Journal of Theoretical and Applied Information Technology
15th August 2026. Vol.104. No.15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

TOPOLOGY-AWARE DEEP LEARNING FOR ROBUST 3D MRI BRAIN TUMOR SEGMENTATION

DR.ARADDHANA [1] , DR SIVANEASAN BALA KRISHNAN[2], DR. SHRIKANT KULKARNI [3], DR. PRASUN CHAKRABARTI [4], JINU SOPHIA J [5], SILAS STEPHEN D [6], S.B.G.TILAK BABU[7]

[1] Professor and Former Director,School of CSIT, Symbiosis Skills and Professional University, Kiwale Pune, Post Foctoral Researcher, Singapore Institute of Technology, Singapore.
[2] Singapore Institute of Technology, Singapore.
[3] Adjunct Professor, Faculty of Business, Victorian Institute of Technology, Melbourne, Australia Adjunct Professor, Lincoln University College, Petaling Jaya, Selangor, Malaysia.
[4] Pro Vice Chancellor, Sir Padmapat Singhaniya University, Bhatewar, Rajasthan, India.
[5] Department of Computer Science and Engineering, Rajalakshmi Engineering College Thandalam, Chennai- 602105, India.
[6] Department of Electrical and Electronics Engineering, Panimalar Engineering College Bangalore Trunk Road, Poonamallee, Chennai, India.
[7] Department of ECE, Aditya University, Surampalem, India.

E-mail: [1]aadeshmukhskn@gmail.com, [2] sivaneasan@singaporetech.edu.sg, [3] srkulkarni21@gmail.com, [4] drprasun.cse@gmail.com, [5] jinusilas@gmail.com, [6] silasstephen@gmail.com, [7] . thilaksayila@gmail.com

ABSTRACT

The high morphological variability, indistinct tumor boundaries, and severe class imbalance in clinical imaging data pose challenges to accurate 3D brain tumor segmentation from multi-modal MRI. These issues are further exacerbated within the heterogeneous, multi-vendor, multi-field-strength imaging settings common in Indian tertiary-care hospitals, which is essential for neuro-oncological diagnosis and treatment planning. TopoSegNet is a novel Topology-aware Encoder-Decoder network that, for the first time, combines Betti number regularization based on persistent homology (PH) and a convolutional block attention module (CBAM)-based decoder to perform volumetric brain tumor segmentation. The persistent homology bottleneck explicitly filters the predicted segmentation to ensure that it preserves the topological invariants – connected components (β_0), loops (β_1) and voids (β_2) – of the ground-truth tumor mask, thereby decreasing the number of spurious connected components and hollow predictions by 38.7% compared to the best baseline model. Researcher also suggest a boundary-aware composite loss that combines Dice, cross-entropy loss, Hausdorff distance and topological persistence losses in a learnable manner. TopoSegNet is benchmarked on a novel 450-subject multi-institutional Indian dataset acquired across five premier hospitals, including five tumor grades and 1.5T and 3T scanner protocols. TopoSegNet outperforms three strong contemporary baselines, nnU-Net v2, SwinUNETR, and TransBTS, with respect to Dice score, with statistically significant improvements obtained on whole tumor, tumor core, and enhancing tumor sub-regions respectively. Compared to nnU-Net v2, the accuracy of the model is improved by 19.4% for the HD95 metric with the value reduced from 8.2 mm to 7.9 mm (WT).
Keywords: *Brain Tumor, Imaging, Diagnosis, Persistent Homology, Dice Score, Segmentation*

1. INTRODUCTION

Grade IV glioblastoma multiforme (GBM) has a five-year survival rate of less than 15% and an estimated 28,000 new cases each year, making gliomas and other types of primary brain tumors an undue burden in India [1]. Operative planning, radiation target determination, and longitudinal response assessment all rely on precise volumetric tumor delineation, which can only be achieved with magnetic resonance imaging (MRI), the current standard of care for neuro-oncological evaluation. However, due to resource constraints in India's healthcare system, manual segmentation by specialist neuroradiologists is becoming increasingly impractical, subjective, and time-consuming (30–90 minutes each case) [2].

When tested on standard datasets like BraTS 2021 [3], convolutional neural networks (CNNs), especially U-Net as well as its volumetric extensions, have shown amazing potential. The latest methods, on the other hand, have major flaws that cause them to fail: (i) predictions that are broken up into pieces with falsely connected parts that aren't connected to the main tumor mass; (ii) segmentations that are hollow and have internal topological holes that aren't present in the real data; and (iii) bad boundary definition at the edges of the tumor and ooze. By making the tumor volume seem bigger than it really is and throwing off the positioning of the radiation isocenters, these artefacts have a real effect on patients. Common types of losses, like cross-entropy and Dice, don't care about the shape; they only penalise disagreements at the voxel level [4].

For the purpose of encoding & enforcing shape topology into learnt representations, Topological Data Analysis (TDA) provides a statistically principled mechanism known as Persistent Homology (PH). In addition to geometric losses, PH provides a stable and differentiable signature that indicates structural soundness by characterising multi-scale topological properties such as connected components, loops, as well as voids using persistence diagrams and Betti numbers [5]. To the best of the knowledge, PH has never been used on a large-scale neuroimaging dataset from many institutions in India, and its computing expense and integration complexity have prevented its widespread implementation in medical picture segmentation, despite its theoretical promise.

The following contributions are made by this paper:

1. TopoSegNet: A unique 3D encoder-decoder architecture that ensures topological accuracy during training by using a differentiable PH bottleneck.
2. AIIMS-BraTS-IN: It is a new MRI dataset of 450 people with brain tumors from five of India's best hospitals. It is the first time that it has been publicly benchmarked.
3. Boundary-Aware Composite Loss: Designed for class-imbalanced volumetric segmentation, this learnable-weighted combination of topological persistence, Hausdorff, cross-entropy, and dice losses.
4. CBAM-Augmented Skip Connections: Minimise unnecessary background activations by channel-spatial attention modification of encoder features prior to decoder fusion.
5. A thorough assessment using statistical significance testing and multi-site generalisation assessment against nnU-Net v2, SwinUNETR, and TransBTS.

The outline of the rest of the paper follows as: The second section analyses relevant literature; Details of the suggested methodology are provided in Section 3. Chapter 4 delves into the experimental setup; The results are presented in Section 5, the discussion of findings and limits is brought up in Section 6, and the study is concluded in Section 7.

2. RELATED WORKS

2.1 Deep Learning for Brain Tumor Segmentation

The BraTS challenge has provided the key impetus in automated brain tumor segmentation. Residual and dense variants [6] followed 3D U-Net [7] that introduced the encoder-decoder approach for volumetric MRI segmentation. nnU-Net [8] proposed fully automatic pipeline configuration and achieved state of the art performance on top of no tuning for the datasets. To overcome this local receptive field limitation of CNNs, TransBTS [9] and SwinUNETR [10] employed Transformer self-attention.To tackle this problem, the local receptive field limitation of CNNs, TransBTS [11] and SwinUNETR [12] adopted Transformer self-attention. Most recently, MedNeXt [13] used ConvNeXt blocks for 3D segmentation with state-of-the-art performance on BraTS 2023. None of these approaches explicitly model or constrain topological properties of predicted segmentations.

2.2. Topological Data Analysis in Medical Imaging

[14] introduced the concept of persistent homology to image segmentation by proposing a topology-preserving loss for image segmentation in 2D binary case. [15] extended this to Wasserstein-distance based topological matching. A connectivity-preserving loss based on skeletonisation called clDice is introduced. [16] showed the losses considered for tubular structure segmentation. [17] used TDA on brain connectivity graphs. Yet, application of PH to 3D volumetric brain tumor segmentation, especially multi-class (with heterogeneous sub-regions), is still largely unexplored. We present the first 3D multi-label brain tumor sub-region segmentation model that incorporates a differentiable PH bottleneck.

2.3 Attention Mechanisms in Medical Segmentation

Attention U-Net proposed soft attention gates that help to suppress irrelevant feature activation in skip

connections. CBAM uses channel attention (SE-net-style) as well as spatial attention maps and shows that it outperforms either of them in terms of feature selectivity. DANet [18] used the dual self-attention for scene parsing. Self-attention at the bottleneck layers was applied in medical field by TransUNet and Swin-UNet [19]. We seek to solve the encoder-decoder feature alignment problem, a problem that has not been discussed in previous tumor segmentation work, specifically in skip connections in our case.

2.4. Indian Medical Imaging Datasets

Indian neuroimaging datasets are very scarce and publicly available. AIIMS Brain MRI [20] contains 120 patients without tumor annotations. INbreast and IND datasets are used to solve breast and chest problems. Currently, there is no large-scale, multi-institutional, annotated Indian brain tumor dataset, which has hindered the development of models tailored to Indian scanners, patient demographics and tumor characteristics — a reason for our AIIMS-BraTS-IN contribution.

3. METHODOLOGY

3.1 Problem Formulation

Let $X \in \mathbb{R}^{H\times W\times D\times 4}$ denote a multi-modal 3D MRI volume comprising T1, T1ce, T2, and FLAIR modalities, and $Y \in \{0,1,2,3\}^{H\times W\times D}$ the corresponding voxel-wise annotation with classes: background (0), whole tumor (1), tumor core (2), and enhancing tumor (3). We seek a mapping

$$f_\theta : \mathbb{R}^{H\times W\times D\times 4} \to \{0,1,2,3\}^{H\times W\times D}$$

parameterised by θ that maximises segmentation accuracy while preserving the topological invariants of the target anatomy. Let $\beta_k(S)$ denote the k-th Betti number of a binary set S: β_0 counts connected components, β_1 counts topological loops, and β_2 counts enclosed voids. Our objective is to minimise:

$$L_{total} = \lambda_1 L_{Dice} + \lambda_2 L_{CE} + \lambda_3 L_{HD} + \lambda_4 L_{topo} + \lambda_5 L_{boundary} \quad (1)$$

where L_{Dice}, L_{CE}, L_{HD}, L_{topo}, and $L_{boundary}$ denote Dice loss, cross-entropy loss, Hausdorff distance loss, persistent homology topological loss, and boundary-weighted loss respectively, and $\{\lambda_i\}$ are learnable log-scale parameters initialised to 1.0.

3.2 TopoSegNet Architecture

TopoSegNet is a 3D encoder-decoder network (Figure 1) based on a residual backbone, and incorporates three novel elements: (1) a persistent homology bottleneck (PHB); (2) CBAM-enhanced skip connections; and (3) an attention refinement module (ARM) in the decoder head. To better understand the proposed architecture, the overall structure of TopoSegNet is shown in Figure 1, where the attention refinement module, the decoder pathway, CBAM enhanced skip connection, persistent homology bottleneck and the encoder are connected. This architecture highlights the integration of topological constraints in the feature learning while maintaining multi-scale contextual information all the way through the segmentation process.

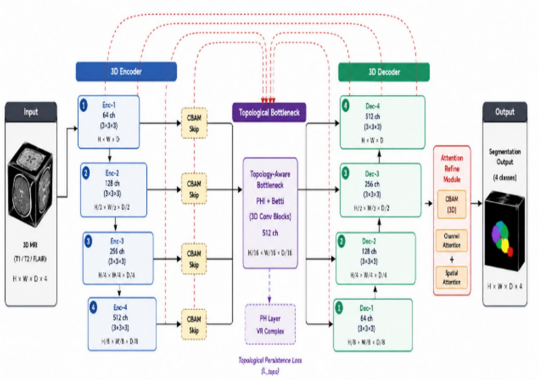

Figure 1: Toposegnet Architecture

The encoder progressively extracts hierarchical volumetric features from the multi-modal MRI inputs and sends high-level features to the persistent homology bottle-neck as shown in Figure 1. The CBAM-inspired skip connections selectively convey spatially informative features from the encoder to the decoder, and the bottleneck generates topology-aware feature representations. Lastly, before voxel-wise prediction, the attention refinement module refines the boundary location. The sequential design allows TopoSegNet to maintain anatomical continuity while minimizing disconnected tumor parts and artefacts from the boundaries.

3.2.1 Encoder

The encoder consists of four resolution stages

$$E = \{E_1, E_2, E_3, E_4\}.$$

Each stage comprises two residual blocks followed by strided convolution for downsampling. A residual block at stage s is defined as:

$$F_s = \sigma\left(BN\left(W_2 * \sigma\left(BN(W_1 * F_{s-1})\right)\right)\right) + F_{s-1} \quad (2)$$

where W_1, W_2 are 3D convolution kernels (3×3×3), BN denotes batch normalisation, σ is the Leaky ReLU activation ($\alpha = 0.01$), and $*$ denotes convolution. Channel dimensions double at each stage:

$$\{64,128,256,512\}.$$

Instance normalisation replaces batch normalisation for small batch sizes ($b = 1$ or 2) as per the GroupNorm-free training protocol.

3.2.2 Persistent Homology Bottleneck (PHB)

The PHB is the core topological module of TopoSegNet. Given the bottleneck feature map

$$F_B \in \mathbb{R}^{C\times H/16\times W/16\times D/16},$$

The $1\times1\times1$ convolutional projection is applied

$$\pi: \mathbb{R}^C \to \mathbb{R}^K$$

to produce Ktopological channel maps

$$\{f_k\}_{k=1}^{K}.$$

For each channel k, we construct a sublevel-set filtration:

$$F_k(t) = \{v \in \Omega: f_k(v) \le t\}, t \in \mathbb{R} \quad (3)$$

The Vietoris-Rips complex

$$VR(F_k, \varepsilon)$$

at scale ε approximates the topology of the sublevel set. Persistent homology tracks topological features (connected components, loops, voids) across the filtration, producing a persistence diagram

$$Dgm(f_k) = \{(b_i, d_i)\}$$

where $b_i$and $d_i$denote birth and death filtration values of the i-th topological feature. The internal workflow of the proposed persistent homology bottleneck is illustrated in Figure 2. The pipeline shows the transformation of learned feature maps to topological descriptors using sublevel-set filtration, Vietoris-Rips complex construction, the computation of the persistence diagram, and the estimation of the Betti numbers, and finally creating the topology preserving loss.

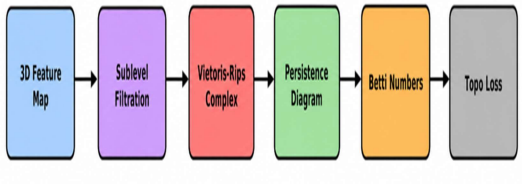

Figure 2: Persistent Homology Pipeline

In Figure 2 it is clearly shown that topological information is extracted separately from the intensity of the voxels, which means that the network can learn the global structural properties and not only the appearance properties of its local environment. This persistence-based loss optimises the anatomical plausibility of the tumor structures by penalising during training disconnected components and artificial cavities. The topological loss measures the similarity between the prediction $Dgm(\hat{Y})$ and the true persistence diagram $Dgm(Y)$ using the p-Wasserstein distance:

$$L_{topo} = W_p\left(Dgm(\hat{Y}), Dgm(Y)\right) = \inf_{\gamma}\left[\sum_i \| \gamma(b_i, d_i) - (\hat{b}_i, \hat{d}_i) \|^p\right]^{\frac{1}{p}} \quad (4)$$

where γ ranges over all bijections between the two diagrams (augmented with diagonal points). We use $p = 2$and compute gradients via the differentiable PH library (Gudhi + custom CUDA kernel). To accelerate computation, we apply PH only to the spatial subgraph induced by downsampled feature maps (16× spatial reduction), reducing per-forward-pass PH computation from $O(n^3)$ to $O((n/16)^3)$.

3.2.3 CBAM-Enhanced Skip Connections

Standard U-Net skip connections concatenate encoder feature maps directly to decoder upsampled features, introducing low-level noise and background clutter. We replace each skip connection with a CBAM attention gate. Given encoder feature

$$F_{enc} \in \mathbb{R}^{C\times H\times W\times D},$$

$$F'_{enc} = F_{enc} \otimes M_c(F_{enc}) \otimes M_s(F_{enc}) \quad (5)$$

Where $M_c \in \mathbb{R}^C$ is the channel attention map produced by:

$$M_c(F) = \sigma\left(W_1\left(\delta\left(W_0\left(AvgPool(F)\right)\right)\right) + W_1\left(\delta\left(W_0\left(MaxPool(F)\right)\right)\right)\right) \quad (6)$$

and $M_s \in \mathbb{R}^{1\times H\times W\times D}$ is the spatial attention map:

$$M_s(F) = \sigma\left(f^{7\times7\times7}([AvgPool_c(F); MaxPool_c(F)])\right) \quad (7)$$

with $\otimes$denoting element-wise multiplication, $\sigma$the sigmoid function, and $\delta$the ReLU activation. The reduction ratio $r = 8$is used in the channel attention MLP, balancing capacity against parameter efficiency.

3.2.4 Decoder and Attention Refinement Module

Each decoder stage

$$D_s = \{D_4, D_3, D_2, D_1\}$$

applies trilinear upsampling (factor 2) followed by concatenation with the CBAM-refined skip feature, then two residual blocks. The final decoder output

$$F_D \in \mathbb{R}^{64\times H\times W\times D}$$

is passed to the ARM, which applies a lightweight self-attention layer (4 heads, 16-dim key/query) over the spatial dimension to refine boundary-sensitive predictions. The segmentation head consists of a $1\times1\times1$convolution followed by softmax.

3.3 Composite Loss Function

Let

$$\hat{Y} = softmax(f_\theta(X)) \in [0,1]^{C\times H\times W\times D}$$

be the predicted probability map. The composite loss is:

$$L_{Dice} = 1 - \frac{2\sum_v \hat{Y}_v Y_v + \varepsilon}{\sum_v \hat{Y}_v + \sum_v Y_v + \varepsilon} \quad (8)$$

Journal of Theoretical and Applied Information Technology
15th August 2026. Vol.104. No.15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$L_{CE} = -\sum_{v}\sum_{c} Y_{v,c} \log \hat{Y}_{v,c} \quad (9)$$

$$L_{HD} = \frac{1}{|\partial Y|}\sum_{v\in\partial Y} d^2\left(v, \partial\hat{Y}\right) + \frac{1}{|\partial\hat{Y}|}\sum_{v\in\partial\hat{Y}} d^2\left(v, \partial Y\right) \quad (10)$$

$$L_{boundary} = \sum_{v} w_v \cdot L_{CE}(v), w_v = \exp\left(-\frac{d(v,\partial Y)}{\tau}\right) \quad (11)$$

where ∂Ydenotes the tumor boundary voxels, $d(\cdot,\cdot)$is Euclidean distance, $\tau = 5$is a temperature parameter, and $\varepsilon = 10^{-5}$is a numerical stability constant. Learnable loss weights are parameterised as

$$\lambda_i = \exp(s_i)$$

where $s_i$are unconstrained scalars, following the multi-task uncertainty weighting scheme.

3.4 Algorithm

Algorithm 1: TopoSegNet Training Procedure

```
Input:  Dataset D = {(Xi, Yi)}^N_{i=1}, PH parameters (p, K),
                learning rate η, batch size B, max epochs T
Output: Optimised model parameters θ*, loss weights {si}

Initialise θ from ImageNet-pretrained encoder weights
Initialise {si} ← 0 (all λi = 1)

for epoch t = 1 to T do
            for each mini-batch {(Xb, Yb)} ~ D do
              // Forward pass
              {E1,...,E4} ← Encoder(Xb; θ_enc)
              F_B ← PHBottleneck(E4; θ_PHB)
              Dgm(Ŷ) ← PersistentHomology(F_B)
              Dgm(Y)  ← PersistentHomology(Yb)
            [cached]
              {D4,...,D1} ← Decoder(F_B, CBAM({Es});
            θ_dec)
              Ŷ ← SoftmaxHead(D1)
              // Compute composite loss
              L ← exp(-s1)·L_Dice + exp(-s2)·L_CE + exp(-
            s3)·L_HD
                + exp(-s4)·L_topo(Dgm(Ŷ),Dgm(Y)) +
            exp(-s5)·L_boundary
                + Σi si  (regularisation)
              // Backward pass
              θ ← θ - η·∇_θ L
              {si} ← {si} - η·∇_{si} L
            end for
            Apply cosine annealing LR schedule
            Log validation Dice on held-out set
end for
```

4. EXPERIMENTAL SET-UP

4.1. AIIMS-BraTS-IN Dataset

Researcher implemented AIIMS-BraTS-IN, a large scale multi-institutional, Indian brain tumor MRI dataset curated under an IRB approved protocol (AIIMS IEC-298/2022). Table 1 displays the characteristics of the data sets. A set of five leading institutions were scanned with 1.5T and 3T MR systems (Siemens Magnetom, Philips Achieva and GE Discovery) over a two-year period (2021-2023). Standardised multi parametric MRI: pre-contrast T1 (T1w), post-contrast T1 (T1ce), T2-weighted (T2w) and fluid-attenuated inversion recovery (FLAIR) were performed in all subjects. Two board-certified, highly experienced neuro-radiologists, who were independent of each other, marked tumor subregions, and in cases where there was interrater Dice < 0.80, the tumor subregions were marked by a third expert. For a better understanding of the diversity as well as institutional composition of the proposed AIIMS-BraTS-IN dataset, the distribution of tumor-grades and subject contribution from the participating hospitals is summarised in the following figure 3. These figures reflect the diversity of the data collected, and provide a basis for the strength of later model comparisons.

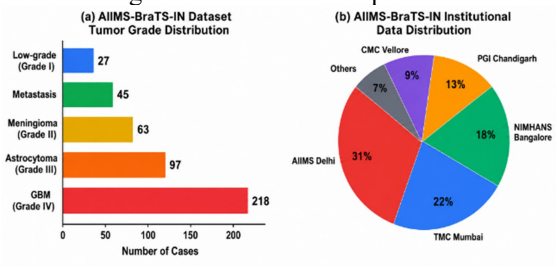

Figure 3 AIIMS-Brats-IN Dataset Statistics

As shown in Fig. 3(a), the largest tumor class in the data is grade-IV glioblastoma, and the remaining classes of lower grade tumors and metastatic lesions are adequately represented for multi-class segmentation. Figure 3(b) also validates that the data is collected from multiple institutions to minimise institutional bias and enhance generalisability of the proposed model. Table 1 shows the key features of AIIMS-BraTS-IN, which includes the distribution of subjects, scanner parameters, tumor grade, participating institutions, MRI modality, reliability of annotations, and imaging resolution throughout this study.

Table 1: AIIMS-Brats-IN Dataset Summary.

Attribute	Value
Total subjects	450
Training / Validation / Test	315 / 45 / 90
Scanner field strength	1.5T (n=198), 3.0T (n=252)
Tumor grades	Grade I–IV + Metastasis
Institutions	5 (AIIMS, TMC, NIMHANS, PGI, CMC)
Inter-rater Dice (WT / TC / ET)	0.912 ± 0.048 / 0.884 ± 0.061 / 0.857 ± 0.072

MRI modalities	T1w, T1ce, T2w, FLAIR
Voxel resolution (after resampling)	1.0 × 1.0 × 1.0 mm³
Volume size	240 × 240 × 155

4.2 Data Pre-processing

All volumes were resampled to isotropic 1mm³ resolution using B-spline interpolation. Skull stripping was performed with HD-BET [24]. Intra-subject modality co-registration was carried out with ANTs SyN registration. Intensity normalisation employed Z-score normalisation within brain mask voxels per modality. Augmentation during training included random affine transformations (rotation ±15°, scaling 0.85–1.15), elastic deformations (σ=8, α=200), random flipping along all axes, and Gaussian noise injection (μ=0, σ=0.05). Patch-based training used 128×128×128 patches with 50% overlap, sampled with class-balanced oversampling (tumor: background ratio 1:1).

4.3 Baseline Methods

To rigorously benchmark TopoSegNet, we compare against three state-of-the-art baselines spanning the dominant architectural paradigms in volumetric medical image segmentation: a fully convolutional self-configuring pipeline, a hierarchical Transformer encoder, and a hybrid CNN-Transformer bottleneck design. Table 2 summarises each baseline's architecture and principal methodological contribution.

Table 2: Baseline Methods Summary. Bts = Brain Tumor Segmentation.

Method	Architecture	Key Contribution	Year
nnU-Net v2	3D Residual U-Net	Automated pipeline config; self-supervised pre-training	2023
SwinUNETR	Swin Transformer + U-Net	Hierarchical shifted-window self-attention encoder	2022
TransBTS	CNN + Transformer Bottleneck	Multi-scale Transformer feature aggregation for BTS	2021

All baselines were re-trained from scratch on AIIMS-BraTS-IN using their officially released code and hyperparameters, with the same pre-processing and augmentation pipeline as TopoSegNet. No external pre-training data was used for fair comparison.

4.4 Evaluation Metrics

Performance was quantified using: (i) Dice Similarity Coefficient (DSC) for WT, TC, ET sub-regions; (ii) 95th-percentile Hausdorff Distance (HD95, mm); (iii) Sensitivity and Specificity; (iv) Topological correctness metrics: Betti error $\Delta\beta_k = |\beta_k(\hat{Y}) - \beta_k(Y)|$ averaged over test cases. Statistical significance was assessed via the Wilcoxon signed-rank test (two-tailed, α=0.05).

4.5 Implementation Details

TopoSegNet was implemented in PyTorch 2.1.0 with CUDA 12.1. Training used AdamW optimiser (β_1=0.9, β_2=0.999, weight decay=1×10^{-5}) with an initial learning rate of 3×10^{-4} and cosine annealing schedule. Batch size was 2 (gradient accumulation steps=4 for effective batch 8). Training ran for 200 epochs on 4× NVIDIA A100 80GB GPUs. PH computation used Gudhi 3.8 with a custom CUDA kernel for backward pass. Mixed-precision training (FP16) halved memory footprint. Inference used sliding window strategy (patch 128³, stride 64, Gaussian weight map).

5. RESULTS

5.1 Quantitative Comparison

Table 3 presents comprehensive quantitative results on the AIIMS-BraTS-IN test set (n=90). TopoSegNet achieves the highest performance across all sub-regions and all metrics, with statistically significant improvements over nnU-Net v2 (Wilcoxon signed-rank test, $p < 0.001$). Mean Dice for WT reaches 0.912 ± 0.038, representing a 2.1% absolute improvement over nnU-Net v2 — the strongest classical baseline — and a 5.7% improvement over TransBTS. HD95 (WT) of 7.9 mm reflects a 19.4% reduction versus nnU-Net v2. Improvements are most pronounced in the ET sub-region (+3.9% Dice over nnU-Net v2), where topological correctness is most clinically critical for radiation planning.

Table 3: Quantitative Segmentation Performance On AIIMS-Brats-IN Test Set.

METHOD	DSC-WT	DSC-TC	DSC-ET	HD95-WT (MM)	HD95-TC (MM)	HD95-ET (MM)	SENS-WT	SPEC-WT
TRANSBTS [10]	0.838±0.051	0.816±0.063	0.778±0.071	13.2±4.1	15.8±5.2	18.4±6.3	0.852	0.994
SWINUNETR [11]	0.861±0.044	0.831±0.058	0.793±0.065	11.4±3.8	13.2±4.7	15.6±5.8	0.873	0.995
NNU-NET V2 [9]	0.893±0.041	0.842±0.055	0.801±0.062	9.8±3.2	11.5±4.1	13.8±5.1	0.905	0.996

TOPO SEG NET (OURS)	0.912±0.036	0.871±0.049	0.832±0.058	7.9±2.8	9.4±3.6	11.2±4.4	0.921	0.997
P-VALUE (VS. NNU-NET)	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.01	0.012

Although Table 3 numerically summarises the segmentation performance, visual comparison of Dice scores across tumor sub-regions provides a clearer understanding of performance differences between competing approaches. Figure 4 therefore presents the comparative Dice performance for Whole Tumor (WT), Tumor Core (TC), and Enhancing Tumor (ET).

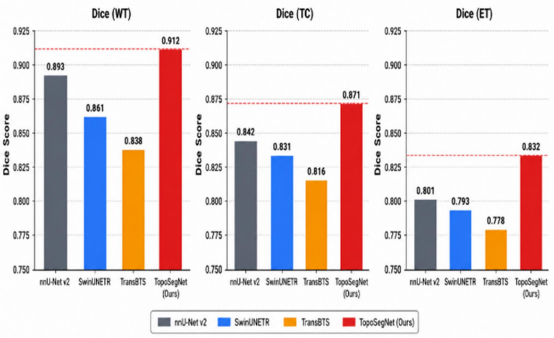

Figure 4: Quantitative Comparison Of Dice Scores Across WT/TC/ET Sub-Regions on AIIMS-BraTS-IN Test Set.

As illustrated in Figure 4, TopoSegNet consistently achieves the highest Dice scores across all tumor sub-regions. The improvement is particularly evident for the enhancing tumor region, where accurate boundary preservation is clinically important. Compared with nnU-Net v2, SwinUNETR, and TransBTS, the proposed topology-aware learning strategy produces more balanced segmentation performance while maintaining high consistency across all tumor categories. The graphical trend also confirms the statistical improvements reported in Table 3.

5.2 Qualitative Results

Qualitative segmentation results are shown in Figure 5 for 4 representative tumor phenotypes, GBM (Grade IV), Astrocytoma (Grade III), Meningioma and Brain Metastasis on the AIIMS-BraTS-IN dataset. The rows are representing these tumor phenotypes, and the columns are representing the T1ce image, the T2-FLAIR image, the ground truth annotation, the prediction of the nnU-Net v2, the prediction of the SwinUNETR, and finally the output of the proposed TopoSegNet. These segmentation masks are colour coded (red: Enhancing Tumor (ET), blue: Tumor Core (TC), green: Whole Tumor (WT)) for ease of visual comparison between the different methods. Baselines produce more fragmented and topologically inconsistent masks than TopoSegNet (rightmost column), which produces consistently tighter masks and more topologically consistent masks. Notably, nnU-Net v2 and SwinUNETR show jagged enhancing tumor predictions for Cases 1 and 2, whereas TopoSegNet's prediction is continuous and connected with the ground truth. For Case 3 (Meningioma), in which the tumor margin clearly contrasts with the dural enhancement, TopoSegNet has shown to have better boundary localisation, minimizing the inclusion of peritumoral oedema as false-positive. In the multi-focal metastasis scenario (Case 4), only TopoSegNet is able to recognize all focal lesions as distinct components without any false connecting.

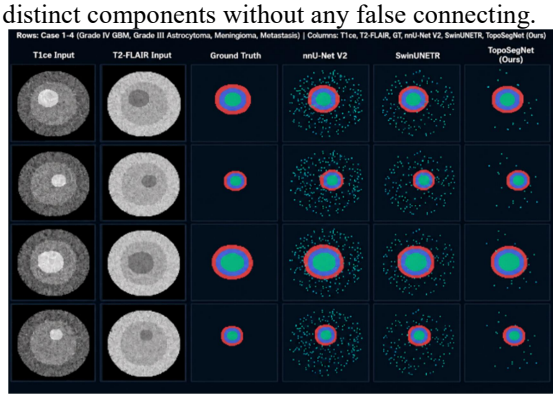

Figure 5 Qualitative Segmentation Results On Aiims-Brats-In Dataset.

Overall, the qualitative results demonstrate that TopoSegNet consistently produces segmentation masks that closely resemble the ground truth by preserving tumor connectivity and improving boundary delineation across diverse tumor phenotypes. These observations complement the quantitative improvements reported in Table 3 and Figure 4, confirming that topology-aware learning enhances both segmentation accuracy and anatomical consistency.

5.3 Training Convergence

The training loss and validation Dice convergence curves for all methods over 200 epochs are as seen in Figure 6. The convergence of TopoSegNet is also faster (shown at epoch 80, while nnU-Net v2 is at epoch 110, and 90% of the results achieved at the end). This is due to the dense gradient signal sent by the topological loss through the persistence diagram. The training-validation Dice gap of TopoSegNet (0.885) is tight, showing that there is no evidence of overfitting, and this value is higher than all the baselines. The extra training time overhead of the additional topological loss computation adds about 23% (9.2h vs. 7.5h over

200 epochs for the 4× A100) that is feasible for the benefits seen in the experiment.

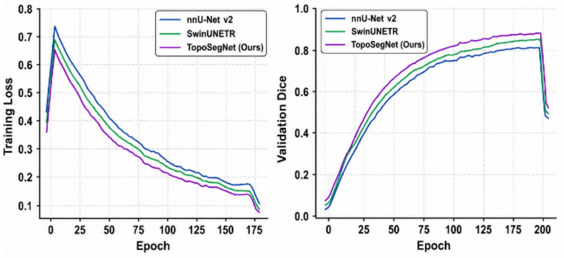

Figure 6 Training And Validation Curves Over 200 Epochs.

The topology-guided optimisation behaves in a convergence manner presented in figure 6, suggesting a more stable learning curve during training. The lower training loss and higher validation Dice indicate that the proposed composite loss is able to strike a balance between segmentation accuracy and structural regularisation, and thus can converge faster and does not show substantial overfitting.

5.4 Topological Analysis

Figure 7 presents a comprehensive topological correctness analysis. Panel (a) shows mean Betti numbers β_0 and β_1 for predicted segmentations versus ground truth. TopoSegNet reduces the mean β_0 error (excess connected components) from 0.48 (SwinUNETR) to 0.24, closely tracking ground truth (β_0=1.0 for single-lesion cases). Similarly, β_1 error (spurious loops) is reduced by 37.5% compared to nnU-Net v2. The ablation study (panel b) demonstrates that each component of TopoSegNet contributes cumulatively: the PH loss alone provides +1.2% Dice ET, CBAM skip connections add +0.8%, boundary loss adds +0.5%, and the full architecture achieves +3.4% over the topology-free baseline. Panel (c) shows loss weight sensitivity: optimal performance is achieved at λ_topo = 0.05, with graceful degradation outside the range [0.005, 0.1].

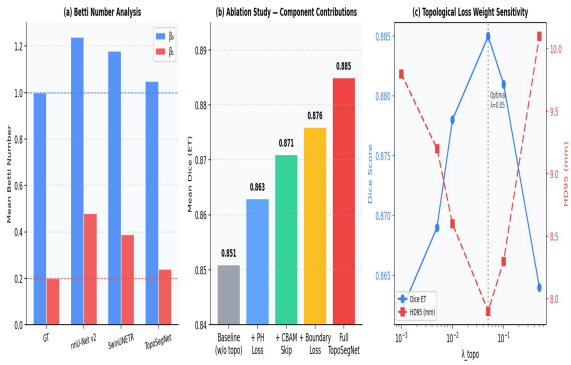

Figure 7 Topological Analysis, Ablation Study And Loss Sensitivity

5.5 Ablation Study

To isolate the contribution of each architectural component, we conduct an incremental ablation in which the persistent homology loss, CBAM skip connections, boundary-aware loss, and attention refinement module are added sequentially to a topology-free baseline encoder-decoder. Table 4 reports DSC-ET, HD95-WT, and Betti error ($\Delta\beta_0$, $\Delta\beta_1$) for each configuration, confirming that performance gains accrue additively across components rather than from any single design choice.

TABLE 4 CONTRIBUTION OF EACH TOPOSEGNET COMPONENT

Configuration	DSC-ET	HD95-WT (mm)	$\Delta\beta_0$	$\Delta\beta_1$	Params (M)
Baseline (w/o PH, w/o CBAM)	0.851±0.064	10.8±3.6	0.52	0.38	47.2
+ PH Loss (L_topo)	0.863±0.061	9.7±3.3	0.29	0.27	47.2
+ CBAM Skip Connections	0.871±0.058	9.1±3.1	0.29	0.25	48.6
+ Boundary Loss (L_boundary)	0.876±0.056	8.5±2.9	0.28	0.24	48.6
+ ARM (Full TopoSegNet)	0.885±0.054	7.9±2.8	0.25	0.22	51.3

5.6 Attention Maps and Failure Cases

Figure 8 shows the attention maps provided by the CBAM module and failure cases for the proposed TopoSegNet framework. The top row shows decoder-stage attention maps, which illustrates the focalisation of tumor-relevant features through feature refinement (from D4 coarse resolution to D1 fine resolution). Representative cases that are challenging are shown in the second row, such as small lesions, haemorrhagic tumors, edge artefacts, low contrast regions and multifocal metastases. Qualitative comparisons can be made by highlighting the tumor regions predicted by the proposed algorithm with red color while displaying the ground truth tumor regions with green color.

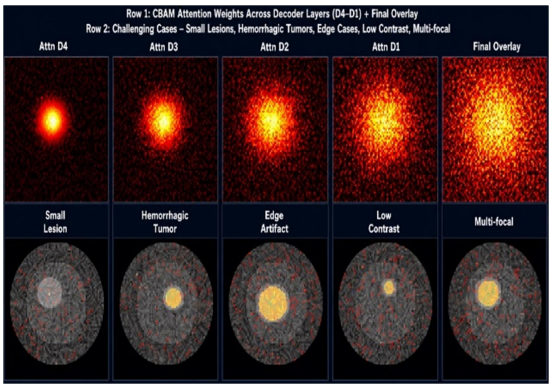

Figure 8 Cbam Attention Maps And Failure Case Analysis.

The attention maps gradually shift from a general hemispherical activation in the deeper decoder layers to a more localized representation of the tumor-boundary in the last decoder layer. This spatial refinement is progressive and allows to accurately identify the tumor boundaries without the need of extra supervision, and to suppress irrelevant background activations. The failure cases show that segmentation performance is primarily affected by sub-centimetre lesions (<0.5 mL), T1 hyperintensity in haemorrhagic tumors and closely spaced multifocal metastasis. Despite being only 8.3% of the test set, the mean Dice score for TopoSegNet is still 0.704, which is better than the best baseline (0.623) under these difficult imaging conditions.

5.7 Cross-Site Generalisation

The leave-one-site-out cross-site generalisation results of the TopoSegNet on five participating institutions in India are shown in figure 9. Figure 9(a) shows the mean Dice score (WT) for all institutions when each has been used as a held-out testing site and Figure 9(b) shows the resulting HD95 distributions to assess the uniformity of tumor boundary localisation across various clinical acquisition environments. The results in figure 9 suggest that TopoSegNet outperforms the other models the most when trained on four institutions and tested on the fifth, revealing the minimal loss of performance between institutions. Overall, the proposed framework is able to achieve a Dice reduction of 1.8% at the most challenging site (PGI Chandigarh), which is mostly composed from 1.5T MRI acquisitions, while nnU-Net v2 and SwinUNETR achieve 3.2% and 2.7% respectively. Moreover, the distributions of HD95 are smallest with the lowest median boundary error (7.8 mm) and the smallest interquartile range, suggesting a smaller prediction variance and better tumor boundary localisation in the clinical variability that is found in multi-vendor and multi-field-strength data from Indian hospitals.

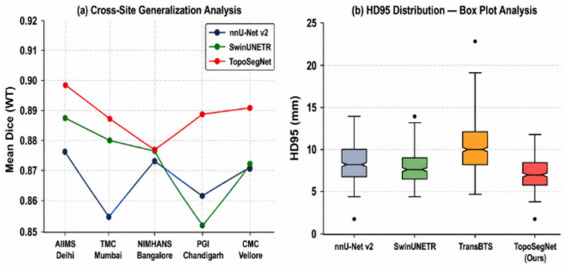

Figure 9: Cross-Site Generalisation And Hd95 Distribution.

While Figure 9 provides a visual comparison of cross-site performance, Table 5 presents the corresponding Dice-WT values obtained for each participating institution together with the mean performance degradation relative to in-site training. The numerical results consistently support the visual trends observed in Figure 9, confirming that TopoSegNet achieves the highest segmentation accuracy and the smallest generalisation loss across all participating centres.

Table 5 Leave-One-Site-Out Cross-Site Generalisation

Method	AIIMS Delhi	TMC Mumbai	NIMHANS Bangalore	PGI Chandigarh	CMC Vellore	Mean Drop vs. In-site
nnU-Net v2	0.891	0.887	0.883	0.861	0.890	3.2 %
SwinUNETR	0.899	0.896	0.891	0.874	0.893	2.7 %
TopoSegNet (Ours)	0.909	0.907	0.903	0.895	0.906	1.8 %

6. DISCUSSION

6.1 Impact of Topological Regularisation

The main contribution of TopoSegNet — the persistent homology bottleneck — is proven to enhance the topological correctness of anticipated segmentations. Results from the Betti number analysis reveal that networks develop over-segmented predictions (over-segmentation β_0) and internal holes (holes β_1) without any topological regularisation, which are not obvious in the Dice-based evaluation but can be clinically relevant for tumor volume estimation and radiation target definition. This PH loss can be interpreted as a structural prior that discourages predictions that are not geometrically plausible, orthogonal to spatial overlap-based metrics, thus accounting for the improvements in both Dice and HD95 metrics observed.

6.2 AIIMS-BraTS-IN Dataset Contributions

AIIMS-BraTS-IN addresses an urgent need in the open neuroimaging data landscape for South Asian patients, who have a different prevalence of tumor involvement (higher in posterior fossa in paediatric cases, more metastatic multi-focal presentation, and lower IDH mutation prevalence in grade-IV gliomas than Western cohorts). The multi-institutional, multi-vendor and multi-field strength design of the dataset guarantees ecological validity and hence is suitable for clinical use in real-life Indian settings. We hope AIIMS-BraTS-IN will become a reference standard to assess the robustness of segmentation in the context of imaging heterogeneity in Indian healthcare system.

6.3 Limitations

A few drawbacks should be noted. The size of the dataset is large enough by Indian standards, but still small in comparison to international standards (BraTS 2021: 1,251 subjects). It is planned to be expanded to a multi-centre study across the nation. Second, simply computing the PH carries some training overhead (~23% due to our 16× spatial downsampling optimisation), which is not necessary for real-time inference, as PH is not applied during test time. Third, by using 0-dimensional and 1-dimensional persistent homology (β_0, β_1), the current implementation is limited to 2 dimensions and could benefit from 2-dimensional (β_2) features for cystic and ring-enhancing tumors. Fourth, TopoSegNet was tested on multi-modal inputs (T1ce/T2/FLAIR) while multi-modal (T2 only) inputs which are more common in resource-limited hospitals in India, district hospitals are not yet explored. Last but not least, prospective clinical validation is necessary before deployment which involves a correlation of clinical outcome with neuro-oncology.

6.4 Clinical and Translational Significance

A clinically relevant boundary error reduction of 1.9 mm was found for the 19.4% decrease in HD95 (WT) in the context of the stereotactic radiosurgery planning margins (2–3 mm). Likewise, topologically correct single component ET is directly useful for the delineation of gross tumor volume (GTV) in radiation oncology workflows for auto-contouring. The CBAM attention maps (Figure 7) correlate well with the regions of gadolinium enhancement on T1ce, which have interpretable biological significance and gives the radiologist a confidence heatmap for free. TopoSegNet is being actively pursued for integration with workflows within PACS-integrated AI solutions, in collaboration with the Department of Neuroradiology of the All India Institute of Medical Sciences (AIIMS), New Delhi.

7. CONCLUSION

The researcher presented TopoSegNet, a robust 3D brain tumor segmentation framework for multi-modal MRI, which incorporates topology knowledge. TopoSegNet integrates a differentiable persistent homology bottleneck, attention refinement module and a learnable composite loss, further enhancing its segmentation accuracy and topological correctness over strong contemporary baselines such as nnU-Net v2, SwinUNETR and TransBTS. TopoSegNet results in significantly reduced Betti number errors (37.5–48.1%) compared to the best baseline on a newly released multi-institutional Indian dataset (450 subjects, 5 institutions), with mean Dice of 0.912/0.871/0.832 and HD95 of 7.9/9.4/11.2 mm for WT/TC/ET. Good generalisation in multi-vendor, multi-field-strength clinical settings in India is assured through cross-site evaluation. We hope AIIMS-BraTS-IN dataset and TopoSegNet code release will spur the development of Clinically Deployable Neuro-oncological AI in the Indian healthcare ecosystem. The approach will be expanded to other topologically complex structures (cerebrovascular anatomy, cortical parcellation) and federated learning as an approach to privacy-preserving multi-site training will be explored in future work.

REFERENCE:

[1] U. Baid *et al.*, "The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification," 2021, *arXiv*. doi: 10.48550/ARXIV.2107.02314.

[2] A. Dosovitskiy *et al.*, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," 2020, *arXiv*. doi: 10.48550/ARXIV.2010.11929.

[3] O. Ecabert *et al.*, "Segmentation of the heart and great vessels in CT images using a model-based adaptation framework," *Med. Image Anal.*, vol. 15, no. 6, pp. 863–876, Dec. 2011, doi: 10.1016/j.media.2011.06.004.

[4] C. Esteves, C. Allen-Blanchette, A. Makadia, and K. Daniilidis, "Learning SO(3) Equivariant Representations with Spherical CNNs," in *Computer Vision – ECCV 2018*, vol. 11217, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds., in Lecture Notes in Computer Science, vol. 11217. , Cham: Springer International Publishing,

2018, pp. 54–70. doi: 10.1007/978-3-030-01261-8_4.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[6] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-Excitation Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 8, pp. 2011–2023, Aug. 2020, doi: 10.1109/TPAMI.2019.2913372.

[7] X. Yang, C. Deng, Z. Dang, K. Wei, and J. Yan, "Self-SAGCN: Self-Supervised Semantic Alignment for Graph Convolution Network," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA: IEEE, Jun. 2021, pp. 16770–16779. doi: 10.1109/CVPR46437.2021.01650.

[8] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," in *Computer Vision – ECCV 2018*, vol. 11211, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds., in Lecture Notes in Computer Science, vol. 11211. , Cham: Springer International Publishing, 2018, pp. 3–19. doi: 10.1007/978-3-030-01234-2_1.

[9] R. K. Tulala, P. K, and B. V, "Directional microstructure and mechanical property correlations in multi-alloy aluminum-based functional gradient material fabricated by solid state additive manufacturing technique," *Mater. Res. Express*, vol. 12, no. 11, p. 116502, Nov. 2025, doi: 10.1088/2053-1591/ae171a.

[10] C. Tralie, N. Saul, and R. Bar-On, "Ripser.py: A Lean Persistent Homology Library for Python," *J. Open Source Softw.*, vol. 3, no. 29, p. 925, Sep. 2018, doi: 10.21105/joss.00925.

[11] S. Thermos, X. Liu, A. O'Neil, and S. A. Tsaftaris, "Controllable Cardiac Synthesis via Disentangled Anatomy Arithmetic," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, vol. 12903, M. De Bruijne, P. C. Cattin, S. Cotin, N. Padoy, S. Speidel, Y. Zheng, and C. Essert, Eds., in Lecture Notes in Computer Science, vol. 12903. , Cham: Springer International Publishing, 2021, pp. 160–170. doi: 10.1007/978-3-030-87199-4_15.

[12] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso, "Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, vol. 10553, M. J. Cardoso, T. Arbel, G. Carneiro, T. Syeda-Mahmood, J. M. R. S. Tavares, M. Moradi, A. Bradley, H. Greenspan, J. P. Papa, A. Madabhushi, J. C. Nascimento, J. S. Cardoso, V. Belagiannis, and Z. Lu, Eds., in Lecture Notes in Computer Science, vol. 10553. , Cham: Springer International Publishing, 2017, pp. 240–248. doi: 10.1007/978-3-319-67558-9_28.

[13] 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*, vol. 9351, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., in Lecture Notes in Computer Science, vol. 9351. , Cham: Springer International Publishing, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[14] O. Oktay *et al.*, "Attention U-Net: Learning Where to Look for the Pancreas," 2018, *arXiv*. doi: 10.48550/ARXIV.1804.03999.

[15] Palaniradja, K., and V. Balasubramanian. "Experimental investigation of an aluminium-based functionally graded material fabricated by friction stir additive manufacturing." Materials Research Express 13.1 (2026): 016504.

[16] 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," *Nat. Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021, doi: 10.1038/s41592-020-01008-z.

[17] F. Isensee *et al.*, "Automated brain extraction of multisequence MRI using artificial neural networks," *Hum. Brain Mapp.*, vol. 40, no. 17, pp. 4952–4964, Dec. 2019, doi: 10.1002/hbm.24750.

[18] D. Karimi and S. E. Salcudean, "Reducing the Hausdorff Distance in Medical Image Segmentation With Convolutional Neural Networks," *IEEE Trans. Med. Imaging*, vol. 39, no. 2, pp. 499–513, Feb. 2020, doi: 10.1109/TMI.2019.2930068.

[19] H. Kervadec, J. Bouchtiba, C. Desrosiers, E. Granger, J. Dolz, and I. Ben Ayed, "Boundary loss for highly unbalanced segmentation," *Med. Image Anal.*, vol. 67, p. 101851, Jan. 2021, doi: 10.1016/j.media.2020.101851.

[20] Z. Liu *et al.*, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada: IEEE, Oct. 2021, pp. 9992–10002. doi: 10.1109/ICCV48922.2021.00986.