TESLA: Test-Time Reference-Free Through-Plane Super-Resolution for Multi-Contrast Brain MRI

Yoonseok Choi[1], Sunyoung Jung[2], Mohammed A. Al-masni[3], Ming-Hsuan Yang[4], and Dong-Hyun Kim[1(✉)]

[1] Department of Electrical and Electronic Engineering, Yonsei University, Seoul, Republic of Korea
{yoonseokchoi,donghyunkim}@yonsei.ac.kr

[2] Department of Artificial Intelligence, Yonsei University, Seoul, Republic of Korea
sunyoungj@yonsei.ac.kr

[3] Department of Artificial Intelligence and Data Science, Sejong University, Seoul, Republic of Korea
m.almasani@sejong.ac.kr

[4] Electrical Engineering and Computer Science, University of California at Merced, Merced, USA
mhyang@ucmerced.edu

Abstract. Through-plane super-resolution (SR) in brain magnetic resonance imaging (MRI) is clinically important during clinical assessments. Most existing multi-contrast SR models mainly focus on enhancing in-plane image resolution, relying on functions already integrated into MRI scanners. These methods usually leverage proprietary fusion techniques to integrate multi-contrast images, resulting in diminished interpretability. Furthermore, the requirement for reference images during testing limits their applicability in clinical settings. We propose a TEst time reference-free through-plane Super-resoLution network using disentAngled representation learning in multi-contrast MRI (TESLA) to address these challenges. Our method is developed on the premise that multi-contrast images consist of shared content (structure) and independent stylistic (contrast) features. Thus, after progressively reconstructing the target image in the first stage, we divide it into shared and independent elements during the structure enhancement phase. In this stage, we employ a pre-trained ContentNet to effectively disentangle high-quality structural information from the reference image, enabling the shared components of the target image to learn directly from those of the reference image through patch-wise contrastive learning during training. Consequently, the proposed model enhances clinical applicability while ensuring model interpretability. Extensive experimental results demonstrate that the proposed model performs favorably against other state-of-the-art multi-contrast SR models, especially in restoring structural fine details in the through-plane direction. The code is publicly available at https://github.com/Yonsei-MILab/TESLA.

J. C. Gee et al. (Eds.): MICCAI 2025, LNCS 15972, pp. 584–593, 2026.
https://doi.org/10.1007/978-3-032-05169-1_56

Keywords: Super-resolution · Contrastive learning · Multi-contrast MRI

1 Introduction

The clinical necessity for Super-Resolution (SR) in the through-plane direction in brain Magnetic Resonance Imaging (MRI) has been highlighted due to its potential to induce patient discomfort during medical examinations. During regular medical examinations, T2 MR scans are commonly performed with thick-slice, while the spatial domain is acquired at relatively high-resolution (HR) to preserve the signal-to-noise ratio of the data [9,12]. Conventional super-resolution approaches [5,7] struggle to preserve sharp boundaries and fail to remove stair-step artifacts. While the Low-Rank Total Variation (LRTV) method [19] shows improvement in MRI SR, it comes at a high computational cost. Deep learning-based SR networks [15,16] designed for computer vision often rely on a single contrast image, making them less suitable for multi-contrast MR data. In contrast, recent reference-based image SR models [6,14,17] have demonstrated significant potential in recovering the high-frequency details of low-resolution (LR) target image (Tar) by utilizing features from reference images (Ref). This concept is especially advantageous for MRI-based SR networks leveraging multi-contrast MR data. These approaches enhance SR performance by using the complementary information from each MR contrast through either channel concatenation or self-attention fusion mechanisms to assess the correlation between LR Tar and HR Ref. However, the aforementioned SR methods have mostly focused on in-plane SR in brain MRI, which can be addressed with the function of the existing MR scanner [2]. Moreover, these implicit fusion techniques may lack interpretability and necessitate HR Ref at test time, thus reducing their clinical applicability.

In this work, we propose a TEst time reference-free through-plane Super-resoLution network using disentAngled representation learning in multi-contrast MRI (TESLA). Our framework consists of Progressive Reconstruction (PR) and Structure Enhancement (SE) stages (See Fig. 1). We assume that multi-contrast MR images can be decomposed into shared content information (*i.e.* structure) and distinct stylistic features (*i.e.* contrast), which we leverage in the second stage. In the PR phase, we progressively reconstruct the LR Tar, aiming to not only minimize structural distortion between LR Tar and HR Tar but also reduce the domain gap of shared content with HR Ref. In the SE phase, we utilize a pre-trained ContentNet to extract valuable content information from HR Ref. Then, we enable the smooth content feature decomposed from coarsely reconstructed SR Tar to learn features for HR Tar through contrastive learning explicitly. We also implement a data consistency (DC) term that allows the final output to retain structural fine details and physical meaning. The contributions of this work are: (1) We present a model that enables SR Tar to directly learn high-quality content information from HR Ref, thereby enhancing the interpretability of the model. (2) Unlike multi-contrast SR techniques, the proposed

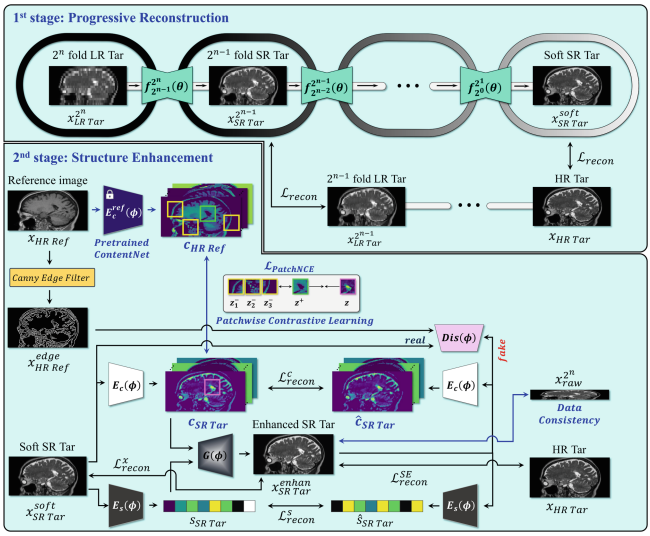

Fig. 1. Overview of the proposed network TESLA. In the first stage, we progressively reconstruct LR Tar. In the second stage, we leverage the high-quality content information disentangled from HR Ref with the pre-trained ContentNet to enrich the structural fine detail of Soft SR Tar with the patch-wise contrastive learning. HR Ref and ContentNet are used only during training; inference requires only PR, the encoders (E_c and E_s), and decoder (G) in SE.

network operates independently of HR Ref during the testing phase, significantly increasing its clinical applicability. (3) Our experimental findings demonstrate that the proposed model outperforms alternatives in accurately preserving fine details of the brain on the IXI, HCP, In-house, and BraTS21 datasets.

2 Method

The proposed method aims to improve the slice thickness resolution of LR Tar by utilizing multi-contrast MR images. The framework comprises two principal components, as depicted in Fig. 1.

Progressive Reconstruction. In the first stage, a simulated 2^n fold LR Tar ($x^{2^n}_{LRTar}$) is progressively reconstructed using an optimized nnU-Net [8] for gradually reducing the domain gap of shared content with HR Ref as well as mitigating structural distortion between LR Tar and HR Tar using

$$\mathcal{L}^{PR}_{recon} = \sum_{i=n}^{1} \mathcal{L}_{l1+ssim}\left(x^{2^{i-1}}_{SRTar}, x^{2^{i-1}}_{LRTar}\right), n = 1, 2, 3, \ldots \quad (1)$$

Table 1. Utilized datasets on several tasks. Tar indicates low-resolution (LR) target images, and Ref denotes high-resolution (HR) reference images. We employed the sagittal plane images as an input on all datasets. Tar and Ref are described in the form of modality/slice thickness.

Task		Through-plane SR			Pseudo-vessel recon	
Dataset		IXI	HCP	BraTS21	IXI	In-house
Train	Tar	LR T2/4.8 mm	LR T2/5.6 mm	LR FLAIR/4.0 mm	LR T2/4.8 mm	–
	Ref	HR T1/1.2 mm	HR T1/0.7 mm	HR T1CE/1.0 mm	HR T1/1.2 mm	–
Test	Tar	LR T2/4.8 mm	LR T2/5.6 mm	LR FLAIR/4.0 mm	–	LR T2/4.0 mm

where $\mathcal{L}_{recon}^{PR}$ means reconstruction loss on PR phase, $\mathcal{L}_{l1+ssim}$ denotes the weighted combination of L1 and SSIM loss, with a ratio 1:1, and $x_{SRTar}^{2^{i-1}} = f_{2^{i-1}}^{2^i}\left(x_{LRTar}^{2^i}\right)$ indicates 2^{i-1} fold reconstructed SR Tar, respectively.

Structure Enhancement. Although CNN-based reconstruction networks can minimize the disparity between the input and the label, they often smooth out structural information in the process [3]. To address this issue, the SE phase employs a pre-trained ContentNet$\left(E_c^{ref}\right)$ that has been fine-tuned to effectively disentangle high-quality content features (c_{HRRef}) from HR Ref. The ContentNet denotes the content encoder of MUNIT [11]. It comprises multiple convolutional layers for downsampling the input and a residual block designed to extract additional semantic structure. We show that integrating L1 and SSIM loss for reconstruction, and adversarial loss for pretraining ContentNet effectively extracts the most dynamic content information from HR Ref. In contrast to MUNIT, which uses LSGAN, we enhance the structural details of the extracted content features from HR Ref by using PatchGAN as the discriminator conditioned on the edge image obtained by the Canny edge filter from HR Ref (See Fig. 4). The losses for pre-training the ContentNet are:

$$\mathcal{L}_{recon}^{x} = \mathbb{E}_{x\sim p(x)}[\mathcal{L}_{l1+ssim}\left(G(E_c(x), E_s(x)), x\right)] \tag{2}$$

$$\mathcal{L}_{recon}^{c} = \mathbb{E}_{c\sim p(c), s\sim q(s)}[\mathcal{L}_{l1+ssim}\left(E_c(G(c, s)), c\right))] \tag{3}$$

$$\mathcal{L}_{recon}^{s} = \mathbb{E}_{c\sim p(c), s\sim q(s)}[\mathcal{L}_{l1+ssim}\left(E_s(G(c, s)), s\right))] \tag{4}$$

$$\begin{aligned}\mathcal{L}_{Adv}^{x} = \mathbb{E}_{c\sim p(c),\, s\sim q(s)}\left[\log\left(1 - Dis\left(G(c, s),\, x_{HRRef}^{edge}\right)\right)\right] \\ + \mathbb{E}_{x\sim p(x)}\left[\log\left(Dis(x,\, x_{HRRef}^{edge})\right)\right]\end{aligned} \tag{5}$$

where $x = x_{HRRef}$, $p(c)$ is given by $c = c_{HRRef} = E_c(x) = E_c^{ref}(x)$, and $q(s)$ is given by $s = E_s(x)$, respectively. Here, E_c and E_s indicate the content and style encoders. Additionally, G and Dis refer to the generator and discriminator, respectively. Equation 2 through Eq. 5 are all applied at the same rate. We applied the same losses as ContentNet to the training process of the SE phase. In this case, we simply replaced x_{HRRef} with x_{SRTar}^{Soft}.

Patchwise Contrastive Learning. Technically, c_{SRTar} and c_{HRRef} are not identical, although they share some content. In this context, we can observe that the degree of correlation between each region in c_{SRTar} and c_{HRRef} varies. When considering a patch representing the cerebrospinal fluid (CSF) in c_{SRTar}, the corresponding region in c_{HRRef} exhibits a stronger correlation compared to other patches in c_{HRRef}. Instead of using pixel-level loss, we maximize the mutual information in the latent space between positive feature pairs extracted from each patch of c_{SRTar} and c_{HRRef} and minimize the corresponding information between negative feature pairs from each patch of them. As such, we use PatchNCE loss [18] based on contrastive learning to enable c_{SRTar} to explicitly learn the mutual information with c_{HRRef}. To extract the feature stack from c_{SRTar} and c_{HRRef}, we employ E_c and the pre-trained E_c^{ref}, respectively. Each layer and spatial position within this feature stack corresponds to a patch of the input image, where deeper layers are associated with larger patches. We select L layers of interest and pass the feature maps through a small two-layer MLP network (H_l), as in SimCLR [4]. This results in the feature stack denoted as $\{z_l\}_L = \left\{H_l\left(E_c^l\left(c_{SRTar}\right)\right)\right\}_L$, where E_c^l indicates the output of the l^{th} selected layer. We index the layers as $l \in \{1, 2, \ldots, L\}$ and denote spatial locations as $s \in \{1, \ldots, S_l\}$, where S_l represents the number of spatial locations in each layer. The corresponding feature is referred to as $z_l^s \in \mathbb{R}^{C_l}$, while the other features are denoted as $z_l^{S\backslash s} \in \mathbb{R}^{(S_l-1)\times C_l}$, with C_l being the number of channels at each layer. Similarly, we encode c_{HRRef} into $\{\hat{z}_l\}_L = \left\{H_l\left(E_c^{ref\cdot l}\left(c_{HRRef}\right)\right)\right\}_L$.

$$\mathcal{L}_{PatchNCE} = \sum_{l=1}^{L}\sum_{s=1}^{S} CE\left(z_l^s, \hat{z}_l, z_l^{S\backslash s}\right) \tag{6}$$

where CE is Cross Entropy loss. The z_l^s,$\hat{z}_l$, and $z_l^{S\backslash s}$ are mapped query, positive, and negatives. To maintain the physical significance of the reconstructed x_{SRTar}^{enhan}, we average x_{SRTar}^{enhan} along the through-plane direction to ensure it matches $x_{raw}^{2^n}$,

$$DC = \mathcal{L}_{l1+ssim}\left(avg\left(x_{SRTar}^{enhan}\right), x_{raw}^{2^n}\right) \tag{7}$$

The total loss in the SE phase is computed by Eqs. 2–7 with equal weights.

Implementation Details. The proposed framework is trained with batch size 10 on an NVIDIA A5000 GPU with 24GB memory for 100 epochs, which takes about 10 h for each experiment. We use Adam optimizer with a learning rate of 1×10^{-4} for all experiments.

Table 2. Quantitative comparison results of different tasks on several datasets. All metrics are expressed in the format of mean(std).

Task	Through-plane SR						Pseudo-vessel recon	
Dataset(scale factor)	IXI(×4)		HCP(×8)		BraTS21(×4)		In-house(×4)	
Method	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
LRTV [19]	0.6707(0.04)	14.48(1.06)	0.7633(0.01)	22.44(0.60)	0.7861(0.04)	17.32(1.55)	0.7755(0.02)	15.72(0.65)
MINet [6]	0.9414(0.03)	28.91(1.65)	0.9236(0.02)	27.46(1.29)	0.9325(0.02)	28.33(2.06)	0.8402(0.01)	24.17(0.71)
MASA [17]	0.9189(0.01)	26.86(0.94)	–	–	0.9299(0.02)	28.02(2.11)	0.9002(0.02)	24.53(0.92)
MC-VarNet [14]	0.9343(0.03)	28.27(1.63)	0.9306(0.02)	27.38(1.27)	0.9298(0.02)	27.90(2.11)	0.8629(0.02)	23.60(0.80)
TESLA(ours)	0.9532(0.01)	29.23(0.77)	0.9489(0.02)	28.50(0.73)	0.9432(0.01)	29.11(1.89)	0.9144(0.01)	25.67(0.87)

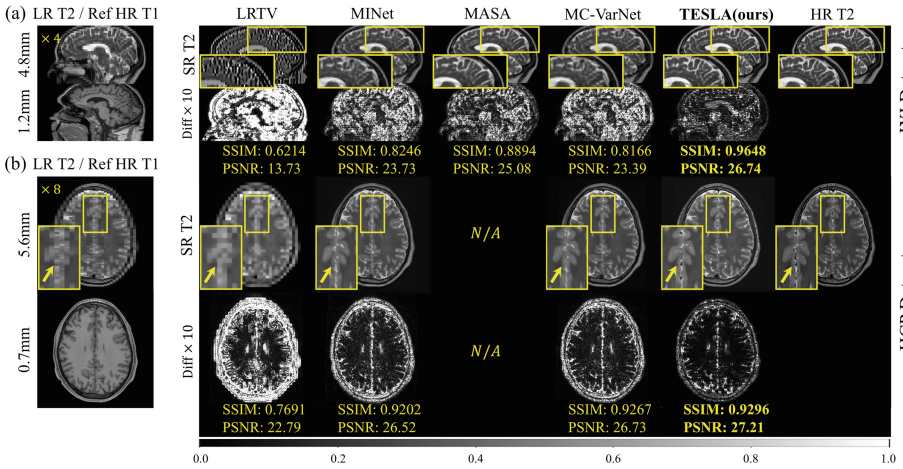

Fig. 2. Qualitative comparison results of through-plane SR on IXI and HCP dataset.

3 Experimental Results

Utilized Datasets for Different Tasks. We employ four datasets: IXI [10], HCP [20], BraTS21 [1], and In-house, each with different conditions for the respective tasks, as shown in Table 1. For Tar and Ref in all datasets, we utilize center 100 key slices per subject, inclusive of brain tissue and absent of substantial artifacts. We normalize the intensity to a range of 0 to 1 without any additional augmentations. To simulate LR Tar, we use b-spline interpolation, which effectively generates a realistic representation of the stair-step artifact in the through-plane direction. Note that all LR Tar is simulated except for the in-house dataset. We utilize the IXI [10], HCP [20], and BraTS21 [1] datasets for the through-plane SR. Unlike HCP and BraTS21, which are aligned, we use Elastix [13] to register between Tar and Ref for the IXI dataset. It is important to note that Tar and Ref are paired during the training process. In general, 50 subjects are randomly selected from each dataset, with 40 designated for training and 10 for testing. For the IXI and BraTS21 datasets, sagittal plane images are utilized as input. In contrast, for the HCP dataset, axial plane images, which

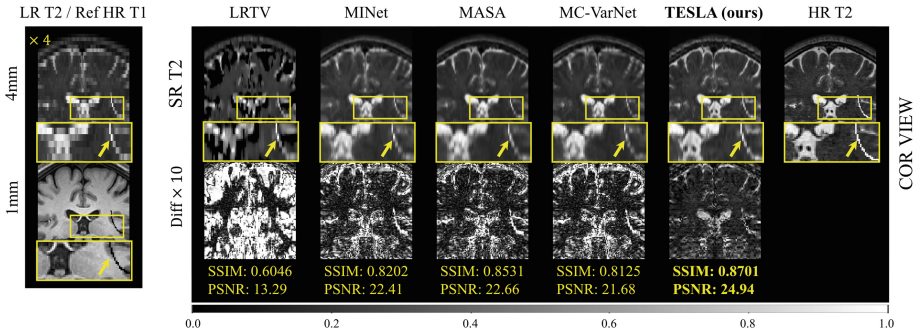

Fig. 3. Qualitative comparison results of pseudo-vessel reconstruction on in-house dataset when the scaling factor is ×4.

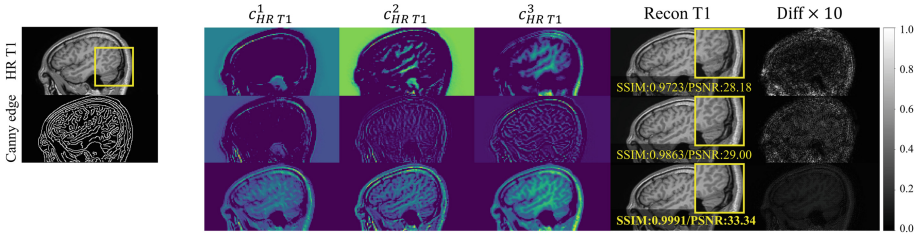

Fig. 4. Qualitative results of the ablation study on the loss combinations in ContentNet on the IXI dataset, which effectively decomposes high-quality structural information from HR Ref. First row: MUNIT (L1 + Perceptual + Adversarial) + LSGAN, Second row: MUNIT (L1 + Perceptual + Adversarial) + PatchGAN, and Third row: MUNIT (L1 + SSIM + Adversarial) + PatchGAN. $c^i_{HRT1}, (i = 1, 2, 3)$ indicates randomly selected content information decomposed from HR T1 on each condition.

are perpendicular to the sagittal direction, are employed as input, as the data are originally acquired in the sagittal orientation. Each dataset has the following matrix sizes: IXI: 128×256, HCP: 320×320, BraTS21: 192×192. The modalities used as Tar and Ref in each dataset are shown in Table 1. In the pseudo-vessel reconstruction task, the model trained on the IXI dataset is evaluated using LR T2 from the in-house dataset, which has one subject with a slice thickness of 4 mm. To validate clinical applicability, the pseudo-vessel is simulated to resemble a quarter ellipse when viewed in the coronal plane, as illustrated in Fig. 3.

Comparative Experiments. To evaluate the effectiveness of the proposed TESLA, we conduct a comparative analysis against one model-based single-contrast SR method: LRTV [19] and three deep learning-based multi-contrast SR networks: MINet [6], MASA [17], MC-VarNet [14]. Note that these three approaches require HR Ref during testing. Furthermore, for the pseudo-vessel reconstruction, an additional HR T1 with a slice thickness of $1.0mm$ is acquired and utilized as HR Ref for testing the comparison models. Figure 2(a) demon-

Table 3. Ablation study on the contribution of each stage and term.

PR	SE	DC	SSIM	PSNR
✓			0.9318(0.01)	25.63(1.15)
✓	✓		0.9521(0.01)	28.34(1.09)
✓	✓	✓	0.9532(0.01)	29.23(0.77)

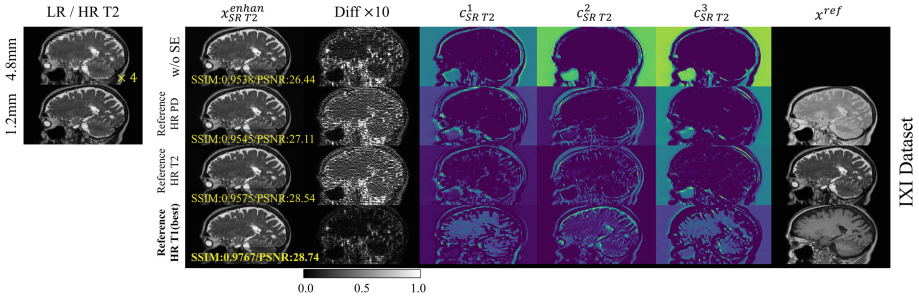

Fig. 5. Qualitative results of the ablation study analyzing the optimal HR Ref on the IXI dataset when the scaling factor is ×4. x_{SRT2}^{enhan} means the final output of the proposed network on each condition. $c_{SRT2}^{i}(i = 1, 2, 3)$ indicates randomly selected content information decomposed from SR T2 on each condition. x^{ref} denotes HD Ref.

strates that, in the through-plane SR task, the proposed model most effectively recovers the intricate CSF structure in the upper region of SR T2 when the slice thickness of LR T2 on the IXI dataset is scaled by a factor of 4. While the compared SOTA models require upsampling in two axes due to their in-plane SR design, TESLA performs SR only along the through-plane axis, yet still achieves superior reconstruction of anatomical details along the z-axis. The yellow arrow in Fig. 2(b) highlights that only the proposed framework successfully reconstructs the hypo-intensity structure adjacent to the longitudinal cerebral fissure in LR T2 when the HCP dataset is scaled by a factor of 8.

Figure 3 depicts the reconstruction outcomes of a pseudo-vessel simulated within LR T2 of an in-house dataset, acquired with a slice thickness of $4mm$, presented in coronal and sagittal views. Especially, Fig. 3 demonstrates that TESLA outperforms other models in reconstructing both the pseudo-vessel and the intricate anatomical structures of the brain. To enhance clinical applicability, we also conduct through-plane SR on the BraTS21 dataset at a scale of 4. Table 2 displays quantitative metrics demonstrating that our method consistently outperforms the alternatives on Structural Similarity Index Measure (SSIM) and Peak Signal-to-Noise Ratio (PSNR) in both the through-plane SR and pseudo-vessel reconstruction tasks.

Ablation Study. We conduct three ablation studies: an assessment of the contribution of each stage of the proposed framework, an evaluation of the loss

combination in ContentNet that most effectively separates high-quality structural information from HR Ref, and an analysis of the optimal MR contrast to use as HR Ref for restoring the structural details of LR Tar in multi-contrast SR networks. As shown in Fig. 4, in contrast to the original MUNIT [11](first row in Fig. 4), where they utilize L1 and perceptual loss as the reconstruction loss with the adversarial loss, we effectively disentangle the structural information of HR Tar by training the MUNIT generator with L1 and SSIM loss as the reconstruction loss. Additionally, we use PatchGAN as the discriminator, conditioned on edge images extracted from HR Tar using the Canny edge filter. The yellow box in Fig. 4 demonstrates that fine-tuning ContentNet under these specific conditions yields artifact-free reconstructions of T1 images that closely resemble the original. Table 3 demonstrates that the proposed model achieves the highest average score in both SSIM and PSNR metrics when incorporating PE, SE, and DC terms. Figure 5 demonstrates that x_{SRT2}^{enhan} exhibits the most dynamic structural features when T1 is used as HR Ref.

4 Conclusion

The proposed method effectively reconstructs the structural details of LR Tar by employing disentangled content information from HR Ref. This method improves the interpretability of network performance enhancements, in contrast to traditional models that solely extract features from HR Ref and utilize specific fusion techniques in a "black-box" scheme. In contrast to state-of-the-art methods, the proposed model does not require HR Ref during the test phase, which substantially benefits clinical practice by tackling the common challenge of absent modalities. Our approach demonstrates significant potential for enhancing through-plane SR capabilities while adequately meeting the clinical requirements in brain MRI.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2025-00561616 and RS-2023-00243034).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bakas, S., et al.: Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. arXiv preprint arXiv:1811.02629 (2018)
2. Behl, N.: Deep resolveG mobilizing the power of networks. MAGNETOM Flash (78) **1** (2021)
3. Chen, C., Chen, X., Cheng, H.: On the over-smoothing problem of cnn based disparity estimation. In: Proceedings of the IEEE International Conference on Computer Vision (2019)

4. Chen, T., Kornblith, S., Norouzi, M., Hinton, G.: A simple framework for contrastive learning of visual representations. In: Proceedings of the International Conference on Machine Learning (2020)
5. Dong, C., Loy, C.C., He, K., Tang, X.: Image super-resolution using deep convolutional networks. IEEE Trans. Pattern Anal. Mach. Intell. **38** (2015)
6. Feng, C.M., Fu, H., Yuan, S., Xu, Y.: Multi-contrast mri super-resolution via a multi-stage integration network. In: Medical Image Computing and Computer Assisted Intervention (2021)
7. Feng, C.M., Wang, K., Lu, S., Xu, Y., Li, X.: Brain mri super-resolution using coupled-projection residual network. Neurocomputing **456** (2021)
8. Futrega, M., Milesi, A., Marcinkiewicz, M., Ribalta, P.: Optimized u-net for brain tumor segmentation. In: International MICCAI Brainlesion Workshop (2021)
9. Gholipour, A., et al.: Super-resolution reconstruction in frequency, image, and wavelet domains to reduce through-plane partial voluming in mri. Med. Phys. **42** (2015)
10. Group, B.I.A.: Ixi dataset (2025). https://brain-development.org/ixi-dataset/. Accessed 17 Feb 2025
11. Huang, X., Liu, M.Y., Belongie, S., Kautz, J.: Multimodal unsupervised image-to-image translation. In: Proceedings of the European Conference on Computer Vision (2018)
12. Jia, Y., Gholipour, A., He, Z., Warfield, S.K.: A new sparse representation framework for reconstruction of an isotropic high spatial resolution mr volume from orthogonal anisotropic resolution scans. IEEE Trans. Med. Imaging **36** (2017)
13. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.: Elastix: a toolbox for intensity-based medical image registration. IEEE Trans. Med. Imaging **29** (2009)
14. Lei, P., Fang, F., Zhang, G., Zeng, T.: Decomposition-based variational network for multi-contrast mri super-resolution and reconstruction. In: Proceedings of the IEEE International Conference on Computer Vision (2023)
15. Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., Timofte, R.: Swinir: image restoration using swin transformer. In: Proceedings of the IEEE International Conference on Computer Vision (2021)
16. Lim, B., Son, S., Kim, H., Nah, S., Mu Lee, K.: Enhanced deep residual networks for single image super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2017)
17. Lu, L., Li, W., Tao, X., Lu, J., Jia, J.: Masa-sr: matching acceleration and spatial adaptation for reference-based image super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2021)
18. Park, T., Efros, A.A., Zhang, R., Zhu, J.Y.: Contrastive learning for unpaired image-to-image translation. In: Proceedings of the European Conference on Computer Vision (2020)
19. Shi, F., Cheng, J., Wang, L., Yap, P.T., Shen, D.: LRTV: MR image super-resolution with low-rank and total variation regularizations. IEEE Trans. Med. Imaging **34** (2015)
20. Van Essen, D.C., et al.: The human connectome project: a data acquisition perspective. Neuroimage **62** (2012)