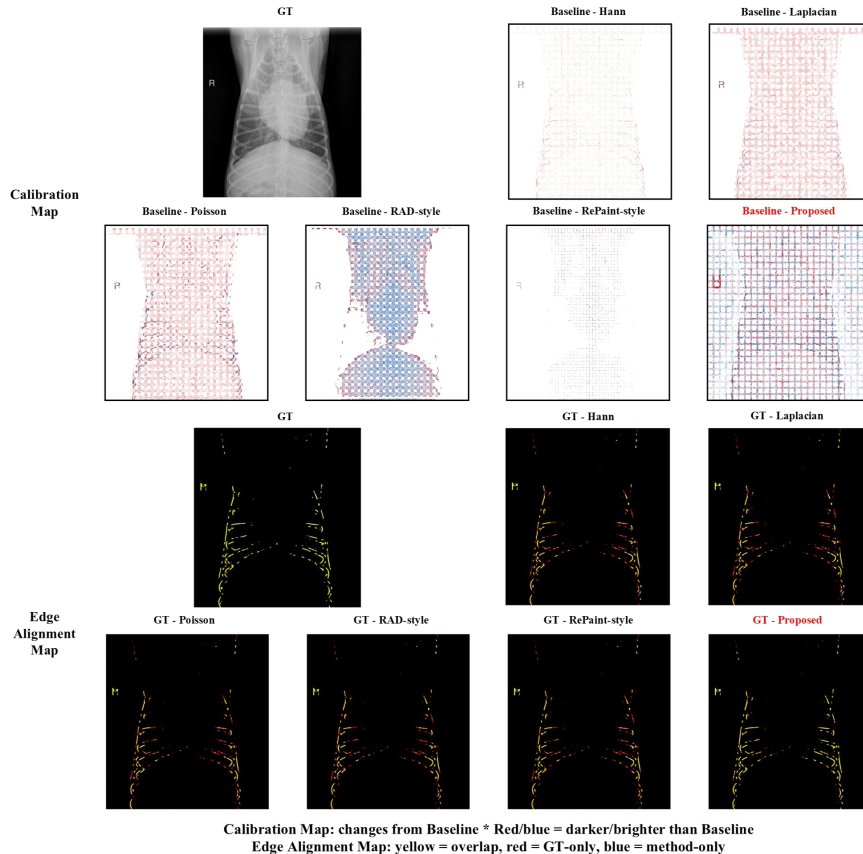
* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

과를 생성하는 경향을 보였다. 이는 제안 기법이 경계 영역의 구조적 불연속을 효과적으로 제거하는 동시에, 전체 영상의 지각적 품질을 유지하거나 개선하는 방향으로 작동함을 시사한다. 전역 화질 지표의 미세한 감소는 경계의 연속성과 구조 보존을 우선시한 복원에 기인한 자연스러운 균형 조정으로 해석할 수 있다.

(2) 정성적 분석 결과

시각적 평가 결과는 정량 분석과 유사한 경향을 보였으며, 병합 영상은 보정 변화 지도와 경계 정합성 지도를 이용하여 분석하였다. 보정 변화 지도는 정제 이전 영상을 기준으로 각 방법의 보정 방향과 변화 정도를 나타내며, 빨간색과 파란색

은 각각 기준 영상보다 어둡거나 밝아진 영역을 의미한다. 이는 정답 영상과의 오차가 아닌 각 방법의 보정 양상을 보여준다. 경계 정합성 지도는 Sobel 기반 경계 비교 결과로 노란색은 공통 경계, 빨간색은 정답 영상만, 파란색은 방법 결과만의 경계를 의미한다. <Figure 8>에서 확인할 수 있듯, Dice coefficient의 향상은 경계의 위치 정합도가 개선된 영역에서 두드러졌다. 보정 변화 지도에서는 비교 방법별로 기준 영상 대비 보정 양상이 다르게 나타나며, 전통적 블렌딩 기법은 경계 부근의 밝기 변화를 완화하는 형태로 보정한다. 제안 방법은 병합 경계 영역 전반에 걸쳐 더 뚜렷한 보정 양상을 보이며, 경계 정합성 지도에서는 정답 영상 경계와 겹치는 영역이 유지되거나 증가하는 경향을 나타낸다. 이는 Dice coefficient가

**Figure 8.** Calibration Map and Edge Alignment Map Qualitative Results in Terms of the Dice Coefficient

Sobel 기반 이진 경계 마스크의 중첩 정도를 평가하는 지표라는 점에서 경계 구조의 공간적 정합성이 개선된 결과로 해석할 수 있다.

<Figure 9>는 HD95가 크게 감소한 병합 영상 사례이다. 보정 변화 지도에서는 제안 방법의 변화가 경계 부근과 패치 격자 영역에 집중되어 나타났다. 경계 정합성 지도는 제안 방법

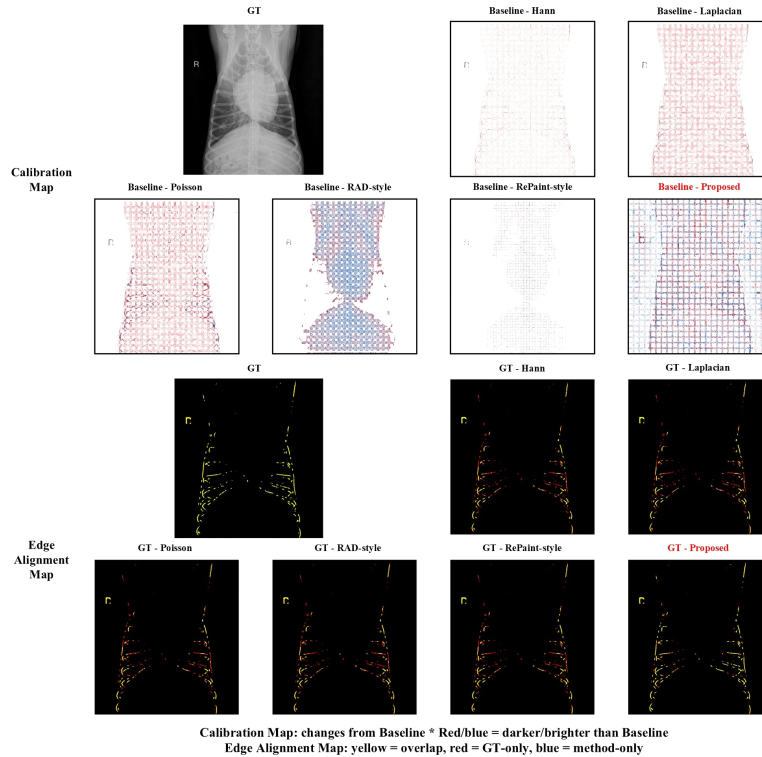


Figure 9. Calibration Map and Edge Alignment Map Qualitative Results in Terms of HD95

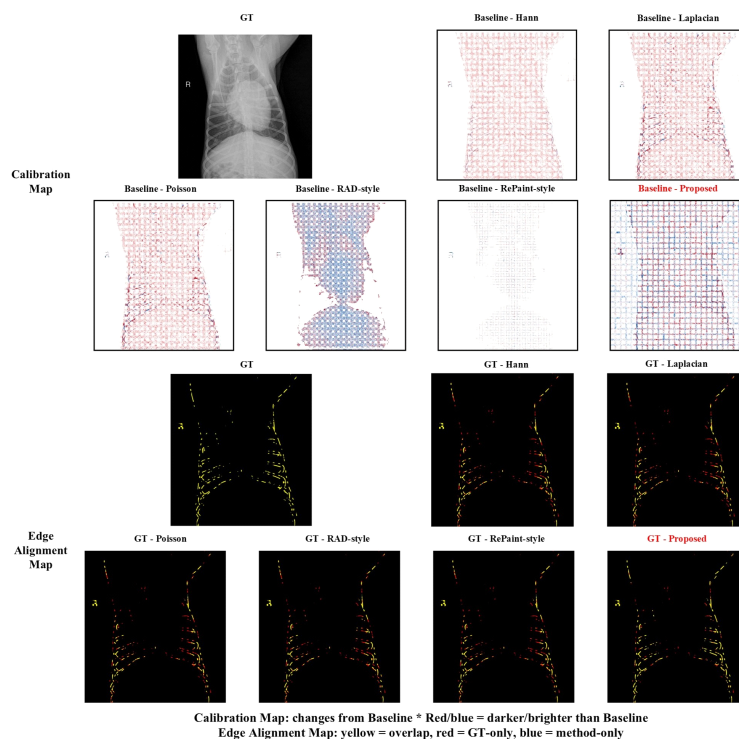


Figure 10. Calibration Map and Edge Alignment Map Qualitative Results in Terms of PSNR

이 정답 영상과의 경계 정합성을 향상시키는 경향을 보였으며, 이는 HD95 감소 결과와 일관된다.

<Figure 10>과 <Figure 11>에서는 PSNR 및 SSIM이 소폭 감

소한 사례로 정제 이후 영상이 고주파 조직 경계를 더 선명하게 재현하였으며 구조적 정합도는 개선되었으나 픽셀 단위 평균 오차가 증가한 결과로 해석된다. PSNR과 SSIM은 영상 전

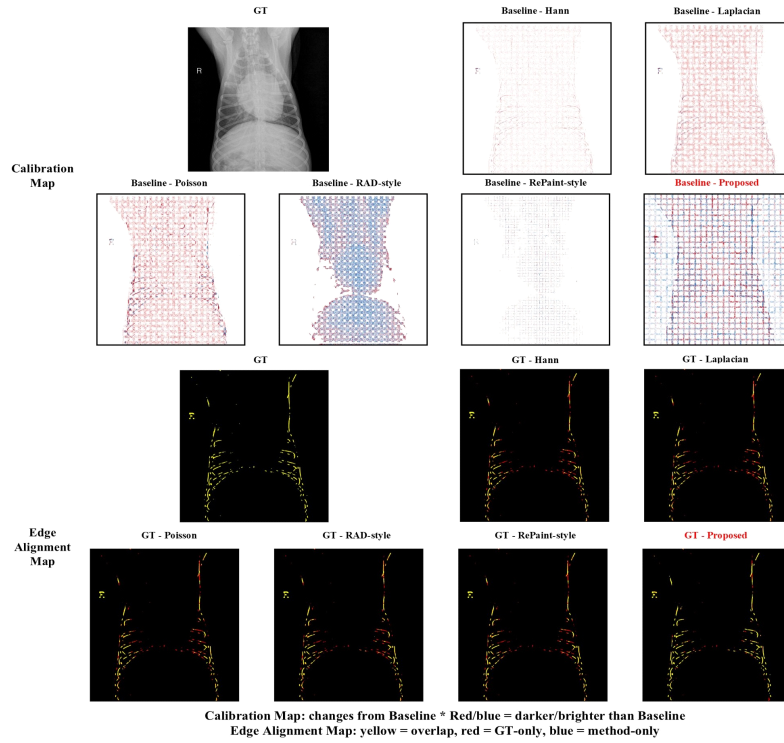


Figure 11. Calibration Map and Edge Alignment Map Qualitative Results in Terms of SSIM

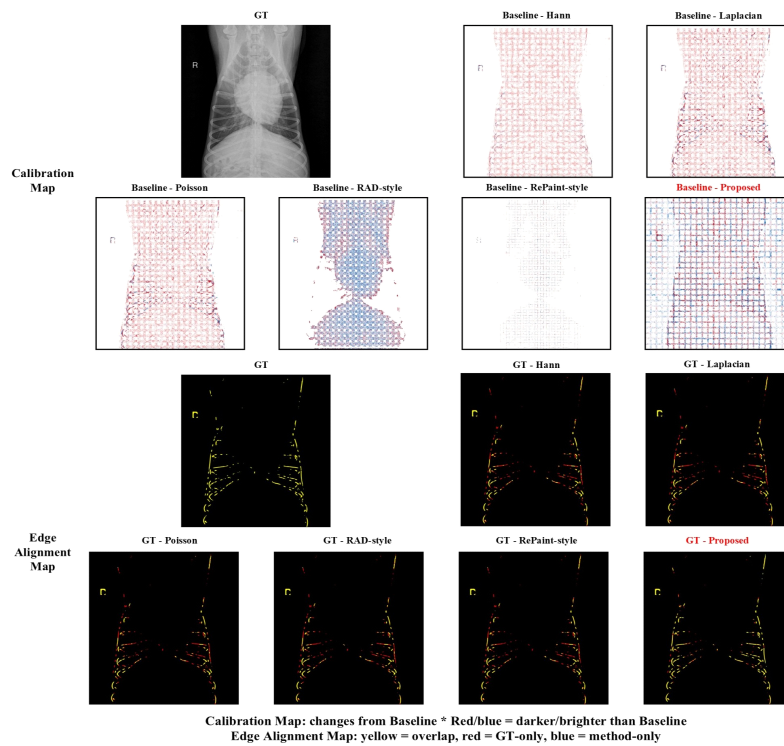


Figure 12. Calibration Map and Edge Alignment Map Qualitative Results in Terms of LPIPS

체의 화소 단위 유사도와 구조적 유사도를 기반으로 계산되기에, 경계 부근의 국소적 보정이나 고주파 성분의 변화가 전역 지표에서 오차로 반영될 수 있다. 보정 변화 지도에서는 제안 방법이 기존 영상 대비 경계 영역을 중심으로 명확한 보정을 수행한 양상이 나타난다. 이는 일부 화소 기반 지표의 감소로 이어질 수 있으나, 경계 정합성 지도에서는 정답 영상 경계와의 중첩이 유지되며 해부학적 경계 구조의 정합성이 보존되는 경향이 나타난다. 따라서 PSNR 및 SSIM의 변화는 제안 방법이 전역 화소 오차 최소화보다 경계 구조 복원에 중점을 두기에 나타나는 지표 간 절충 현상으로 해석할 수 있다.

또한 <Figure 12>에서는 LPIPS가 개선된 병합 영상 사례가 나타난다. 보정 변화 지도에서는 제안 방법이 정제 이전 영상 대비 경계 및 격자 영역을 중심으로 보정을 수행한 양상이 확인되며, 경계 부위의 시각적 이질감 완화를 위한 변화로 해석할 수 있다. 경계 정합성 지도에서도 제안 방법은 정답 영상 경계와의 중첩을 유지하며 불필요한 경계 차이를 줄이는 경향을 보인다. 이는 제안 방법이 단순히 경계선을 흐리게 만드는 것이 아닌 경계 주변의 구조적 연속성과 시각적 자연스러움을 함께 개선함을 보여준다.

(3) 정답 영상 기반 차이 분석

정답 영상 기반 차이 분석은 기존 영상 대비 보정량을 나타내는 보정 변화 지도와 달리, 정답 영상과 각 방법 결과 영상 사이의 절대 차이를 시각화하여 병합 영상에 남아 있는 오차 분포를 확인하기 위해 수행하였다.

<Figure 13>은 LPIPS가 개선된 사례의 절대 차이 결과를 보여준다. 첫 번째 영상은 정답 영상을 나타내며, 나머지는 정답 영상과 각 방법 결과 간 절대 차이 지도($|GT - Method|$)로, 밝은 영역일수록 잔여 오차가 크다. 전반적으로 잔여 오차는 해부학적 구조 경계와 고주파 질감 영역에 집중되었고, 평탄한

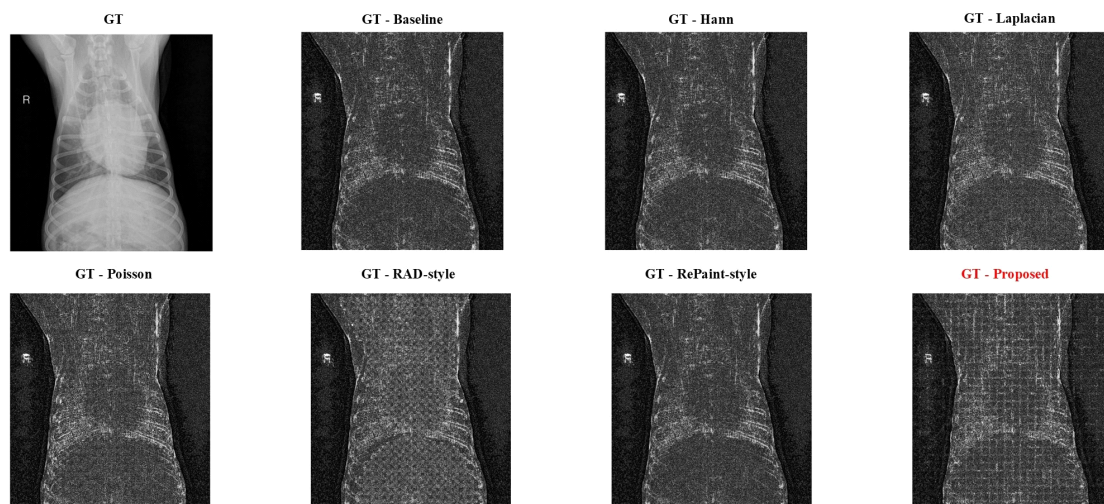
배경에서는 낮게 관찰되었다. 정제 이전 영상과 전통적 블렌딩 방법은 경계 및 질감 영역을 따라 오차를 보였으며, RAD-style은 질감 차이가 비교적 넓게 나타났다. RePaint-style은 정제 이전 영상과 유사한 분포를 보였는데, 이는 기존 영상을 거의 변화시키지 않아 경계 보정이 충분치 않았음을 의미한다.

제안 방법의 절대 차이 지도에서는 경계 정제가 적용된 위치를 따라 패치 격자 형태의 잔여 오차가 확인되었다. 이는 경계 부근에 재구성된 질감이 인접 영역과 달리 나타나는 현상으로, 경계 불연속을 완화하더라도 일부 픽셀 단위 오차가 남을 수 있음을 의미한다. 그러나 절대 차이 지도는 픽셀 오차를 반영할 뿐이며, 경계 위치 정합성을 나타내는 Dice coefficient와 HD95에서는 제안 방법이 가장 우수한 성능을 보였다.

4.4 병합된 이미지 일부 영역 단위 평가

본 평가는 병합 영상 내 실제 패치 경계 인접 부위에서 발생하는 불연속 현상을 국소 수준에서 분석하기 위해 수행되었다. 64×64 패치 3×3 에 해당하는 192×192 크기의 블록을 stride 128로 슬라이딩 추출하여 평가 영역을 구성하고, 각 블록의 중심을 기준으로 4.2절과 동일한 비교 방법의 결과를 동일 영역에서 비교하였다. 또한 Tukey 기반 병합 마스크로 블록 내 실제 경계선이 포함된 관심 영역을 선별하여 경계 정제 효과를 평가하였다.

정량 평가는 세 가지 항목으로 수행하였다. 경계 위치 정합도는 관심 영역 내 Sobel edge 기반 이진 경계 마스크의 위치 차이를 통해 HD95로, 영상 품질은 PSNR과 SSIM으로, 시각적 유사도는 LPIPS로 측정하였다. 또한 지표 변화가 큰 사례에 대해서는 보정 변화 지도를 생성하여 정답 영상과의 차이를 시각적으로 분석하였다.



GT image and absolute difference maps, $|GT - Method|$
Brighter pixels indicate larger residual errors relative to GT.

Figure 13. Merged-image GT-based absolute difference map results

(1) 정량적 분석 결과

국소 영역 단위 평가에서도 전체 영상 평가와 유사한 경향이 관찰되었으며, 특히 경계 기반 정합 지표에서 개선 폭이 두드러졌다(<Table 7>).

<Table 7>의 결과와 같이, HD95는 정제 이전 5.288 px에서 정제 이후 4.502 px로 감소하였으며, 통계적으로 유의한 차이를 보였다. 이는 제안 방법이 패치 블록 내부의 경계 영역에서 Sobel edge 기반 경계 마스크의 위치 차이를 줄이는 데 효과적임을 의미한다. LPIPS 또한 0.216에서 0.205로 낮아졌으며, 통계적으로 유의한 향상을 나타냈다. 제안 방법이 국소 경계 영역의 지각적 품질을 개선하는 경향을 보였음을 의미한다.

반면, PSNR과 SSIM은 소폭 감소하였으며, 통계적으로 유의한 향상은 확인되지 않았다. 이는 경계 부근의 구조와 질감을 보정하는 과정에서 일부 픽셀 단위 오차가 증가할 수 있음을 보여준다. 결과적으로 제안 방법은 국소 영역 평가에서도

HD95와 LPIPS에서 가장 우수한 성능을 기록하여 전역 픽셀 기반 지표보다 경계 위치 정확성과 지각적 품질 측면에서 더 뚜렷한 개선을 보였다.

(2) 정성적 분석 결과

정성적 분석에서는 국소 영역의 경계 정합 및 지각적 품질 변화를 시각적으로 확인하였다. <Figure 14>는 HD95가 개선된 국소 영역 사례를 보여준다. 정제 이전 영상에서는 패치 경계가 뚜렷하게 드러났으나, 정제 이후 해당 경계가 제거되고 정답 영상의 윤곽을 따르도록 정렬되었다. 정제 전후로 경계선 중심의 오차 대역이 축소되고 오차 분포가 균일하게 나타났다.

고주파 구조가 포함된 영역에서는 정제 이후 영상이 경계 부근의 미세 질감과 윤곽을 명확하게 복원하는 경향을 보였다. <Figure 15>와 <Figure 16>에서 확인할 수 있듯이, 일부 사

Table 7. Localized quantitative comparison of boundary refinement methods in merged-image regions. Statistical significance was evaluated using a one-sided Mann - Whitney U test

Metric	Baseline	Hann	Laplacian	Poisson	RAD-style	RePaint-style	Proposed	p-value
HD95(↓)	5.288 ± 4.337	5.023 ± 4.135	4.903 ± 4.159	4.999 ± 4.553	5.327 ± 4.729	5.182 ± 4.230	4.502 ± 1.923	3.33×10^{-3} (**)
PSNR(↑)	41.598 ± 3.270	41.462 ± 3.304	41.138 ± 3.396	40.874 ± 3.483	40.336 ± 3.694	41.482 ± 3.292	40.553 ± 3.352	1.00×10^0 (-)
SSIM(↑)	0.942 ± 0.012	0.941 ± 0.012	0.939 ± 0.013	0.937 ± 0.014	0.922 ± 0.023	0.940 ± 0.012	0.932 ± 0.013	1.00×10^0 (-)
LPIPS(↓)	0.216 ± 0.140	0.226 ± 0.148	0.236 ± 0.151	0.229 ± 0.149	0.236 ± 0.155	0.224 ± 0.149	0.205 ± 0.143	2.17×10^{-2} (*)

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

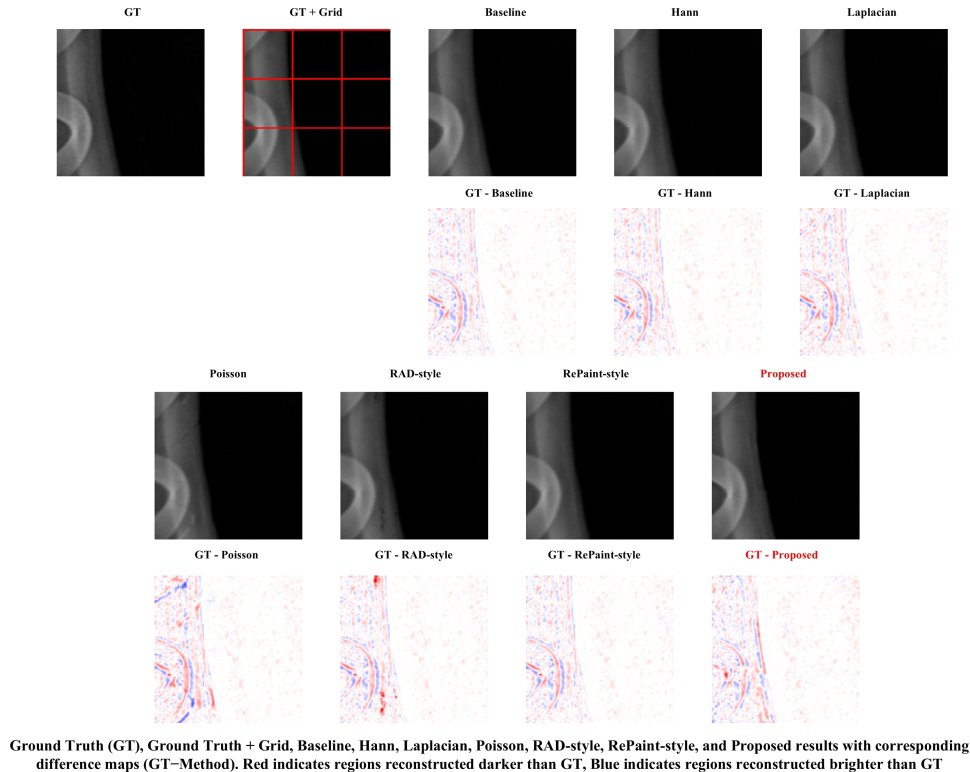
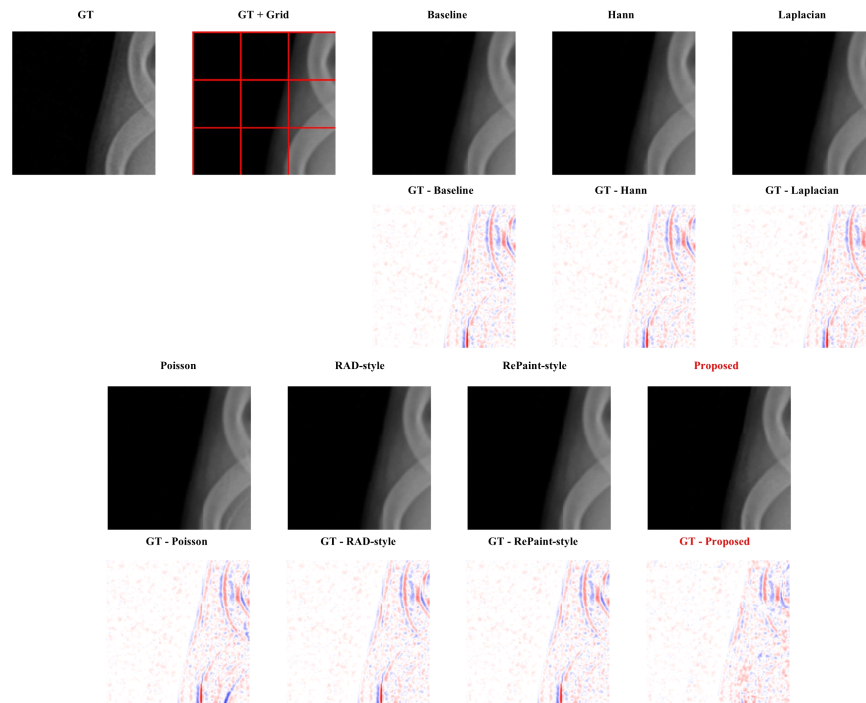


Figure 14. Representative Localized-region Qualitative Results in Terms of HD95

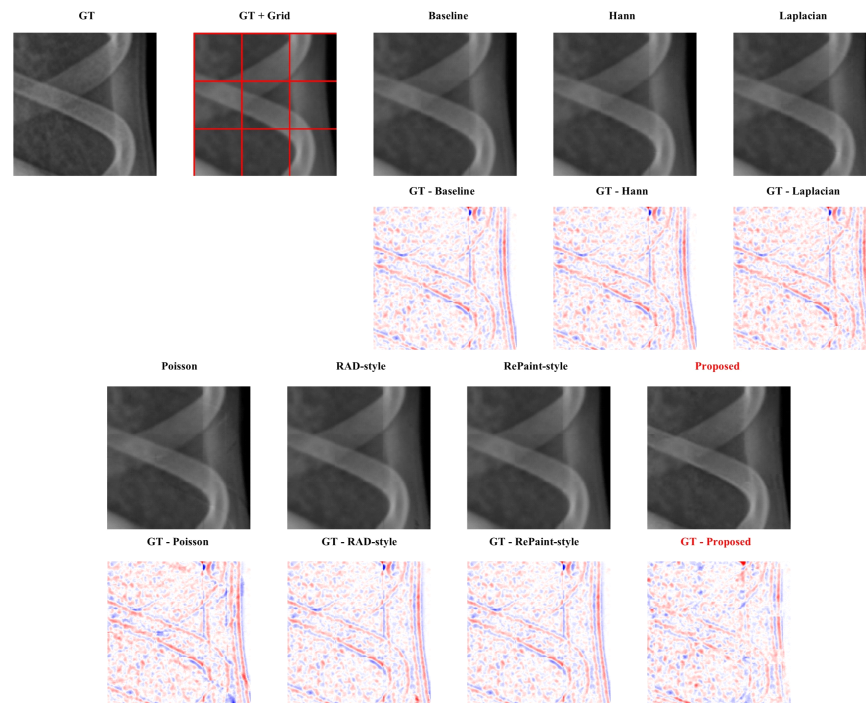
레에서는 PSNR과 SSIM이 소폭 감소하였으나, 경계 부근의 구조적 정합성은 향상되는 양상이 나타났다. 보정 변화 지도에

서는 제안 방법이 정제 이전 결과 대비 경계 인접 영역을 중심으로 영상을 보정한 양상이 확인된다. 이러한 보정은 전역적



Ground Truth (GT), Ground Truth + Grid, Baseline, Hann, Laplacian, Poisson, RAD-style, RePaint-style, and Proposed results with corresponding difference maps (GT-Method). Red indicates regions reconstructed darker than GT, Blue indicates regions reconstructed brighter than GT

Figure 15. Representative Localized-region Qualitative Results in Terms of PSNR



Ground Truth (GT), Ground Truth + Grid, Baseline, Hann, Laplacian, Poisson, RAD-style, RePaint-style, and Proposed results with corresponding difference maps (GT-Method). Red indicates regions reconstructed darker than GT, Blue indicates regions reconstructed brighter than GT

Figure 16. Representative Localized-region Qualitative Results in Terms of SSIM

인 픽셀 단위 유사도 지표에서 일부 감소로 반영될 수 있으나, HD95 감소 결과와 같이 Sobel edge 기반 경계 마스크의 위치 정확성 측면에서는 개선으로 나타났다. 이는 제안 방법이 단순한 전역 화질 보존보다 경계 영역의 구조 복원에 중점을 두기 때문으로 해석할 수 있다.

<Figure 17>의 결과를 보면, LPIPS가 개선된 사례에서 정제 이후 영상은 정답 영상과 유사한 질감 및 밝기 분포를 보인다. 정답 영상과의 비교 결과에서는 경계 주변의 오차가 감소하고, 실제 윤곽 부근에서만 국소적인 차이가 남았다. 이는 제안한 기법이 단순한 평활화가 아닌 구조 정보를 유지하면서 질감을 재구성함을 의미한다.

4.5 계산 복잡도

제안한 경계 정제 기법은 병합 과정의 수직 및 수평 경계선을 중심으로 경계 패치를 추출하고, 해당 영역에 확산 기반 정제를 수행한다. 전체 연산량은 정제 대상 경계 패치 수와 각 패치의 확산 샘플링 비용으로 결정된다. X-ray를 행 R 개, 열 C 개로 분할 시 수평 경계는 $(R-1)C$ 개, 수직 경계는 $R(C-1)$ 개가 생성되므로 전체 경계 수는 $N_{boundary} = 2RC - R - C$ 이다. 예를 들어, X-ray 영상을 3×3 패치로 구성하는 경우 12개의 경계선이 생기며, 20×20 패치로 나눌 경우 760개의 경계

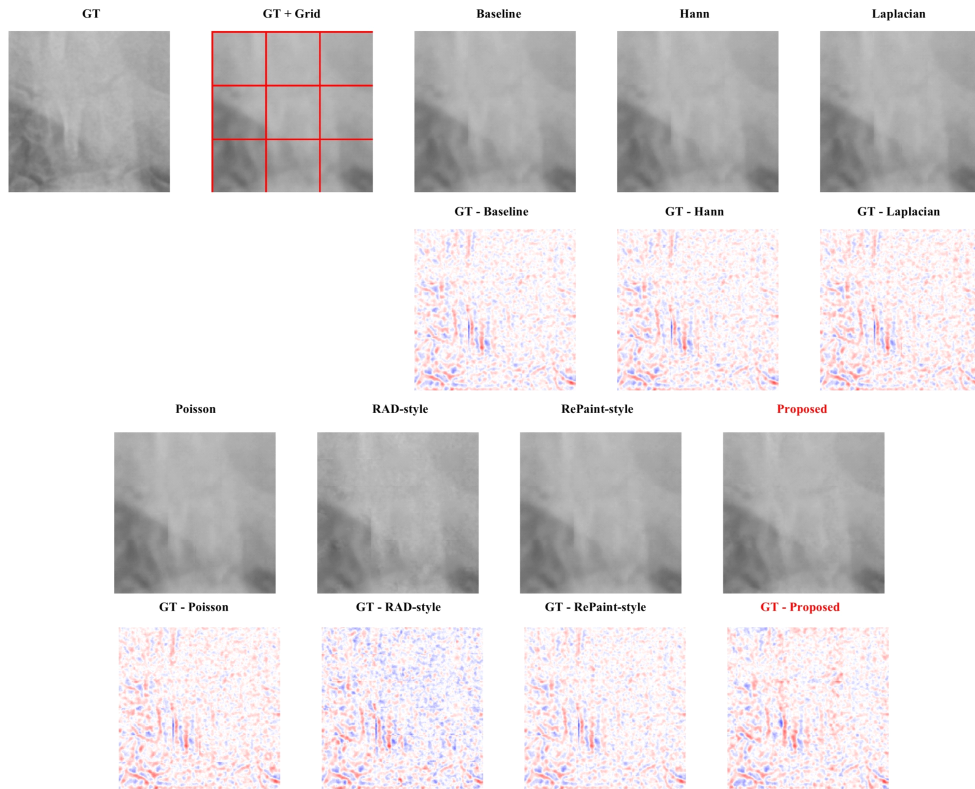
선이 생긴다. 경계 패치 수는 점근적으로 $O(RC)$ 이다.

각 경계 패치는 학습된 역확산 과정을 T 단계에 걸쳐 반복 수행하여 복원되므로, 경계 정제 단계의 전체 추론 비용은 경계 패치 수와 확산 샘플링 단계 수의 곱에 비례하여 $O(RC \cdot T)$ 로 표현된다. 본 연구는 $T=1,000$ 단계의 확산 샘플링을 적용하였으며, 추론 비용은 경계 패치 수와 확산 단계 수 T 에 비례하며, 특히 확산 모델에서는 반복적인 확산 단계 수 T 가 중요한 연산 요인으로 작용한다.

5. 결 론

본 연구는 패치 단위 X-ray 영상 복원 과정에서 발생하는 경계 불연속 문제를 해결하기 위해 확산 기반 경계 정제 기법을 제안하였다. 제안 방법은 SR3 및 SRDiff 기반의 초기 복원 결과로부터 경계 중심 패치를 구성하고, Tukey 함수 기반 가중 마스크로 국소 경계 영역만을 선택적으로 복원한다. 이를 통해 확산확률모델이 경계 일관성 확보에 효과적으로 기여함을 입증하였다.

제안 기법의 효과는 패치 단위, 병합 영상 단위 및 국소 영역 단위에서 수행된 실험을 통해 정량적·정성적으로 검증되었다. 비교 방법으로는 Hann window, Laplacian pyramid blending,



Ground Truth (GT), Ground Truth + Grid, Baseline, Hann, Laplacian, Poisson, RAD-style, RePaint-style, and Proposed results with corresponding difference maps (GT-Method). Red indicates regions reconstructed darker than GT, Blue indicates regions reconstructed brighter than GT

Figure 17. Representative Localized-region Qualitative Results in Terms of LPIPS

Poisson image editing, RAD-style 및 RePaint-style을 포함하였다. 실험 결과, 패치 단위에서는 FSIM과 LPIPS가 유의하게 개선되었으며, Dice coefficient와 HD95는 평균값 기준으로 개선되는 경향을 보였다. 병합 영상 단위에서는 Dice coefficient와 HD95가 유의하게 개선되었으며, 국소 영역 단위에서도 HD95와 LPIPS가 유의하게 개선되었다. 이는 제안 방법이 패치 경계 부위의 구조적 정합성과 시각적 품질을 동시에 개선함을 보여준다. 반면 PSNR과 SSIM은 경계 부근 고주파 성분과 구조적 연속성을 복원하는 과정에서 소폭 감소하였다. 이는 전역 화질 지표와 경계 정합 지표 간의 절충 현상으로 해석된다. 정성적 분석에서도 정제 이전의 인공적인 경계선이 완화되고, 정답 영상의 경계 구조와 자연스럽게 연결되는 윤곽이 재구성됨을 확인하였다. 제안 기법은 별도의 후처리 없이 경계 영역의 통계적 연속성과 시각적 자연성을 향상시켜, 패치 기반 의료 영상 복원에서 구조적 신뢰도와 진단적 일관성을 높이는 데 활용될 수 있다.

다만 제안 기법은 모든 경계 패치에 확산 기반 정제를 반복 수행하므로, 패치 분할 수와 확산 샘플링 단계 수에 비례하여 연산 비용이 증가하기에 패치 분할 수가 많은 고해상도 영상에서는 연산 부담이 커질 수 있다. 또한 정답 영상 기반 차이 분석에서 확인되었듯이, 경계 정제가 국소적으로 적용되는 과정에서 일부 병합 영상에는 경계 위치를 따라 미세한 잔여 오차가 남을 수 있다. 향후 연구에서는 경계 영역의 통계적 차이에 기반한 임계값을 설정하여 이를 초과하는 경계 패치만 선택적으로 정제함으로써 정제 대상 수를 줄이고, 정제 결과의 삽입 방식과 병합 후 경계 일관성을 개선하여 잔여 오차를 완화하고자 한다. 나아가 수의학 분야에서 널리 활용되는 X-ray 영상에 대한 다양한 AI 기반 분석을 가속하는 기반 기술로 확장될 수 있다.

참고문헌

- Ahmadian, K. and Reza-Alikhani, H. R. (2022), X-Ray Medical Image Super-Resolution via Self-Organization Neural Networks and Geometric Directional Gradient, *IET Image Processing*, **16**(14), 3910-3928.
- Buglakova, E., Archit, A., D'Imprima, E., Mahamid, J., Pape, C., and Kreshuk, A. (2025), Tiling Artifacts and Trade-Offs of Feature Normalization in the Segmentation of Large Biological Images, *arXiv preprint arXiv:2503.19545*.
- Burt, P. J. and Adelson, E. H. (1983), A Multiresolution Spline with Application to Image Mosaics, *ACM Transactions on Graphics*, **2**(4), 217-236.
- Ho, J., Jain, A., and Abbeel, P. (2020), Denoising Diffusion Probabilistic Models, *Advances in Neural Information Processing Systems*, **33**, 6840-6851.
- Kim, S., Suh, S., and Lee, M. (2025), RAD: Region-Aware Diffusion Models for Image Inpainting, *Proceeding Computer Vision and Pattern Recognition Conference*, 2439-2448.
- Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., and Shi, W. (2017), Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, *Proceeding IEEE Conference on Computer Vision and Pattern Recognition*, 4681-4690.
- Li, G., Rao, C., Mo, J., Zhang, Z., Xing, W., and Zhao, L. (2024), Rethinking Diffusion Model for Multi-Contrast MRI Super-Resolution, *Proceeding IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11365-11374.
- Li, H., Yang, Y., Chang, M., Chen, S., Feng, H., Xu, Z., Li, Q., and Chen, Y. (2022), SRDiff: Single Image Super-Resolution with Diffusion Probabilistic Models, *Neurocomputing*, **479**, 47-59.
- Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., and Timofte, R. (2021), SwinIR: Image Restoration Using Swin Transformer, *Proceeding IEEE/CVF International Conference on Computer Vision*, 1833-1844.
- Lim, B., Son, S., Kim, H., Nah, S., and Mu Lee, K. (2017), Enhanced Deep Residual Networks for Single Image Super-Resolution, *Proceeding IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 136-144.
- Lugmayr, A., Danelljan, M., Romero, A., Yu, F., Timofte, R., and Van Gool, L. (2022), RePaint: Inpainting Using Denoising Diffusion Probabilistic Models, *Proceeding IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11461-11471.
- Ma, C., Rao, Y., Lu, J., and Zhou, J. (2021), Structure-Preserving Image Super-Resolution, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **44**(11), 7898-7911.
- Mac Aodha, O., Campbell, N. D., Nair, A., and Brostow, G. J. (2012), Patch Based Synthesis for Single Depth Image Super-Resolution, *Proceeding European Conference on Computer Vision*, 71-84, Springer, Berlin, Heidelberg.
- Meng, C., He, Y., Song, Y., Song, J., Wu, J., Zhu, J. Y., and Ermon, S. (2021), SDEdit: Guided Image Synthesis and Editing with Stochastic Differential Equations, *arXiv preprint arXiv:2108.01073*.
- Pérez, P., Gangnet, M., and Blake, A. (2003), Poisson Image Editing, *ACM Transactions on Graphics*, **22**(3), 313-318.
- Pielawski, N. and Wählby, C. (2020), Introducing Hann Windows for Reducing Edge-Effects in Patch-Based Image Segmentation, *PLOS ONE*, **15**(3), e0229839.
- Ronneberger, O., Fischer, P., and Brox, T. (2015), U-Net: Convolutional Networks for Biomedical Image Segmentation, *Proceeding International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234-241, Springer, Cham.
- Saharia, C., Ho, J., Chan, W., Salimans, T., Fleet, D. J., and Norouzi, M. (2022a), Image Super-Resolution via Iterative Refinement, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **45**(4), 4713-4726.
- Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E., Ghasemipour, S. K. S., Ayan, B. K., Mahdavi, S. S., Lopes, R. G., Salimans, T., Ho, J., Fleet, D. J., and Norouzi, M. (2022b), Palette: Image-to-Image Diffusion Models, *ACM SIGGRAPH 2022 Conference Proceedings*, 1-10.
- Uzunova, H., Ehrhardt, J., Jacob, F., Frydrychowicz, A., and Handels, H. (2019), Multi-Scale GANs for Memory-Efficient Generation of High-Resolution Medical Images, *Proceeding International Conference on Medical Image Computing and Computer-Assisted Intervention*, 112-120, Springer, Cham.
- Weese, J. and Lorenz, C. (2016), Four Challenges in Medical Image Analysis from an Industrial Perspective, *Medical Image Analysis*, **33**, 44-49.

- You, C., Li, G., Zhang, Y., Zhang, X., Shan, H., Li, M., Ju, S., Zhao, Z., Zhang, Z., Cong, W., Vannier, M. W., Saha, P. K., Hoffman, E. A., and Wang, G. (2019), CT Super-Resolution GAN Constrained by the Identical, Residual, and Cycle Learning Ensemble (GAN-CIRCLE), *IEEE Transactions on Medical Imaging*, **39**(1), 188-203.
- Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., and Fu, Y. (2018), Image Super-Resolution Using Very Deep Residual Channel Attention Networks, *Proceeding European Conference on Computer Vision (ECCV)*, 286-301.
- Zhao, Y., Wang, R., Jia, W., Yang, J., Wang, W., and Gao, W. (2018), Local Patch Encoding-Based Method for Single Image Super-Resolution, *Information Sciences*, **433**, 292-305.

저자소개

조연경: 광운대학교 정보융합학부 데이터사이언스학과에서 2026년 학사학위를 취득하고, 현재 광운대학교 인공지능융용학과 석사과정에 재학 중이다. 연구분야는 컴퓨터 비전이다.

장효영: 광운대학교 정보융합학부 콘텐츠테크놀로지학과에서 2024년 학사학위를 취득하고, 2026년 광운대학교 인공지능융용학과에서 석사학위를 취득하였다. 연구분야는 컴퓨터 비전, 생성모델이다.

김지훈: 광운대학교 정보융합학부에서 2025년 학사학위를 취득하고, 현재 광운대학교 인공지능융용학과 석사과정에 재학 중이다. 연구분야는 컴퓨터 비전이다.

김정년: 광운대학교 정보융합학부 데이터사이언스학과에서 2022년 학사학위를 취득하고, 현재 광운대학교 인공지능융용학과 석사과정에 재학 중이다. 연구분야는 컴퓨터 비전이다.

김민형: 전남대학교 수의과대학 수의학과에서 2024년 학사 학위, 2026년 석사 학위를 취득하였다. 연구분야는 수의영상진단의학이다.

한수민: 전남대학교 수의과대학 수의학과에서 2024년 학사 학위, 2026년 석사 학위를 취득하였다. 연구분야는 수의영상진단의학이다.

박승조: 전남대학교 수의과대학에서 학사, 석사 및 박사학위를 취득하였다. 현재 전남대학교 수의과대학 교수로 재직 중이다. 주요 연구분야는 수의영상의학이며, 방사선영상, 초음파영상, 컴퓨터단층촬영(CT), 자기공명영상(MRI)을 활용한 영상 진단 및 영상 분석이다.

이상민: 한국외국어대학교 산업공학과에서 2004년 학사학위를 취득하고, KAIST 산업공학과에서 2007년 석사학위, 고려대학교 산업경영공학과에서 2019년 박사학위를 취득하였다. 독일 베를린공대 생산관리 연구실에서 국제 연구 프로그램에 참여한 바 있고, 2007년부터 2020년 사이 티맥스소프트와 삼성전자에 SW 엔지니어와 AI 엔지니어 임무를 수행하였다. 주요 연구 분야는 인공지능 기반 물류·제조 최적화, 영상 이상탐지 외 병리 및 수의학 AI 연구이다.