VEP-AGUNet3D: Vascular-Enhancement Prior Improves Post-Treatment Glioma Segmentation

Fizza Hassan*, Mufti Mahmud*†‡
* ICS Department, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia
† SDAIA-KFUPM JRC for AI, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia
‡ IRC for Bio Systems and Machines, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia
Emails: g202318310@kfupm.edu.sa, mufti.mahmud@kfupm.edu.sa, muftimahmud@gmail.com

***Abstract*—Accurate segmentation of residual or recurrent tumour in post-treatment glioma MRI remains challenging due to resection cavities, treatment-related changes, and heterogeneous enhancement patterns. While deep learning methods achieve strong performance on pre-operative benchmarks, their performance often degrades in post-treatment settings, particularly for small and irregular residual disease. In this work, we introduce a vascular enhancement prior (VEP), a lightweight auxiliary channel derived automatically from routine multiparametric MRI, which emphasizes enhancement- and vessel-like patterns while suppressing edema- and fluid-dominated regions. VEP is computed from T1, T1-Gd, T2, and FLAIR images and integrated into an attention-gated 3D U-Net without modifying the backbone architecture, loss function, or inference procedure. We evaluate three input configurations: a 3-channel structural baseline (T1, T1-Gd, FLAIR), a 4-channel structural baseline (T1, T1-Gd, T2, FLAIR), and a 3-channel+VEP configuration in which VEP is used in place of raw T2. Experiments on the BraTS 2024 post-treatment glioma dataset show that the VEP-based model achieves the best overall external-cohort performance among the evaluated configurations, improving both overlap and boundary accuracy relative to the structural baselines. In the corrected external evaluation, the raw 4-channel baseline performs comparably to the 3-channel baseline, while the VEP-based model achieves the strongest overall results. These findings suggest that structured integration of modality-derived information can be beneficial for robust post-treatment glioma segmentation.**

***Index Terms*—post-treatment glioma segmentation, vascular enhancement prior, attention-gated 3D U-Net, multi-parametric MRI, training stochasticity, patch-based learning**

I. Introduction

Glioblastoma multiforme (GBM) remains associated with poor prognosis despite maximal safe resection and concurrent chemoradiotherapy [1]. In the post-treatment setting, accurate delineation of residual or recurrent tumour on MRI is clinically important but difficult: resection cavities, haemorrhage, treatment-related change, and postoperative enhancement can mimic tumour and obscure subtle disease along cavity margins [2], [3]. This makes manual contouring time-consuming, expertise-dependent, and variable, motivating robust automated segmentation methods.

Automated glioma segmentation has progressed substantially on pre-operative MRI, largely driven by BraTS-style benchmarks [4]. Encoder–decoder backbones such as U-Net [5], attention-gated variants, and self-configuring frameworks such as nnU-Net [6] achieve strong performance in pre-operative settings. However, performance typically degrades in post-operative and post-treatment cohorts, especially for small, irregular, and fragmented residual tumour near resection margins [2], [3]. This motivates approaches that are more robust to treatment-related appearance changes and more selective in how they use multimodal MRI information.

Most BraTS-style segmentation methods use early fusion by concatenating structural MRI modalities as input channels [4], [6], [7]. The BraTS 2024 post-treatment glioma dataset provides multi-institutional mpMRI (T1, T1-Gd, T2, and FLAIR) with harmonised tumour-related labels, enabling systematic evaluation in treated cases [8]. While prior work suggests that adding raw modalities does not always yield clear gains in post-operative segmentation [2], [3], it remains plausible that carefully designed derived representations may provide more useful information than naive early fusion alone. In particular, subtraction maps, synthesis-based representations, and other auxiliary channels have been shown in related medical imaging tasks to provide complementary signal beyond the raw structural inputs [9]–[11].

In this work, we test whether a deliberately engineered *vascular enhancement prior* (VEP) improves post-treatment glioma segmentation. VEP is computed automatically from routine mpMRI by estimating enhancement from T1-Gd relative to T1 and suppressing regions dominated by edema or fluid using T2 and FLAIR, producing an auxiliary channel intended to emphasize enhancement- and vessel-like signal without requiring additional acquisitions or manual annotations. We integrate VEP into an attention-gated 3D U-Net (AGUNet3D) and compare it against matched structural baselines using 3-channel input (T1, T1-Gd, FLAIR) and 4-channel input (T1, T1-Gd, T2, FLAIR).

Our contributions are:

- We introduce a lightweight, fully automatic vascular enhancement prior (VEP) derived from routine mpMRI to emphasize enhancement- and vessel-like patterns relevant to residual tumour after treatment.
- We integrate VEP into an attention-gated 3D U-Net without changing the backbone architecture, loss, or inference protocol, enabling controlled comparison against structural baselines.
- We show that the VEP-based model achieves the strongest overall external-cohort performance among the evaluated configurations, outperforming both 3-channel

and 4-channel structural baselines in combined overlap and boundary accuracy.
- We provide an interpretable example of how structured, domain-informed priors can complement raw multimodal MRI in the challenging post-treatment setting.

II. Related Work

Recent literature reports the widespread use of Artificial Intelligence (AI) across various domains, including healthcare. Examples include improved COVID-19 detection from chest X-rays through optimisation [12], transfer learning [13], segmentation, and one-shot learning [14], securing smart healthcare environments [15], fall monitoring [16], and Industry 5.0-enabled smart education [17]. Within the context of brain tumour detection, convolutional encoder–decoder architectures such as U-Net [5] and the self-configuring nnU-Net framework [6] have established strong baselines for brain tumour segmentation on pre-operative MRI, in particular on BraTS-style benchmarks [4]. These models exploit multi-scale context and skip connections to recover fine anatomical detail, and have been extended with attention mechanisms that suppress irrelevant background structures and highlight tumour-relevant regions [18]. Across successive BraTS challenges, such encoder–decoder and attention-based variants have consistently achieved state-of-the-art performance when trained on the standard four structural MRI modalities.In particular, Zhang et al. introduced Attention Gate U-Net (AGU-Net) and AGResU-Net with attention units on the skip connections and demonstrated consistent improvements over vanilla U-Net and ResU-Net on pre-operative BraTS benchmarks [19].

In contrast, early post-operative imaging remains considerably more challenging for automated segmentation. Resection cavities, blood products, and post-surgical inflammatory changes introduce complex signal patterns that can resemble tumour and degrade both overlap and boundary-based metrics [2], [3]. Helland et al. combined a patch-based nnU-Net with a full-volume attention-gated U-Net for early post-operative glioblastoma and reported substantially lower performance than is typical on pre-operative BraTS cohorts, particularly for small, irregular residual tumour along the cavity wall [3]. Other post-operative and post-treatment studies, including transfer-learning approaches on limited cohorts [20], joint pre-/post-operative pipelines [21], and patient-specific fine-tuning strategies [22], further highlight both the difficulty of this setting and the scarcity of large, standardised datasets. Recent extensions of BraTS toward treated glioma, such as Sørensen et al. [23] and the BraTS-GLI post-treatment challenge [8], begin to address this gap by providing multi-institutional post-operative and post-radiotherapy cohorts with harmonised tumour and cavity annotations.

Most existing approaches for BraTS-style tumour segmentation adopt an early-fusion strategy in which the available MRI sequences are concatenated as input channels and processed by a single encoder, as in U-Net, nnU-Net, and many BraTS challenge winning entries [5], [6]. Although some works explore reduced modality setups or alternative fusion schemes, systematic ablations of modality combinations in the post-operative GBM setting remain rare. In early post-operative glioblastoma segmentation, Helland et al. used only T1w and T1w-CE as input, noting that adding more sequences had been shown in prior work not to improve segmentation performance in that setting [2], [3]. At the same time, several studies in medical image segmentation have shown that derived or synthetic channels (for example, subtraction images, distance maps, or auxiliary organ-at-risk masks) can provide complementary information beyond the raw modalities [10], [11], [24].

Recent work has highlighted that segmentation performance is influenced not only by model design but also by stochastic factors in the training pipeline, including random initialization, data ordering, sampling, and augmentation. Åkesson et al. systematically quantified random effects in deep learning based medical image segmentation by repeating training with different seeds and showed that performance variability across runs can be non negligible even when architecture and hyperparameters are fixed [25]. In patch based training, additional variance can arise from the patch sampling policy and inference strategy, which can bias what the model observes during optimization if sampling trajectories are narrow or effectively repeated [26]. Practical pipelines therefore rely on stochastic sampling and augmentation to increase exposure to rare patterns and hard negatives; toolkits such as TorchIO formalize patch based sampling and augmentation for medical images [27]. Data augmentation has also been shown to improve robustness and generalisation in medical segmentation, including realistic adversarial augmentation schemes [28], and systematic reviews report consistent benefits across tasks and modalities [29]. For domain generalization, augmentation strategies that simulate distribution shift, for example via frequency domain perturbations, further support robustness on unseen domains [30]. These findings motivate explicitly considering and reporting controlled randomization when developing segmentation methods for heterogeneous post treatment imaging.

Overall, prior work shows that post-treatment glioma segmentation remains challenging and that robustness depends on both model design and training choices. However, to our knowledge, no previous study has introduced a deliberately engineered prior channel derived from routine MRI specifically for post-treatment glioma segmentation and evaluated it against matched structural baselines in this setting. In this work, we address this gap by integrating a lightweight vascular-enhancement prior (VEP) derived from routine BraTS sequences (T1, T1-Gd, T2, and FLAIR) into an attention-gated 3D U-Net backbone and evaluating it against matched structural baselines on heterogeneous post-treatment imaging.

III. Methodology

This section describes the BraTS–GLI data and binary label definition, preprocessing, the construction of the vascular enhancement prior (VEP), the AGUNet3D backbone, and the training and evaluation protocols.

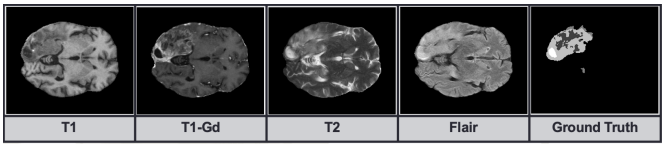

Fig. 1. Example BraTS–GLI post-treatment case showing T1, T1-Gd, T2, and FLAIR, and the corresponding tumour ground truth.

A. Data and Label Definition

We use the BraTS 2024 post-treatment glioma cohort (BraTS–GLI) [8], which provides multi-institutional mpMRI for treated glioma patients. Each case includes four structural modalities: native T1-weighted (T1), post-contrast T1-weighted (T1-Gd), T2-weighted (T2), and T2-FLAIR (FLAIR). The dataset release contains both labelled and unlabelled cases. This study uses only the labelled portion for supervised learning. In our experiments, the original multi-class annotations (tumour subregions and cavity) are converted into a binary tumour *super-region* by merging all tumour-related labels into a single foreground class and treating all remaining voxels, including the resection cavity, as background. Figure 1 shows an example case.

For model development, we follow the dataset release structure by using the labelled training cohort together with a separately released labelled cohort for generalisation testing [8]. Specifically, for the main experiments we split 1350 labelled cases subject-wise into training, validation, and internal test sets using a 70%/15%/15% partition (944/203/203). The validation set is used for early stopping and VEP variant selection, while the internal test set is reserved for unbiased internal evaluation. We additionally evaluate on an *additional held-out cohort* of 271 labelled cases that is not used for training, hyperparameter tuning, or model selection and is used only for final evaluation.

B. Preprocessing

We use the BraTS–GLI volumes in NIfTI format as provided by the organizers, including the standardized preprocessing distributed with the release (e.g., consistent spatial alignment and brain-focused volumes) [8]. All modalities and labels are reoriented to a consistent right–anterior–superior (RAS) convention. Each input modality is normalized independently per volume by clipping intensities to the 0.5th and 99.5th percentiles and linearly rescaling to $[-1, 1]$ (with final clipping), which reduces inter-scan variability while preserving within-scan contrast.

C. Vascular Enhancement Prior (VEP)

We introduce a derived channel that emphasises enhancement-driven and vessel-like patterns relevant for residual tumour after treatment. This *vascular enhancement prior* (VEP) is computed offline from the four structural modalities (T1, T1-Gd, T2, FLAIR) and saved as an additional NIfTI volume per case. During training and inference, VEP is loaded as an input modality; in our main configuration, VEP replaces raw T2 so the network input consists of T1, T1-Gd, FLAIR, and VEP. Figure 2 summarises the pipeline.

TABLE I
VEP WEIGHT VARIANTS USED FOR ABLATION.

VEP ID	Description	w_E	w_S	w_R
W00	Balanced fusion	0.60	0.10	0.15
W01	No structural cue	0.60	0.00	0.15
W02	No suppression	0.60	0.10	0.00
W03	Enhancement only	0.60	0.00	0.00
W04	Lower enhancement weight	0.55	0.10	0.15
W05	Higher enhancement weight	0.65	0.10	0.15
W06	Lower structural weight	0.60	0.05	0.15
W07	Higher structural weight	0.60	0.15	0.15
W08	Lower suppression weight	0.60	0.10	0.10
W09	Higher suppression weight	0.60	0.10	0.20
W10	Enhancement-heavy	0.75	0.05	0.15
W11	Structural-heavy	0.55	0.25	0.15

a) Input normalisation: For VEP construction, each modality is independently normalised to $[0, 1]$ via percentile clipping (2nd–98th) followed by linear rescaling.

b) Enhancement map: Enhancement is estimated from T1-Gd relative to T1 as

$$E = \max\big(I_{\text{T1-Gd}} - I_{\text{T1}},\, 0\big), \tag{1}$$

followed by 3D Gaussian smoothing with $\sigma = 1.0$.

c) Structural cue: A sparse cue S is obtained by thresholding non-zero voxels of E at the 85th percentile and removing connected components smaller than 5 voxels.

d) Edema/fluid suppression: We form

$$C = 0.5\, I_{\text{T2}} + 0.5\, I_{\text{FLAIR}}, \tag{2}$$

min–max normalise C to $[0, 1]$, and invert:

$$R = 1 - \text{norm}(C). \tag{3}$$

e) Coarse brain mask: A coarse mask M is estimated from T1-Gd using a low-percentile threshold (5th) followed by two erosions and three dilations.

f) Fusion: The final VEP is

$$I_{\text{VEP}} = \text{norm}\big(w_E E + w_S S + w_R R\big) \odot M, \tag{4}$$

where $\odot$ denotes element-wise multiplication.

g) Weight ablation: We treat (w_E, w_S, w_R) as hyperparameters and generate 12 deterministic VEP variants per case (Table I). Each variant is saved as a separate NIfTI and used to train a separate model under identical settings.

D. Model Architecture

All experiments use the same attention-gated 3D U-Net backbone (AGUNet3D), implemented as a symmetric four-level encoder–decoder with attention gates on the three deepest skip connections and an ungated shallow skip for preserving fine detail. The network is a symmetric encoder–decoder with four resolution levels, skip connections, and attention gates on the three deepest skips to suppress irrelevant activations while preserving fine detail via an ungated shallow skip.

Each encoder and decoder block contains two $3 \times 3 \times 3$ convolutions, each followed by BatchNorm3d and ReLU.

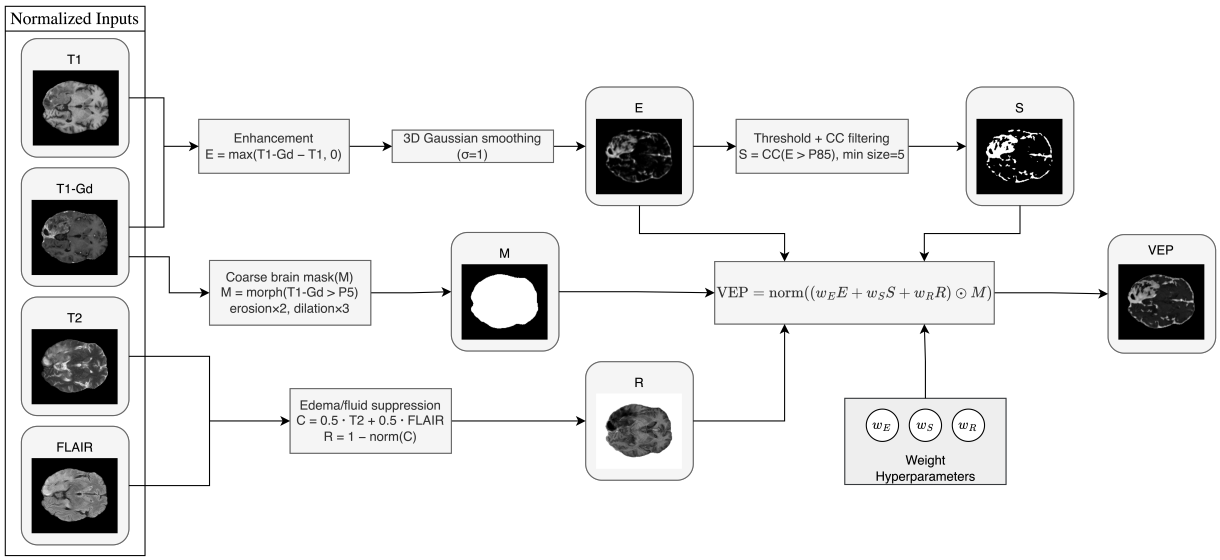

Fig. 2. VEP construction from T1, T1-Gd, T2, and FLAIR. Enhancement E, structural cue S, suppression R, and mask M are fused as $\mathrm{VEP} = \mathrm{norm}(w_E E + w_S S + w_R R) \odot M$.

Downsampling uses $2 \times 2 \times 2$ max-pooling, and upsampling uses transposed convolutions. A final $1 \times 1 \times 1$ convolution outputs two logits per voxel (background vs. tumour super-region). Convolutional layers use Kaiming normal initialisation.

We evaluate three input configurations: *AGUNet3D–3ch* (T1, T1-Gd, FLAIR), *AGUNet3D–4ch* (all four structural modalities), and *AGUNet3D–3ch+VEP* (T1, T1-Gd, FLAIR, VEP). In the implemented 4-channel configuration, the channel order is kept consistent between training and inference. At inference, voxel-wise probabilities are obtained via softmax and thresholded at 0.5 to form a binary prediction.

E. Training Strategy

We train AGUNet3D in a supervised, patch-based setting using binary tumour super-region masks. To mitigate foreground–background imbalance, we employ a combined Dice and focal loss (focal $\gamma = 2$ with equal weighting). Foreground-biased patch sampling is used while retaining background patches as hard negatives.

Data augmentation is applied on-the-fly and includes random flips, random 90° rotations, and additive Gaussian noise. Optimisation uses AdamW (learning rate 5×10^{-4}, weight decay 10^{-5}) with mixed precision and gradient-norm clipping (max norm 1.0). We select the checkpoint with the best validation Dice and apply early stopping if validation Dice does not improve by at least 10^{-4} for 25 consecutive epochs.

All experiments were carried out on a workstation equipped with three NVIDIA GeForce RTX 3090 GPUs, each with 24 GB of GDDR6X memory. Training was performed using CUDA version 13.0 with NVIDIA driver version 580.95.05. All reported models were trained under the same training protocol to ensure a fair comparison across input configurations and VEP variants.

F. Evaluation and Benchmarks

Inference uses sliding-window prediction with window size 96^3 and overlap 0.5. We report Dice for all evaluations and the 95th percentile Hausdorff distance (HD95) for the binary tumour super-region.

a) Evaluation metrics: For a predicted binary mask P and ground-truth mask G, the Dice similarity coefficient (DSC) is defined as

$$\mathrm{Dice}(P, G) = \frac{2|P \cap G|}{|P| + |G|}, \tag{5}$$

where $|\cdot|$ denotes the number of foreground voxels.

Boundary accuracy is evaluated using the 95th percentile Hausdorff distance (HD95), which is more robust to outliers than the maximum Hausdorff distance. Let ∂P and ∂G denote the sets of surface voxels of the prediction and ground truth, respectively. The directed surface distance from ∂P to ∂G is defined as

$$d(\partial P, \partial G) = \{ \min_{y \in \partial G} \|x - y\|_2 \mid x \in \partial P \}, \tag{6}$$

where $\|\cdot\|_2$ denotes the Euclidean distance. The HD95 is then computed as

$$\mathrm{HD95}(P, G) = \max \Big(\mathrm{percentile}_{95}\big(d(\partial P, \partial G)\big), \; \mathrm{percentile}_{95}\big(d(\partial G, \partial P)\big) \Big). \tag{7}$$

Lower HD95 values indicate better boundary agreement. For external held-out cohort evaluation, HD95 is computed in millimetres using voxel spacing.

IV. Results

We evaluate AGUNet3D on the BraTS–GLI post-treatment glioma dataset using the Dice similarity coefficient (Dice) and

TABLE II
ABLATION STUDY OF VEP WEIGHT CONFIGURATIONS EVALUATED ON THE VALIDATION SET. DICE IS MAXIMISED AND HD95 IS MINIMISED.

VEP ID	Dice ↑	HD95 ↓
W04	**0.873**	5.37
W09	0.871	5.76
W00	0.871	5.76
W08	0.869	**5.29**
W06	0.868	5.69
W11	0.867	5.36
W01	0.867	5.87
W10	0.867	5.42
W07	0.866	5.72
W03	0.862	5.81
W05	0.861	6.28
W02	0.853	6.26

TABLE III
PERFORMANCE OF AGUNET3D MODELS ON THE EXTERNAL HELD-OUT COHORT.

Model	Input Modalities	Dice ↑	HD95 ↓
AGUNet2D-3ch (Literature) [2]	T1, T1-Gd, FLAIR	0.4978	14.56
AGUNet3D-3ch	T1, T1-Gd, FLAIR	0.7887 ± 0.2774	14.18 ± 21.47
AGUNet3D-4ch	T1, T1-Gd, T2, FLAIR	0.7934 ± 0.2685	15.98 ± 21.89
AGUNet3D-3ch+VEP (W04)	T1, T1-Gd, FLAIR, VEP	**0.8032 ± 0.2723**	**12.73 ± 19.92**

the 95th percentile Hausdorff distance (HD95) for the binary tumour super-region.

A. Vascular Enhancement Prior (VEP) Ablation

Table II summarises the ablation study over 12 VEP weight configurations on the validation set. For each configuration, we report the best validation Dice achieved during training and the corresponding HD95 value at the same epoch.

The baseline balanced configuration (W00) achieved a Dice of 0.871 and an HD95 of 5.76. The highest validation Dice was obtained with W04 (lower enhancement weight), which achieved a Dice of 0.873 and an HD95 of 5.37. The lowest HD95 was obtained with W08 (lower suppression weight), at 5.29. Based on validation Dice, W04 was selected for subsequent evaluation on the external held-out cohort.

B. External Held-Out Cohort: Comparison with Baseline Models

Table III reports performance on the external held-out cohort. For the AGUNet3D models, values correspond to the seed-42 runs. Dice is reported as mean ± standard deviation across all 271 external cases, while HD95 is reported as mean ± standard deviation over cases with defined HD95. In addition to the standard 3-channel and 4-channel AGUNet3D variants, we include the best-performing VEP configuration selected from the validation ablation study (W04). A representative 2D result from the literature is included for contextual reference only and should not be interpreted as a direct numerical comparison because of differences in architecture dimensionality, data splits, and evaluation protocols.

On the external held-out cohort, AGUNet3D-3ch+VEP (W04) achieved the best overall performance, with a Dice of 0.8032 ± 0.2723 and an HD95 of 12.73 ± 19.92 mm. AGUNet3D-4ch followed closely, with a Dice of 0.7934 ± 0.2685 and an HD95 of 15.98 ± 21.89 mm, while AGUNet3D-3ch achieved a Dice of 0.7887 ± 0.2774 and an HD95 of 14.18 ± 21.47 mm.

V. DISCUSSION

The ablation results indicate that VEP performs best when enhancement, structural cues, and suppression are used together. Removing any one of these components reduces performance, with the clearest drop observed when suppression is removed. This is consistent with the challenges of post-treatment MRI, where residual tumour often appears alongside edema, fluid, and cavity-related changes with overlapping intensities. In this setting, the suppression term appears to play an important role by reducing the influence of these confounding regions, while the enhancement and structural terms help preserve tumour-relevant signal.

On the external held-out cohort, the VEP-based model achieved the strongest overall performance among the evaluated configurations. In contrast, the raw 4-channel baseline performed comparably to the 3-channel baseline in Dice, but did not match the VEP-based model in overall performance. This suggests that simply adding an additional structural modality is not sufficient to guarantee better segmentation in the post-treatment setting, whereas a structured derived representation can provide more useful guidance.

A plausible explanation is that VEP encourages the network to focus on enhancement-relevant and vessel-like patterns while suppressing fluid- and edema-dominated regions that are common after treatment. Qualitative inspection of difficult cases further suggested that VEP often acts as a selective prior, reducing spurious tumour expansion into post-treatment confounding regions. At the same time, this selectivity may reduce sensitivity in very small lesions or in cases where the binary tumour label extends beyond the compact enhancement-focused signal emphasized by the prior. Thus, the main benefit of VEP may be improved selectivity rather than uniformly higher sensitivity across all cases.

Additional runs with different random seeds suggested the same broad pattern. Relative to the structural baselines, the VEP-based model remained competitive in overlap accuracy and showed stronger overall boundary performance. Although the Dice improvement over the 3-channel baseline was modest, the combined Dice and HD95 results support the practical value of the prior.

All models were trained under the same protocol, including foreground-biased patch sampling, on-the-fly augmentation, and standard random initialization, to ensure fair comparison across input configurations. These elements were included to improve exposure to diverse spatial contexts and appearance variations in heterogeneous post-treatment data. However, because their individual effects were not isolated in a separate matched ablation, they should be interpreted here as part of the common training procedure rather than as an independently validated source of improvement.

VI. Conclusion

This paper presents a vascular enhancement prior (VEP) for post-treatment glioma segmentation and integrates it into an attention-gated 3D U-Net framework. By explicitly encoding enhancement-driven information, structural cues, and edema suppression derived from routine multi-parametric MRI, the proposed approach improves overall performance relative to the evaluated structural baselines. The results support the view that integrating modality-derived information through a structured prior can be beneficial in the challenging post-treatment setting.

Several limitations remain. The VEP is constructed using fixed operations and manually defined weights, which may not generalise optimally across all imaging protocols or post-treatment appearances. In addition, the evaluation focuses on binary tumour super-region segmentation and does not address finer subregion delineation or longitudinal disease progression. The case-level analysis also suggests a trade-off: while VEP can improve selectivity in complex post-treatment regions, it may be overly conservative in some small or subtle lesions. Future work will explore adaptive or learnable formulations of the prior, extension to multi-class segmentation, and incorporation of longitudinal information to further improve robustness in post-treatment glioma analysis.

References

[1] R. Stupp *et al.*, "Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma," *New England Journal of Medicine*, vol. 352, no. 10, pp. 987–996, 2005.

[2] R. H. Helland, D. Bouget, R. S. Eijgelaar, P. C. De Witt Hamer, F. Barkhof, O. Solheim, and I. Reinertsen, "Glioblastoma segmentation from early post-operative mri: Challenges and clinical impact," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, vol. 15006 of *Lecture Notes in Computer Science*, pp. 284–294, 2024.

[3] R. H. Helland *et al.*, "Segmentation of glioblastomas in early post-operative multi-modal mri with deep neural networks," *Scientific Reports*, vol. 13, p. 19751, 2023.

[4] B. H. Menze *et al.*, "The multimodal brain tumour image segmentation benchmark (brats)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015.

[6] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnu-net: A self-configuring method for deep learning–based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, 2021.

[7] F. Isensee, P. F. Jäger, P. M. Full, P. Vollmuth, and K. H. Maier-Hein, "nnu-net for brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12659, pp. 118–132, Springer, 2021.

[8] "Brats 2024 post-treatment glioma dataset." https://www.synapse.org/Synapse:syn53708249/wiki/626323, 2024. Accessed: Aug 1, 2025.

[9] B. M. Ellingson, "Contrast-enhanced T1-weighted digital subtraction for increased lesion conspicuity and quantifying treatment response in malignant gliomas," in *Glioma Imaging*, pp. 49–60, Springer, 2019.

[10] A. Chartsias, T. Joyce, M. V. Giuffrida, and S. A. Tsaftaris, "Multimodal mr synthesis via modality-invariant latent representation," in *International Conference on Information Processing in Medical Imaging (IPMI)*, pp. 371–383, 2017.

[11] W. Lei, H. Mei, Z. Sun, S. Ye, R. Gu, H. Wang, R. Huang, S. Zhang, S. Zhang, and G. Wang, "Automatic segmentation of organs-at-risk from head-and-neck ct using separable convolutional neural network with hard-region-weighted loss," *Medical Physics*, vol. 46, no. 10, pp. 4630–4639, 2019.

[12] A. K. Singh, A. Kumar, M. Mahmud, M. S. Kaiser, and A. Kishore, "COVID-19 Infection Detection from Chest X-Ray Images Using Hybrid Social Group Optimization and Support Vector Classifier," *Cogn. Comput.*, vol. 16, no. 4, pp. 1765 – 1777, 2024.

[13] N. Prakash, M. Murugappan, G. Hemalakshmi, M. Jayalakshmi, and M. Mahmud, "Deep transfer learning for COVID-19 detection and infection localization with superpixel based segmentation," *Sustain. Cities Soc.*, vol. 75, 2021.

[14] V. N. M. Aradhya, M. Mahmud, D. Guru, B. Agarwal, and M. S. Kaiser, "One-shot Cluster-Based Approach for the Detection of COVID–19 from Chest X–ray Images," *Cogn. Comput.*, vol. 13, no. 4, pp. 873 – 881, 2021.

[15] F. Farhin, M. S. Kaiser, and M. Mahmud, "Secured smart healthcare system: Blockchain and bayesian inference based approach," *Adv. Intell. Syst. Comput.*, vol. 1309, pp. 455 – 465, 2021.

[16] M. J. Al Nahian *et al.*, "Towards Artificial Intelligence Driven Emotion Aware Fall Monitoring Framework Suitable for Elderly People with Neurological Disorder," *Proc. BI 2020*, vol. 12241 LNAI, pp. 275 – 286, 2020.

[17] Y. Supriya, D. Bhulakshmi, S. Bhattacharya, T. R. Gadekallu, P. Vyas, R. Kaluri, S. Sumathy, S. Koppu, D. J. Brown, and M. Mahmud, "Industry 5.0 in Smart Education: Concepts, Applications, Challenges, Opportunities, and Future Directions," *IEEE Access*, vol. 12, pp. 81938 – 81967, 2024.

[18] D. Bouget, A. Pedersen, S. A. M. Hosainey, O. Solheim, and I. Reinertsen, "Meningioma segmentation in t1-weighted mri leveraging global context and attention mechanisms," *Frontiers in Radiology*, 2021.

[19] J. Zhang, Z. Jiang, J. Dong, Y. Hou, and B. Liu, "Attention gate resunet for automatic mri brain tumor segmentation," *IEEE Access*, vol. 8, pp. 58533–58545, 2020.

[20] M. Ghaffari, G. Samarasinghe, M. Jameson, F. Aly, L. Holloway, P. Chlap, E.-S. Koh, A. Sowmya, and R. Oliver, "Automated post-operative brain tumour segmentation: A deep learning approach," *Magnetic Resonance Imaging*, 2022.

[21] E. Lotan, B. Zhang, S. Dogra, W. D. Wang, D. Carbone, G. Fatterpekar, E. K. Oermann, and Y. W. Lui, "Development and practical implementation of a deep learning–based pipeline for automated pre- and post-operative glioma segmentation," *American Journal of Neuroradiology*, vol. 43, no. 1, pp. 24–32, 2022.

[22] S. Kundu, D. Toumpanakis, J. Wikstrom, R. Strand, and A. K. Dhara, "Atten-sevnetr for volumetric segmentation of glioblastoma and interactive refinement to limit over-segmentation," *IET Image Processing*, 2024.

[23] P. J. Sørensen, C. N. Ladefoged, V. A. Larsen, F. L. Andersen, M. B. Nielsen, H. S. Poulsen, J. F. Carlsen, and A. E. Hansen, "Repurposing the public brats dataset for postoperative brain tumour treatment response monitoring," *Tomography*, 2024.

[24] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle, "Brain tumor segmentation with deep neural networks," *Medical Image Analysis*, vol. 35, pp. 18–31, 2017.

[25] J. Åkesson, J. Töger, and E. Heiberg, "Random effects during training: Implications for deep learning-based medical image segmentation," *Computers in Biology and Medicine*, vol. 180, p. 108944, 2024.

[26] F. Madesta, R. Schmitz, T. Rösch, and R. Werner, "Widening the focus: Biomedical image segmentation challenges and the underestimated role of patch sampling and inference strategies," in *Bildverarbeitung für die Medizin 2021*, Informatik aktuell, p. 253, Springer Vieweg, 2021.

[27] F. Pérez-García, R. Sparks, and S. Ourselin, "Torchio: A python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning," *Computer Methods and Programs in Biomedicine*, vol. 208, p. 106236, 2021.

[28] C. Chen, C. Qin, C. Ouyang, Z. Li, S. Wang, H. Qiu, L. Chen, G. Tarroni, W. Bai, and D. Rueckert, "Enhancing mr image segmentation with realistic adversarial data augmentation," *Medical Image Analysis*, vol. 82, p. 102597, 2022.

[29] F. Garcea, A. Serra, F. Lamberti, and L. Morra, "Data augmentation for medical imaging: A systematic literature review," *Computers in Biology and Medicine*, vol. 152, p. 106391, 2023.

[30] Q. Qiao, W. Wang, M. Qu, K. Su, B. Jiang, and Q. Guo, "Medical image segmentation via single-source domain generalization with random amplitude spectrum synthesis," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024*, vol. 15009 of *Lecture Notes in Computer Science*, pp. 435–445, Springer, Cham, 2024.