Academic Journal of Emerging Technologies
ISSN Print: 3104-4417
ISSN Online: 3104-4425
CD ISSN: 3104-4433
https://doi.org/10.63313/AJET.9078

FMC-Net: A Frequency-Domain Completion Network for Brain Tumor Segmentation from Incomplete MRI

Taoyu Wang. Zhichao Hu. Jincheng Li. Yuhao Ma and Jie Xu
Tianjin University of Technology and Education, Tianjin, China
Email: 1808665869@qq.com

How to cite this paper: Wang, T., Hu, Z., Li, J., Ma, Y., & Xu, J. (2026). FMC-Net: A frequency-domain completion network for brain tumor segmentation from incomplete MRI. Academic Journal of Emerging Technologies, 4(1), 44–53. ISSN Print: 3104-4417, ISSN Online: 3104-4425. https://doi.org/10.63313/AJET.9078
Published: 2026-11-16

Copyright © 2026 by author(s) and Erytis Publishing Limited.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

Abstract

Multimodal MRI supports brain tumor segmentation, but some sequences may be unavailable in routine care. FMC-Net uses a shared encoder with availability-weighted aggregation, spectral completion via Fourier-domain feature-wise linear modulation, and learnable fusion across frequency bands. Five-fold cross-validation on BraTS 2020 showed that FMC-Net (8.39M parameters) outperformed standard 3D U-Net and a HeMIS-style U-Net in mean Dice across all 15 nonempty modality combinations. With FLAIR and contrast-enhanced T1 only, mean Dice reached 71.38%, a 2.15-point gain over standard 3D U-Net ($p < 0.001$). The gain reached 14.71 points with T1 alone. Ablation showed that random modality dropout was necessary for robustness. Without fine-tuning, enhancing tumor Dice reached 76.87% on BraTS 2021. The model supports CPU inference.

Keywords

Missing modalities; brain tumor segmentation; frequency-domain completion; lightweight network; Fourier transform; MRI

1. Introduction

Multimodal MRI supports glioma diagnosis, treatment planning, and follow-up [1, 2]. T1-weighted (T1), contrast-enhanced T1-weighted (T1ce), T2-weighted (T2), and fluid-attenuated inversion recovery (FLAIR) images provide complementary views of tumor anatomy (Figure 1). In routine care, limited scan time, motion, scanner availability, and contraindications to contrast can leave one or more sequences unavailable.

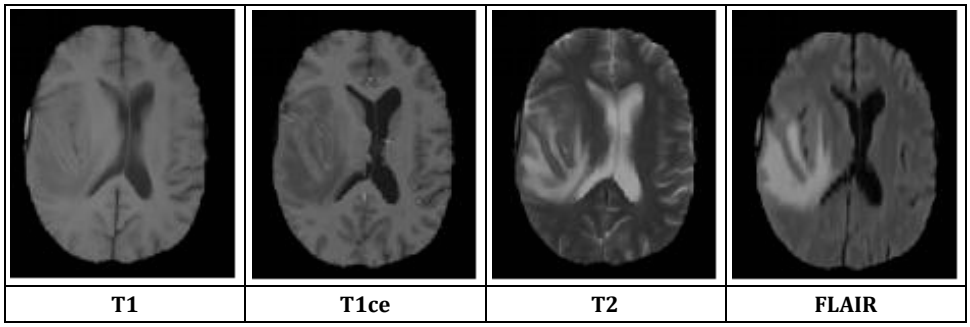


Figure 1. Representative images from the four MRI sequences used for segmentation.

Convolutional and Transformer segmenters [5–12] generally assume complete inputs, and zero-filling missing sequences can degrade performance. HeMIS aggregates modality features, while U-HVED jointly completes and segments inputs [3, 4]. Other approaches use cross-modal fusion [13–15], knowledge distillation [16, 17], modality synthesis [18], or specialized fusion and representation learning [19–23]. Per-combination results and transfer without fine-tuning are reported less consistently. Frequency-domain operations can model global relationships efficiently [24, 25]. FMC-Net uses spectral completion and frequency-selective fusion in an 8.39M-parameter model. We evaluate all 15 nonempty modality combinations and transfer to two external datasets.

2. Methods

2.1. Task and Architecture

Given any nonempty subset of the four MRI sequences, FMC-Net predicts the whole tumor (WT), tumor core (TC), and enhancing tumor (ET). The output labels are background (0), WT (1), TC (2), and ET (3). Missing inputs are zero-filled and indicated by a binary mask A, with $A_m = 1$ for available modality m. Figure 2 shows the shared encoder, bottleneck, Spectral Completion Module (SCM), Cross-Modal Multi-Frequency Fusion (CMFF) module, and decoder.

The shared encoder extracts three levels of skip features (24, 48, and 96 channels). Masked skip features are averaged across available modalities. Stacked deepest features are fused by a 1 × 1 × 1 convolution and downsampled to a 192-channel bottleneck at D/8 × H/8 × W/8. SCM and CMFF process the bottleneck features before four decoder stages restore resolution by trilinear upsampling and skip connections. A 1 × 1 × 1 convolution produces the class logits.

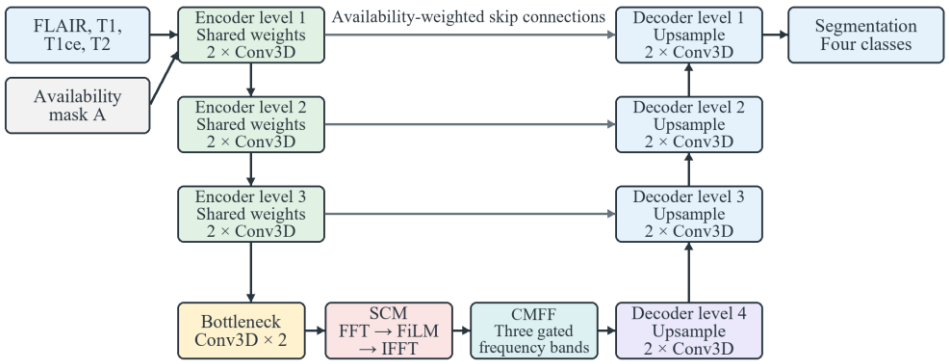


Figure 2. Overall architecture of FMC-Net.

2.2. Shared Encoding and Aggregation

The same encoder processes each sequence independently. For encoder level k and modality m, Equations (1) and (2) mask and average the skip features. Weight sharing reduces encoder parameters by 75% compared with four independent encoders. Aggregation adds no trainable parameters.

$$f_{k,m} \leftarrow f_{k,m} \times A_m \tag{1}$$

$$skip_k = \frac{\sum_m A_m \cdot f_{k,m}}{max(\sum_m A_m, 1)} \tag{2}$$

2.3. Spectral Completion Module

SCM uses shared low-frequency anatomy and complementary higher-frequency contrast to infer missing spectral information (Figure 3). A 3D real fast Fourier transform (FFT) maps bottleneck features x of shape B × C × D × H × W.

$$F = RFFT_{3D}(x) \in \mathbb{C}^{B \times C \times (\lfloor D/2 \rfloor + 1) \times H \times W} \tag{3}$$

A two-layer multilayer perceptron (MLP) with 64 hidden units and ReLU activation maps the binary B × 4 mask to channel-wise feature-wise linear modulation (FiLM) coefficients γ and β. The reported FiLM parameter count is approximately 640.

$$\gamma = MLP_\gamma(A), \quad \beta = MLP_\beta(A), \quad \gamma, \beta \in \mathbb{R}^{B \times C} \tag{4}$$

FiLM modulates magnitude while preserving phase. An inverse real FFT then restores spatial features. The FFT has O(N log N) complexity and gives the modulation access to global information.

$$|F'|_{b,c} = (1 + \gamma_{b,c}) \cdot |F|_{b,c} + \beta_{b,c} \tag{5}$$

$$x' = IRFFT_{3D}(F') \in \mathbb{R}^{B \times C \times D \times H \times W} \tag{6}$$



Figure 3. Spectral Completion Module (SCM).

2.4. Cross-Modal Multi-Frequency Fusion

CMFF separates the completed features into high-frequency edges, mid-frequency transitions, and low-frequency shape cues (**Figure 4**). In Equations (7)–(9), G_3 and G_5 are 3^3 and 5^3 Gaussian filters, and the low-band branch uses a 3^3 depthwise convolution.

$$x_{high} = x - G_3(x) \quad (7)$$

$$x_{mid} = G_3(x) - G_5(x) \quad (8)$$

$$x_{low} = DepthwiseConv_3(x) \quad (9)$$

Independent 1 × 1 × 1 convolutions give each band a sigmoid gate G and transformed feature V. Residual fusion combines the gated bands and preserves x if all gates close.

$$G_{band} = \sigma(Conv_{1\times1\times1}(x_{band})) \in [0,1], \quad V_{band} = Conv_{1\times1\times1}(x_{band}) \quad (10)$$

$$x_{out} = x + \sum_{band} G_{band} \odot V_{band} \quad (11)$$

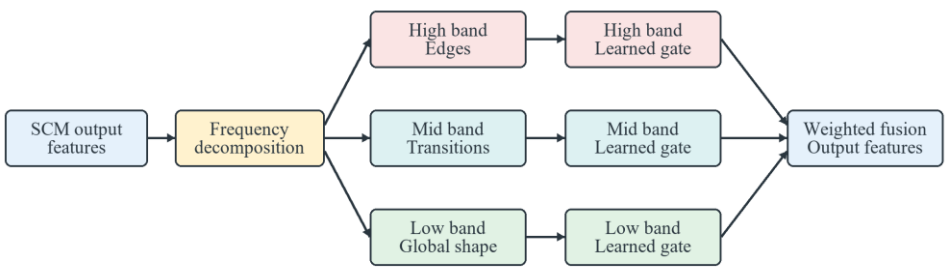


Figure 4. Cross-Modal Multi-Frequency Fusion (CMFF).

2.5. Modality Dropout and Loss

During training, each modality is independently dropped with probability 0.3 while retaining at least one sequence. Equation (12) gives equal weight to foreground Dice loss (WT, TC, ET) and four-class voxel-wise cross-entropy. Training uses no deep supervision.

$$L_{total} = 0.5 \times L_{Dice} + 0.5 \times L_{CE} \quad (12)$$

3. Experiments

3.1. Implementation Details

Five-fold cross-validation (CV) used all 369 BraTS 2020 cases, shuffled with seed 42. Each sequence was z-score normalized over nonzero voxels. Training used random 128^3 crops, AdamW, FP16, and up to 200 epochs on two 48 GB RTX A6000 GPUs. Batch sizes were 8 for Baseline and 6 for M. The L configuration used batches of 4 with two-step gradient accumulation.

Inference used 128^3 or 192^3 center crops and averaged logits across eight axis-flip test-time augmentations (TTA). We scored WT, TC, and ET Dice and their mean, plus 95th-percentile Hausdorff distance (HD95) for selected settings. Missing channels were zero-filled for each tested combination.

The matched baselines were a standard early-fusion 3D U-Net (Vanilla, 3.02M parameters), a HeMIS-style U-Net with independent encoders and mean–variance pooling (3.71M), and a shared-encoder U-Net without SCM or CMFF (3.04M) for ablation.

3.2. Segmentation with Missing Modalities

3.2.1. Comparison across All 15 Modality Combinations

Table 1 compares all 15 nonempty modality combinations under the same training and evaluation protocol.

Table 1. Dice scores (%) across 15 modality combinations on BraTS 2020 (128^3 crops, five-fold CV).

F	T1	T1c	T2	V WT	H WT	FMC WT	V TC	H TC	FMC TC	V ET	H ET	FMC ET
●	○	○	○	64.4	56.1	**66.0**	31.3	24.5	**34.9**	29.2	20.6	**33.3**
○	●	○	○	18.0	18.9	**36.5**	17.9	19.7	**31.4**	7.9	7.7	**19.6**
○	○	●	○	42.3	26.1	**43.7**	53.3	27.2	**58.1**	64.6	46.7	**68.3**
○	○	○	●	50.8	43.8	**57.7**	27.3	26.6	**36.9**	27.0	25.1	**33.7**
●	●	○	○	66.5	64.3	**68.0**	39.8	37.8	**45.1**	29.8	32.8	**37.5**
●	○	●	○	74.6	74.1	**75.8**	62.4	60.1	**65.1**	70.3	70.8	**72.5**
●	○	○	●	67.1	66.6	**69.4**	38.2	40.5	**44.3**	31.4	34.0	**40.5**
○	●	●	○	49.5	47.8	**55.2**	57.7	56.5	**64.8**	68.9	69.0	**71.3**
○	●	○	●	57.9	50.2	**63.0**	36.9	36.0	**44.2**	28.0	24.6	**37.7**
○	○	●	●	66.1	63.5	**69.1**	62.4	59.5	**65.1**	70.4	69.5	**72.0**
●	●	●	○	**75.8**	74.7	75.6	63.9	60.7	**66.4**	**72.7**	71.5	72.6
●	●	○	●	68.4	67.8	**70.5**	43.3	43.1	**47.2**	32.5	36.0	**42.3**
●	○	●	●	76.5	75.8	**77.1**	64.1	62.5	**65.4**	71.2	71.0	**71.7**
○	●	●	●	67.8	65.1	**70.0**	62.7	61.1	**65.9**	71.3	70.1	**73.0**
●	●	●	●	**77.1**	75.7	76.9	64.3	62.5	**65.8**	**72.6**	70.5	72.1

F = FLAIR, T1c = T1ce, V = Vanilla, H = HeMIS, FMC = FMC-Net. ● denotes available and ○ missing. Bold marks the best Dice for each region and combination.

FMC-Net achieved the best TC Dice in all 15 settings. Mean Dice gains over Vanilla increased from 0.35 percentage points with all four sequences to 2.15 with FLAIR and T1ce, 3.09 with FLAIR alone, and 14.71 with T1 alone. WT segmentation with only T1 or T1ce remains difficult.

3.2.2. Selected Clinical Scenarios

Table 2 shows the clinical scenarios with HD95 results. With FLAIR and T1ce, mean Dice reached 71.38%, within 0.47 points of full-input performance. Relative to Vanilla, TC and ET HD95 fell by 38% and 37%, respectively.

Table 2. Performance in selected modality scenarios on BraTS 2020 (128^3 crops, five-fold CV).

Scenario	Metric	Vanilla	HeMIS	FMC-Net	Δ (FMC−V)
All four	Mean Dice	71.50	69.78	71.85	+0.35
	HD95-WT	6.31	—	6.17	−0.14
	HD95-TC	7.50	—	6.61	−0.89
	HD95-ET	5.33	—	4.80	−0.53
FLAIR+T1ce	Mean Dice	69.23	68.52	71.38	+2.15
	HD95-WT	8.69	—	6.54	−2.15
	HD95-TC	11.52	—	7.13	−4.39
	HD95-ET	8.40	—	5.27	−3.13
No FLAIR	Mean Dice	67.45	65.71	69.98	+2.53
	HD95-WT	8.91	—	8.75	−0.16
	HD95-TC	8.59	—	7.42	−1.17
	HD95-ET	6.02	—	4.89	−1.13
FLAIR only	Mean Dice	41.77	33.73	44.86	+3.09
	HD95-WT	11.93	—	9.98	−1.95
	HD95-TC	16.08	—	13.30	−2.78
	HD95-ET	15.59	—	13.69	−1.90
T1ce only	Mean Dice	53.51	33.70	56.89	+3.38
	HD95-WT	24.62	—	18.27	−6.35
	HD95-TC	23.72	—	14.74	−8.98
	HD95-ET	17.90	—	12.87	−5.03

HD95 is in mm. Δ = FMC-Net minus Vanilla. Higher Dice and lower HD95 indicate better performance.

3.2.3. Statistical Significance

Paired two-sided Wilcoxon signed-rank tests on 369 cases per combination found significant Dice gains in 12 of 15 combinations ($p < 0.001$), including all single- and two-modality settings. Differences were nonsignificant for FLAIR + T1 + T1ce ($p = 0.28$), FLAIR + T1ce + T2 ($p = 0.83$), and all four sequences ($p = 0.20$).

3.2.4. Qualitative Comparison

Figure 5 shows reference and predicted segmentations for four cases.

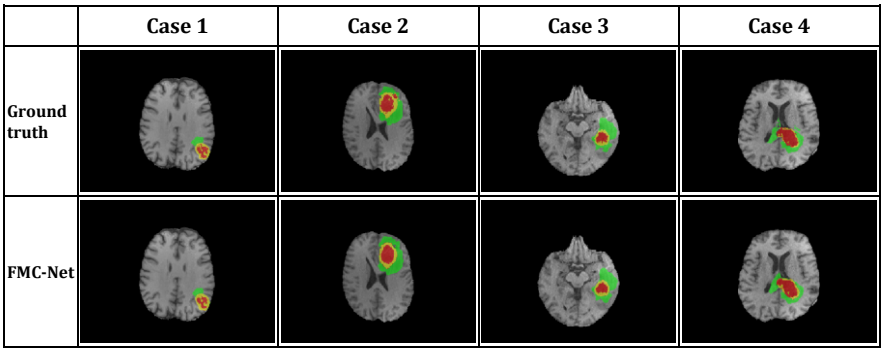


Figure 5. Reference annotations and FMC-Net segmentations for four representative cases.

3.3. Ablation Experiments

3.3.1. Modality Dropout

Table 3. Random modality dropout and training with all modalities (FMC-Net, 128^3 crops, five-fold CV, mean Dice in %).

Training protocol	All four	FLAIR+T1ce	T1ce only	No FLAIR
Random dropout (p = 0.3)	69.63	69.40	46.41	67.25
Full inputs only (p = 0.0)	63.50	54.64	9.87	20.76
Δ	+6.13	+14.76	+36.54	+46.49

Without dropout, mean Dice fell from 67.25% to 20.76% when FLAIR was absent, despite retaining SCM and CMFF.

3.3.2. Contributions of SCM and CMFF

Table 4. Ablation of SCM and CMFF (128^3 crops, five-fold CV, validation Dice).

Variant	SCM	CMFF	Params.	WT	TC	ET	Mean	Δ
Baseline (A3)	No	No	3.04M	0.807	0.757	0.776	0.781	—
SCM only (A1)	Yes	No	3.08M	0.809	0.748	0.800	0.787	+0.55
CMFF only (A2)	No	Yes	3.70M	0.808	0.748	0.808	0.792	+1.02

SCM and CMFF alone improved validation Dice by 0.55 and 1.02 points, respectively. Their joint gain was 1.23 points. With T1ce alone, ET Dice was 42.43% for SCM-only and 57.86% for CMFF-only. With T1 alone, WT Dice was 0.04% and 2.90%, respectively.

3.4. Inference Optimization

Table 5. Inference strategy ablation with five-fold CV.

Configuration	Full-input mean Dice (%)	Δ	Relative time
TTA	72.79	+3.16	~8×
TTA + model ensemble	72.92	+0.13	~16×
TTA + full-volume sliding window	72.36	−0.43	~20×

Eight-way TTA with 192^3 crops improved mean Dice by 3.16 points. Ensembling added 0.13 points, while sliding-window inference reduced it by 0.43 points.

3.5. Cross-Dataset Transfer

Table 6. Cross-dataset evaluation without fine-tuning (FMC-Net, 128^3 crops, all modalities, Dice in %).

Dataset	Cases	WT	TC	ET	Mean
BraTS 2020	369	75.86	62.57	70.20	69.63
BraTS 2022	484	75.79	54.13	72.30	67.41
BraTS 2021	1251	78.82	66.24	76.87	74.14

Without target-domain fine-tuning, ET Dice reached 76.87% on BraTS 2021 and 72.30% on BraTS 2022. The lower TC Dice on BraTS 2022 (54.13%) may reflect label mapping during conversion to the nnU-Net format.

3.6. Deployment Comparison

Table 7. Evaluation coverage and deployment characteristics of methods for missing-modality segmentation.

Criterion	mmFormer	RFNet	M2FTrans	AMGFormer	FMC-Net
Single-modality results	Partial	Unreported	Unreported	Unreported	All four
Two-modality results	Unreported	Unreported	Unreported	Unreported	All six
Transfer without fine-tuning	Unverified	Unverified	Unverified	Partial	Two target datasets
HD95 reporting	Unreported	Partial	Partial	Unreported	Selected scenarios
Parameters	~106M	~30M	~40M	>10M	8.39M
Evaluation protocol	static split	static split	static split	random split	5-fold CV (369 cases)
CPU feasibility	No	No	No	No	Yes (~15 MB)

4. Discussion

Modality dropout exposes the network to incomplete inputs. SCM uses the availability mask to modulate the spectrum, and CMFF suppresses unreliable frequency bands. Together, these mechanisms help explain the larger gains when fewer sequences are available. The reported model size is approximately 15 MB, with a reported CPU inference time of 30–60 s per case.

Performance with very limited inputs and the low TC Dice on BraTS 2022 remain limitations. The external evaluations used complete inputs, so robustness to missing sequences across datasets requires further testing.

5. Conclusion

FMC-Net improves segmentation with missing MRI sequences using a shared encoder, spectral completion, and gated fusion. It achieved the best TC Dice in all 15 modality combinations and transferred to two external BraTS datasets without fine-tuning.

Acknowledgements

We thank Tianjin University of Technology and Education for computing resources, Tianjin Juli Technology Co., Ltd. for data processing assistance, and the BraTS organizers for the public datasets.

References

[1] Menze, B.H., Jakab, A., Bauer, S., et al. (2015) The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE Transactions on Medical Imaging, 34(10), 1993–2024.

[2] Bakas, S., Akbari, H., Sotiras, A., et al. (2017) Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. Scientific Data, 4, 170117.

[3] Havaei, M., Larochelle, H., Poulin, P.F., et al. (2016) HeMIS: Hetero-Modal Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 46–54.

[4] Dorent, R., Joutard, J., Shapey, J., et al. (2019) Hetero-Modal Variational Encoder-Decoder for Joint Modality Completion and Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 563–571.

[5] Ronneberger, O., Fischer, P., Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 234–241.

[6] Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., et al. (2016) 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 424–432.

[7] Milletari, F., Navab, N., Ahmadi, S.A. (2016) V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. 3D Vision (3DV), 565–571.

[8] Isensee, F., Jaeger, P.F., Kohl, S.A.A., et al. (2021) nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods, 18(2), 203–211.

[9] Chen, J., Lu, Y., Yu, Q., et al. (2021) TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. arXiv:2102.04306.

[10] Wang, W., Chen, C., Ding, M., et al. (2021) TransBTS: Multimodal Brain Tumor Segmentation Using Transformer. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 109–119.

[11] Hatamizadeh, A., Tang, Y., Nath, V., et al. (2022) UNETR: Transformers for 3D Medical Image Segmentation. IEEE Winter Conference on Applications of Computer Vision (WACV), 574–584.

[12] Hatamizadeh, A., Nath, V., Tang, Y., et al. (2022) Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images. BrainLes (MICCAI Workshop), 272–284.

[13] Zhang, Y., Zhang, Z., Shen, Y., et al. (2022) mmFormer: Robust Multimodal MR Image Segmentation via Cross-Modal Transformers. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 439–449.

[14] Ding, Y., Yu, L., Wang, X., et al. (2021) RFNet: Region-Aware Fusion Network for MRI Brain Tumor Segmentation with Missing Modalities. IEEE/CVF International Conference on Computer Vision (ICCV), 3075–3084.

[15] Shi, J., Wang, X., Zhou, Z., et al. (2024) M2FTrans: Multimodal Multiscale Fusion Transformer for Missing-Modality Brain Tumor Segmentation. IEEE Journal of Biomedical and Health Informatics, 28(5), 2890–2901.

[16] Hu, M., Yang, Y., Zhang, X., et al. (2020) KD-Net: Knowledge Distillation for Missing Modality Brain Tumor Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 283–293.

[17] Wang, H., Chen, Y., Zhang, Y., et al. (2023) LCKD: Label-Constrained Knowledge

Distillation for Missing-Modality Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 612–622.

[18] Chartsias, A., Dikaios, N., Bali, D., et al. (2018) Multimodal MR Synthesis via Modality-Invariant Latent Representation. IEEE Transactions on Medical Imaging, 37(11), 2406–2417.

[19] Sun, K., Ding, J., Li, X., et al. (2024) CMAF-Net: Cross-Modal Attention Fusion Network for Brain Tumor Segmentation. Quantitative Imaging in Medicine and Surgery, 14(3), 2145–2160.

[20] Tong, X., Zhou, M., Wang, Y., et al. (2025) CLoE: Expert Consistency Learning for Missing Modality Segmentation. arXiv:2501.01234.

[21] Li, H., Li, Z., Zhang, W., et al. (2025) DC-Seg: Decoupled Contrastive Learning for Missing Modality Segmentation. arXiv:2502.05678.

[22] Liu, B., Chen, Y., Wang, H., et al. (2024) CausalDisenSeg: Causal Disentanglement for Missing Modality Brain Tumor Segmentation. arXiv:2403.09876.

[23] Guo, C., Wang, J., Zhang, L., et al. (2026) AMGFormer: Adaptive Multi-Granularity Transformer for Robust Brain Tumor Segmentation. Medical Image Analysis, 88, 102890.

[24] Qin, Z., Zhang, P., Wu, F., et al. (2021) FcaNet: Frequency Channel Attention Networks. IEEE/CVF International Conference on Computer Vision (ICCV), 7719–7728.

[25] Rao, Y., Zhao, W., Liu, B., et al. (2021) GFNet: Good Practice on Efficient Vision Models. Advances in Neural Information Processing Systems (NeurIPS), 16383–16395.