

HANS-Net: Hyperbolic Convolution and Adaptive Temporal Attention for Accurate and Generalizable Liver and Tumor Segmentation in CT Imaging

Arefin Ittesafun Abian^{ID}, Ripon Kumar Debnath^{ID}, Md. Abdur Rahman^{ID}, Mohaimenul Azam Khan Raiaan^{ID}, Md Rafiqul Islam^{ID}, Asif Karim^{ID}, Reem E. Mohamed^{ID}, Sami Azam^{ID}

arXiv:2507.11325v2 [eess.IV] 16 Oct 2025

Abstract—Accurate liver and tumor segmentation on abdominal CT images is critical for reliable diagnosis and treatment planning, but remains challenging due to complex anatomical structures, variability in tumor appearance, and limited annotated data. To address these issues, we introduce Hyperbolic-convolutions Adaptive-temporal-attention with Neural-representation and Synaptic-plasticity Network (HANS-Net), a novel segmentation framework that synergistically combines hyperbolic convolutions for hierarchical geometric representation, a wavelet-inspired decomposition module for multi-scale texture learning, a biologically motivated synaptic plasticity mechanism for adaptive feature enhancement, and an implicit neural representation branch to model fine-grained and continuous anatomical boundaries. Additionally, we incorporate uncertainty-aware Monte Carlo dropout to quantify prediction confidence and lightweight temporal attention to improve inter-slice consistency without sacrificing efficiency. Extensive evaluations of the LiTS dataset demonstrate that HANS-Net achieves a mean Dice score of 93.26%, an IoU of 88.09%, an average symmetric surface distance (ASSD) of 0.72 mm, and a volume overlap error (VOE) of 11.91%. Furthermore, cross-dataset validation on the AMOS 2022 dataset obtains an average Dice of 85.09%, IoU of 76.66%, ASSD of 19.49 mm, and VOE of 23.34%, indicating strong generalization across different datasets. These results confirm the effectiveness and robustness of HANS-Net in providing anatomically consistent, accurate, and confident liver and tumor segmentation.

Index Terms—Liver segmentation, tumor segmentation, hyperbolic convolutions, neural representation, synaptic plasticity

I. INTRODUCTION

Manuscript Submitted 16th July, 2025
(Corresponding author: Sami Azam)

This work did not involve human subjects or animals in its research.

Arefin Ittesafun Abian and Ripon Kumar Debnath contributed equally and are joint-first authors. They are affiliated with the Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh, along with Md. Abdur Rahman Bangladesh (e-mail: aabian191042@bscse.uiu.ac.bd, rdebnath192071@bscse.uiu.ac.bd, mrahman202260@bscse.uiu.ac.bd).

Mohaimenul Azam Khan Raiaan is with the Faculty of Science and Technology, Charles Darwin University, Casuarina, NT 0909, Australia, and also with the Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh (e-mail: mohaimenulazamkhan.raiaan@cdu.edu.au).

Md Rafiqul Islam, Asif Karim and Sami Azam are with the Faculty of Science and Technology, Charles Darwin University, Casuarina, 0909, NT, Australia (e-mail: mdrafiqul.islam@cdu.edu.au, asif.karim@cdu.edu.au, Sami.Azam@cdu.edu.au).

Reem E. Mohamed is with the Faculty of Science and Information Technology, Charles Darwin University, Sydney, NSW, Australia (e-mail: reem.sherif@cdu.edu.au).

THE liver is a vital glandular organ that performs functions of digestion, metabolism, detoxification, and immunity [1]. Liver cancer ranks among the five most frequently diagnosed cancers and constitutes the fourth leading cause of cancer-related mortality worldwide, with hepatocellular carcinoma (HCC) being the most prevalent form [2]. Precise segmentation of liver tumors is essential in the diagnosis and radiation therapy of HCC. According to the World Health Organization (WHO), cancer accounted for 8.8 million deaths in 2015, with 788,000 specifically attributed to liver cancer [3], [4]. With increased public health awareness, the need for effective liver cancer diagnosis and treatment has intensified [5], [6]. Historically, specialists have manually identified liver tumors, an accurate, yet tedious and labor-intensive process prone to errors under substantial workload [5], [7]. The development of computer-aided diagnostic systems has introduced intelligent, automated, or partially automated approaches to the analysis of medical images, significantly improving precision, effectiveness, and consistency in the detection and segmentation of tumors. [7]–[9].

Various image processing and machine learning techniques have been developed to improve liver tumor segmentation and support early cancer intervention [4], [6], [10]. Among these methodologies, attention-based and multi-branch models have earned significant interest, integrating sophisticated architectures such as S2DA-Net [11], EG-UNETR [12], Generalized U-Net [13], and ConvNeXt-2U [14]. These models are structured to encapsulate essential semantic features in various imaging phases, increasing both the precision and the robustness of feature extraction [13]. Currently, other innovative strategies have taken advantage of multiphase fusion and multiscale representation techniques, such as PGC-Net [15] and camouflaged feature mining modules [8], which collectively advance the automated identification of complex anatomical structures, encompassing both liver and tumor regions [5].

Recent developments in liver tumor segmentation have made progress, yet several challenges persist in comprehensively addressing the complexity of the task. Current liver and tumor segmentation studies frequently encounter limitations in the effective management of complex and heterogeneous tumor-liver anatomies, resulting in only moderate accuracy levels. These models depend on flattened representations, which hinders their

ability to accurately capture the complex spatial organization and progression of tumors across 3D slices [15]. Furthermore, they are often incapable of segmenting tumors and livers simultaneously [16]. Most methodologies overlook the enhancement of inter-slice consistency without compromising efficiency. They also do not improve adaptive features and construct continuous and fine-grained anatomical boundaries [8], [13]. Moreover, they typically lack the strategies of temporal attention necessary to capture sequential relationships in volumetric scans and face challenges in resolving fine-grained tumor boundaries due to insufficient detail encoding [5]. Finally, the absence of uncertainty estimation prevents the scope of quantifying the confidence in prediction [17].

The reliance on flattened representations in existing liver tumor segmentation models limits their ability to model complex anatomies, maintain inter-slice consistency, and construct fine-grained anatomical boundaries. This motivated us to develop HANS-Net (Hyperbolic-convolutions Adaptive-temporal-attention with Neural-representation and Synaptic-plasticity Network), a novel segmentation framework designed to jointly segment both tumors and liver tissues with improved anatomical coherence and confidence estimation. While individual components of our architecture have seen use in various domains, to the best of our knowledge, no prior work has jointly applied them to medical image segmentation. For instance, hyperbolic learning methods have been explored for representation learning and graph neural networks [18], but their application to dense 2D or 3D segmentation remains largely unexplored. Similarly, biologically inspired mechanisms such as synaptic plasticity are rarely integrated into medical imaging models; existing segmentation architectures (e.g., U-Net, TransUNet, EG-UNETR) do not employ plasticity-based adaptive modulation. Furthermore, while recent efforts have investigated implicit neural representations [19], temporal attention [20], and uncertainty modeling independently, none have unified them in a single end-to-end system.

HANS-Net addresses these shortcomings through a comprehensive design. It employs hyperbolic convolutions to represent anatomical hierarchies in a non-Euclidean space, enabling accurate modeling of tumor complexity, and a wavelet-inspired decomposition module captures multiscale information. While 3D volumetric approaches naturally preserve spatial continuity, they are computationally intensive and often require extensive annotated data. In contrast, 2D slice-based methods offer computational efficiency and can effectively capture inter-slice relationships when equipped with appropriate mechanisms. Accordingly, HANS-Net integrates adaptive temporal attention (ATA) along with multi-view fusion, making the approach particularly suitable for resource-constrained clinical settings while maintaining competitive segmentation performance. A synaptic plasticity module enhances feature adaptivity, and an implicit neural representation branch improves boundary smoothness. Finally, Monte Carlo dropout is used to quantify uncertainty in the predictions. HANS-Net achieves superior performance (Dice: 93.26%), surpassing state-of-the-art (SOTA) accuracy and demonstrating strong cross-dataset generalization (Dice: 85.09% on AMOS 2022), highlight-

ing its robustness and potential for real-world clinical deployment.

The main contributions of this work are as follows.

- Introduces HANS-Net, the first liver tumor segmentation framework that unites hyperbolic geometry and implicit neural representations for continuous, hierarchy-aware learning in curved space.
- Proposes a learnable wavelet-inspired decomposition module that captures both global liver contours and fine tumor textures, enabling efficient multiscale analysis without heavy computation.
- Integrates a bio-inspired synaptic plasticity mechanism that adaptively adjusts feature weights based on activation dynamics, promoting tumor-focused learning without extra supervision.
- Employs a coordinate-based implicit MLP branch with positional encoding to refine segmentation outputs, generating anatomically plausible tumor boundaries at continuous resolution.
- Enhances clinical trust via pixel-wise uncertainty estimation using Monte Carlo dropout and improves inter-slice consistency with lightweight temporal attention mimicking 3D context.

The remainder of this paper is organized as follows. Section II reviews the related work conducted in this area. Section III details our proposed HANS-Net model and its component designs. Section IV describes the datasets, preprocessing, and experimental setup. Section V presents quantitative and qualitative results, including cross-dataset experiments, uncertainty estimation, and ablation studies. Section VI discusses the implications of our findings and their relevance to current research, including potential future directions. Finally, Section VII summarizes our findings and concludes the study.

II. RELATED WORKS

In this section, we present a review of recent studies on liver tumor segmentation and highlight the use of multi-branch models, attention mechanisms, and multiscale feature extraction approaches, as these have been the recent focus for handling tumor complexity and heterogeneity.

A. Multi-branch and attention-based models

Automatic segmentation of liver tumors contributes significantly to supporting clinicians in making diagnosis and treatment decisions [3]. Although Convolutional neural networks (CNN) have been used effectively in medical image segmentation, they struggle to capture long-range pixel dependencies [11]. In contrast, transformer-based models require a substantial number of parameters and involve significant computational expenses [11]. Multibranch and attention-based models have been suggested as a promising alternative to overcome the drawbacks presented by both CNNs and Transformers, as highlighted in recent studies [2], [5], [11]–[13], [16], [17], [21]–[25].

For example, Liu *et al.* [11] introduced the S2DA-Net, which incorporates a double-branched encoder with a group of multi-head cross-attention aggregation modules to improve accuracy

for liver tumors. The model achieved 69.4% Dice per case and 80.0% global Dice on LiTS2017 and 3D-IRCADb, respectively. In another work, Cheng et al. [12] proposed an edge-guided segmentation network, EG-UNETR, by synthesizing TransBridge (TB) and a feature semantic enhancement module. The TB utilized multi-head position-aware attention and fusion attention to facilitate the integration of cross-level and multiscale contextual feature fusion. The model scored Dice scores of 84.45% and 85.01% on the LiTS2017 and 3D-IRCADb, respectively. Meta-heuristic algorithm and multiscale hybrid attention models are also popular in other segmentation tasks [21]. For instance, Wan et al. [22] proposed MHAED-Net that builds an encoder-decoder network with the convolution normalization activation block for efficient segmentation. Similarly, Kuang et al. [17] developed an innovative multiphase channel stacked dual attention module called MCDA, which captures essential semantic information across different phases. The model achieved a Dice score of 79.22%. On the other hand, Wang et al. [5] proposed SBCNet, which incorporates a contextual encoding module to identify tumor variability through the use of a multiscale adaptive kernel, and a boundary enhancement to improve boundary perception by integrating contour learning with the Sobel operator; SBCNet scored a Dice of 82.5%. In another study, Deepak et al. [13] proposed G-Unet, which is capable of incorporating other models such as CNN, residual networks (ResNets), and densely connected convolutional neural networks (DenseNet) into the general U-Net framework. G-Unet reached a global Dice of 72.9%. In another paper, Zhang et al. [23] introduced the Decoupled Pyramid Correlation Network (DPC-Net), which uses a Pyramid Feature Encoder to extract multi-level features and decouples them along spatial and semantic dimensions. The network employed Spatial and Semantic Correlation modules to capture feature correlations, achieving a Dice score of 76.4% on the LiTS dataset. Finally, Hu et al. [16] introduced MSML-AttUNet, a hierarchical attention network that utilizes multilayer dilated convolutions for multiscale feature extraction and expands the receptive field for richer contextual representation; it achieved a Dice score of 87.74%.

Oktay et al. [24] introduced Attention U-Net, a spatial attention mechanism to focus on salient features to improve boundary delineation. However, it lacks dynamic modulation based on feature activity, a gap we address using our biologically inspired Synaptic Plasticity Module. TransUNet [2] and Swin UNETR [25] utilized transformers to model global context, improving long-range dependency capture. In contrast, Chen et al. [14] proposed ConvNeXt-2U, which integrated two U-Nets with interconnected skip connections and 3D ConvNeXt blocks for segmentation. Nevertheless, these architectures in Euclidean space fail to capture the hierarchical spatial relationships of anatomical structures; this is precisely where our Hyperbolic Convolution Layer excels by embedding features in a non-Euclidean geometric space.

B. Multi-Phase Fusion

Recent studies [12], [26], [27] highlight the effectiveness of models that incorporate temporal information and hierarchical representations in liver tumor segmentation. To begin with, Hu et

al. [26] proposed the TMPLiTS method, a trustworthy multiphase segmentation framework that fuses information from multiple temporal CT scans using a multi-expert mixture scheme, achieving a Dice score of 81.07%. Similarly, architectures such as EG-UNETR [12] and 3D U²-Net [27] utilize multiple encoding branches and nested skip connections to integrate multiscale features, allowing both local and global context modeling. Although these models are effective in aggregating spatial information, they often require extensive parameter tuning and remain limited in handling temporal coherence across 3D volumes. In contrast, HANS-Net incorporates a wavelet-inspired decomposition module to facilitate efficient and principled multiscale feature extraction. Beyond spatial scaling, it introduces an Adaptive Temporal Attention mechanism to explicitly model inter-slice dependencies in volumetric scans, an aspect frequently overlooked in slice-based or purely spatial approaches. This dual focus on multiscale structure and temporal dynamics enables HANS-Net to deliver superior segmentation accuracy while preserving anatomical consistency across frames.

C. Geometry and Multiscale Feature Models

Multiscale and geometry-aware architectures are increasingly explored in recent studies [8], [15], [28], [29] for their capacity to model complex anatomical variations in liver and tumor segmentation. Yang et al. [8] introduced MCFMFNet, a multiscale feature aggregation framework that utilizes multiple convolutional kernels with diverse receptive fields, combined through an element-wise fusion strategy. They also proposed a camouflaged feature mining module to focus attention on tumor regions that are often misclassified as background, achieving Dice scores of 0.8413 and 0.7365 across datasets. Similarly, You et al. [15] proposed PGC-Net, which uses a Pyramid Vision Transformer to extract multiscale regional and contour features. These features are refined via a contour-induced parallel graph reasoning mechanism that propagates complementary information through graph convolution and maps it back to the original pixel space. This method achieved Dice scores of 73.63% and 74.16% on the LiTS17 and 3DIRCADb datasets, respectively. Beyond these transformer-based and multiscale CNN approaches, geometry-aware models such as MMF-Net [28] and ResUNet++ [29] employ pyramid pooling, atrous spatial pyramid modules, and residual refinements to improve contextual understanding across scales. While these designs enhance segmentation performance, they are inherently limited to Euclidean space representations and lack mechanisms to model hierarchical anatomical relationships. Their reliance on stacked convolutional layers can also constrain the ability to capture long-range geometric dependencies, especially in cases of irregular tumor-liver boundaries. In contrast, HANS-Net introduces hyperbolic convolutions into the medical image segmentation domain for the first time, addressing these limitations by enabling compact and hierarchy-aware embeddings. This approach captures the latent anatomical structure more effectively, contributing to improved boundary continuity and segmentation robustness, especially important in anatomically complex regions. Combined with HANS-Net's wavelet-inspired

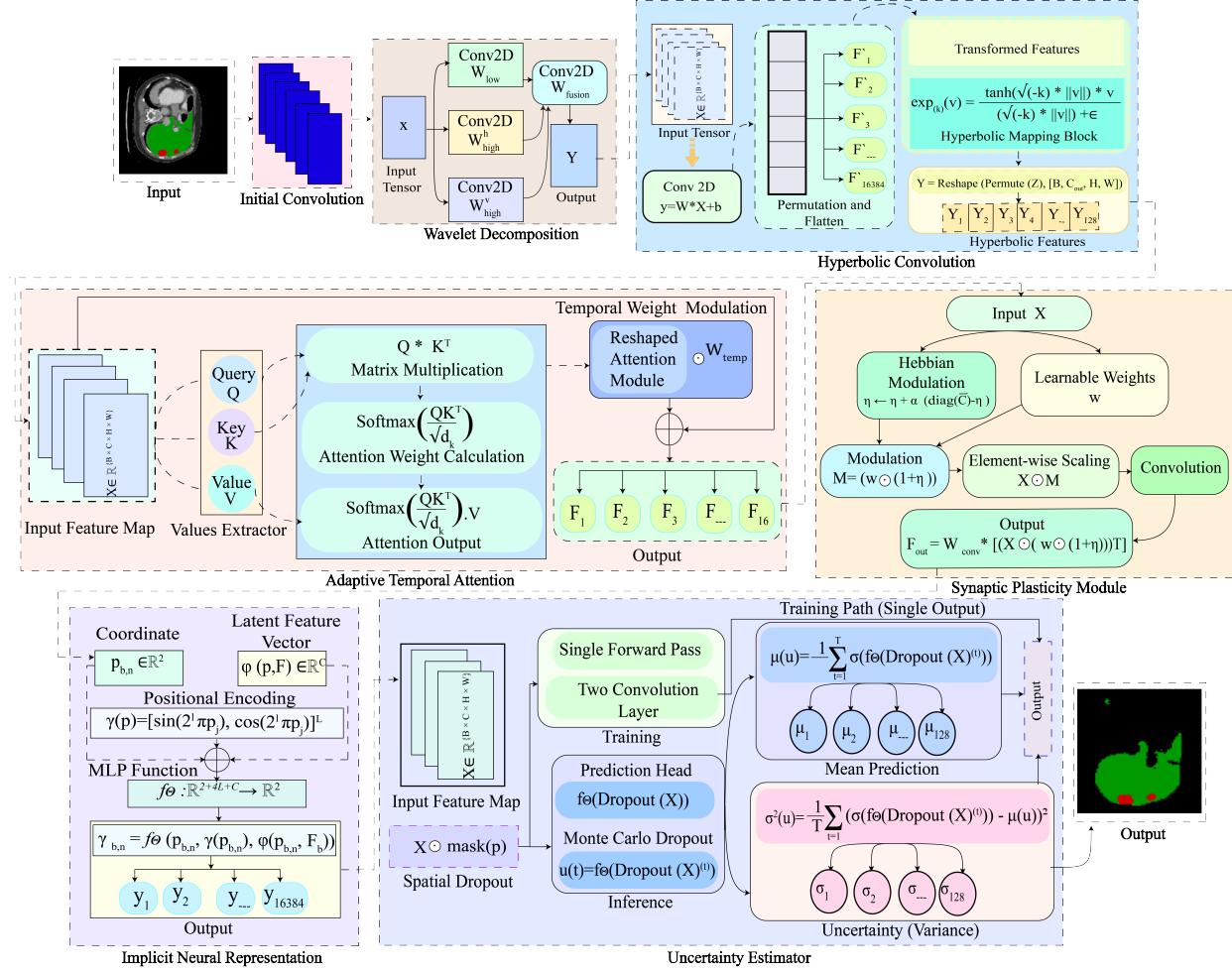


Fig. 1. Overview of the proposed HANS-Net architecture that combines hyperbolic convolutions, wavelet decomposition, temporal attention, synaptic plasticity, implicit neural representation, and uncertainty estimator.

decomposition and temporal modeling strategies, this geometry-aware design offers a significant advancement over traditional multiscale CNNs and transformer-based methods.

Our proposed approach mitigates several shortcomings found in existing studies by demonstrating novel insights and solutions as summarized in Table III. Unlike traditional models, our proposed method, HANS-Net, integrates hyperbolic geometry to better capture the hierarchical structure of liver and tumor patterns, which is not addressed by existing methods. Additionally, the wavelet-inspired decomposition module provides multiscale feature extraction without the high computational cost. Furthermore, the inclusion of a biologically inspired synaptic plasticity mechanism allows HANS-Net to adaptively focus on relevant tumor regions. Moreover, the implicit neural representation branch precisely generates boundaries around anatomical regions. Lastly, the introduction of an uncertainty-aware Monte Carlo dropout enables the model to quantify the confidence of prediction. These advances significantly improve the precision, efficiency, and clinical applicability of liver tumor segmentation, effectively addressing critical gaps present in contemporary methodologies.

III. METHODOLOGY

The overall architecture of HANS-Net is illustrated in Figure 1, which outlines the integrated pipeline consisting of six key modules: Wavelet Decomposition, Hyperbolic Convolution, Adaptive Temporal Attention, Synaptic Plasticity Module, Implicit Neural Representation, and Uncertainty Estimation.

To begin with, the Wavelet-Inspired Decomposition Module splits features into multi-frequency components to preserve both coarse and fine details. Building upon this, the Hyperbolic Convolution Module encodes spatial features into a hyperbolic space to capture hierarchical anatomical relationships. In addition, the Adaptive Temporal Attention module dynamically fuses sequential slices for inter-slice consistency in volumetric data. Furthermore, the Synaptic Plasticity-Enhanced Feature Modulation adaptively strengthens relevant pathways using biologically inspired weight updates. In parallel, the Implicit Neural Representation module enables continuous, high-resolution prediction by modeling outputs as a function of spatial coordinates. Finally, the Uncertainty Estimator, implemented via Monte Carlo Dropout, quantifies prediction confidence to support reliable clinical inter-

pretation. Together, these modules enable HANS-Net to segment liver and tumor structures simultaneously with high accuracy, anatomical boundary fidelity, and robust generalization across datasets.

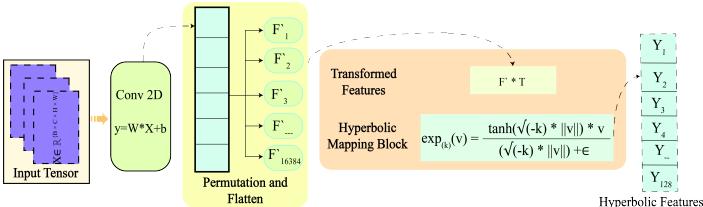


Fig. 2. The Hyperbolic Convolution Module first applies convolution, then flattens and transforms the input before passing it through a hyperbolic mapping block to obtain hyperbolic features.

A. Base Model

Our model processes the input sequentially through a series of specialized modules. It commences with an initial convolution and proceeds with a Wavelet Decomposition Module to extract multi-frequency features. These features are subsequently processed through three hierarchical hyperbolic convolution blocks with pooling to encapsulate structural relationships. The resultant output undergoes refinement via an Adaptive Temporal Attention module and is modulated by a Synaptic Plasticity Module to facilitate enhancement of adaptive features. Concurrently, a parallel Implicit Neural Representation branch was incorporated to capture fine-grained and continuous anatomical boundaries, while an uncertainty estimate quantifies the degree of prediction confidence.

B. Wavelet Decomposition Module

We designed the Wavelet Decomposition Module using convolutional layers to perform low-pass and high-pass filtering, followed by a 1×1 convolution to fuse the resulting feature components into a combined representation [30]. The overall process can be mathematically formulated in two distinct steps. First, the input tensor $\mathbf{x} \in \mathbb{R}^{B \times C \times H \times W}$, where B is the batch size, C is the number of input channels, and H and W are the spatial dimensions, is passed through three convolutional filters to extract the low-pass and high-pass features as shown in Equation (1):

$$\mathbf{F} = \text{Concat}\left(\text{Conv2D}(\mathbf{x}, \mathbf{W}_{\text{low}}), \text{Conv2D}(\mathbf{x}, \mathbf{W}_{\text{high}}^h), \text{Conv2D}(\mathbf{x}, \mathbf{W}_{\text{high}}^v)\right) \quad (1)$$

Next, the concatenated frequency components $\mathbf{F} \in \mathbb{R}^{B \times C_{\text{in}} \times H \times W}$ are fused using a 1×1 convolutional operation, defined in Equation (2):

$$\mathbf{Y}_{b,c,i,j} = \sum_{k=1}^{C_{\text{in}}} \sum_{m=1}^1 \sum_{n=1}^1 \mathbf{F}_{b,k,i+m-1,j+n-1} \cdot \mathbf{W}_{\text{fusion},m,n}^{(c,k)} + \mathbf{b}_c \quad (2)$$

where $\mathbf{Y} \in \mathbb{R}^{B \times C_{\text{out}} \times H \times W}$ is the output tensor, $\mathbf{W}_{\text{fusion}} \in \mathbb{R}^{C_{\text{out}} \times C_{\text{in}} \times 1 \times 1}$ is the adaptive convolutional kernel, and $\mathbf{b} \in \mathbb{R}^{C_{\text{out}}}$ is an optional bias. Here, $b \in \{1, \dots, B\}$ indexes the batch, $c \in \{1, \dots, C_{\text{out}}\}$ the output channels, $i, j \in \{1, \dots, H, W\}$ the spatial positions, $k \in \{1, \dots, C_{\text{in}}\}$ the input channels, and $m, n \in \{1\}$ due to the 1×1 kernel size. This operation performs a weighted channel-wise summation at each spatial location, resulting in the fused output \mathbf{Y} .

The input tensor \mathbf{F} is formed by concatenating features extracted using frequency-specific filters: $\mathbf{W}_{\text{low}} \in \mathbb{R}^{C_{\text{low}} \times K \times K}$ for low-frequency components, $\mathbf{W}_{\text{high}}^h \in \mathbb{R}^{C_{\text{high}} \times K \times K}$ for high-frequency horizontal details and $\mathbf{W}_{\text{high}}^v \in \mathbb{R}^{C_{\text{high}} \times K \times K}$ for high-frequency vertical features. These spectral components are concatenated to form a rich representation in \mathbf{F} , which is then fused by $\mathbf{W}_{\text{fusion}}$. This fusion enables the model to integrate multi-frequency information and enhances representational expressiveness. The resulting tensor \mathbf{Y} serves as the input to the subsequent Hyperbolic Convolution module.

C. Hyperbolic Convolution

We constructed a HyperbolicConv2d layer to integrate standard convolution with hyperbolic geometry for hierarchical feature learning [31]. Given an input tensor $\mathbf{x} \in \mathbb{R}^{B \times C_{\text{in}} \times H \times W}$, where B is the batch size, C_{in} is the number of input channels and H and W are the spatial dimensions, a 2D convolution is first applied with an adaptable kernel $\mathbf{W} \in \mathbb{R}^{C_{\text{out}} \times C_{\text{in}} \times K \times K}$ to produce the feature map \mathbf{F} . Then we flatten the feature map \mathbf{F} along the spatial dimensions and permute to form \mathbf{F}' for transformation in hyperbolic space. Next, the permuted feature map \mathbf{F}' is linearly

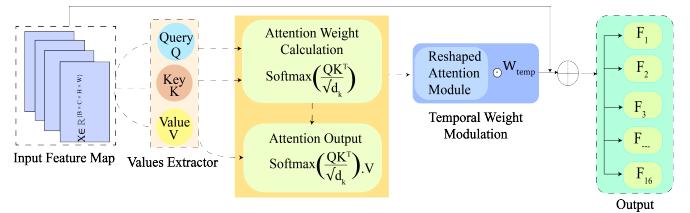


Fig. 3. Work flow Temporal Attention Mechanism that processes input features through query-key-value extraction, attention weight calculation, and temporal weight modulation to generate output features

transformed by a able matrix $\mathbf{T} \in \mathbb{R}^{C_{\text{out}} \times C_{\text{out}}}$, and the hyperbolic exponential map $\exp_{\kappa}(\cdot)$ is applied to each feature vector to embed it in the hyperbolic space, as defined in Equation (3):

$$\mathbf{Z} = \exp_{\kappa}(\mathbf{F}' \mathbf{T}), \quad \exp_{\kappa}(\mathbf{v}) = \frac{\tanh(\sqrt{-\kappa} \|\mathbf{v}\|)}{\sqrt{-\kappa} \|\mathbf{v}\| + \epsilon} \mathbf{v} \quad (3)$$

where κ is the curvature parameter of the hyperbolic space, $\|\mathbf{v}\|$ refers to the Euclidean norm of the vector \mathbf{v} , and ϵ is a small constant for numerical stability. Finally, the hyperbolic transformed features \mathbf{Z} are permuted and reshaped to recover the original spatial dimensions, producing the final output $\mathbf{Y} \in \mathbb{R}^{B \times C_{\text{out}} \times H \times W}$, as described in Equation (4):

$$\mathbf{Y} = \text{Reshape}(\text{Permute}(\mathbf{Z}), [B, C_{\text{out}}, H, W]) \quad (4)$$

As illustrated in Figure 2, the Hyperbolic Convolution Module projects feature maps into a hyperbolic space to better capture the latent anatomical hierarchy. This embedding enables compact representations of tree-like tumor boundaries and nested liver structures, which are challenging for Euclidean CNNs. The module output \mathbf{Y} serves directly as the input for the subsequent Adaptive Temporal Attention stage.

D. Adaptive Temporal Attention

In volumetric data, the slice thickness and spacing can vary across scanners and acquisition protocols. In our study, all scans are first resampled following area-based resampling and nearest-neighbor interpolation during preprocessing for uniformity; and to handle inter-slice dependencies across CT volumes, we introduce an Adaptive Temporal Attention (ATA) mechanism, shown in Figure 3. This ATA dynamically weighs neighboring slices based on learned temporal importance, allowing the network to enhance spatial continuity across volumes while maintaining computational efficiency [20]. Unlike fully 3D convolutional approaches that process entire volumes simultaneously at high computational cost, our ATA mechanism operates on 2D slices while explicitly modeling inter-slice relationships through learned attention weights, thereby preserving volumetric continuity without sacrificing efficiency. The temporal attention component effectively reconstructs inter-slice dependencies by attending to contextually relevant neighboring slices during feature extraction. Equation (5) formulates the attention-enhanced output:

$$\mathbf{F}_{\text{out}} = \mathbf{X} + \left(\mathbf{W}_{\text{temp}} \odot \text{reshape} \left(\text{softmax} \left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}} \right) \cdot \mathbf{V} \right) \right) \quad (5)$$

where $\mathbf{X} \in \mathbb{R}^{B \times C \times H \times W}$ is the input feature map, B is the batch size, C is the number of channels, and $H \times W$ represents the spatial resolution. The query \mathbf{Q} , key \mathbf{K} , and value \mathbf{V} tensors are derived from \mathbf{X} using optimized 1×1 convolutions and reshaped to $\mathbb{R}^{B \times N \times d_k}$, where N is the dot product of $H \cdot W$, and $d_k = C/8$ is the reduced embedding dimension. A scaled dot product attention mechanism computes spatial attention weights via $\text{softmax}(\mathbf{Q}\mathbf{K}^T/\sqrt{d_k})$, which are then applied to \mathbf{V} to generate an attention-weighted representation. This output is reshaped to the original spatial format and further modulated by a trained temporal weight vector $\mathbf{W}_{\text{temp}} \in \mathbb{R}^{B \times 1 \times 1 \times 1}$, derived from the temporal context using a multilayer perceptron. Thus, instead of assuming uniform temporal spacing, ATA dynamically adjusts the contribution of neighboring slices according to their feature correlations. This allows the attention map to dynamically adjust for non-uniform slice intervals, including irregular thickness, by modulating contextual influence according to observed spatial correlations rather than fixed slice indices. Furthermore, by extracting slices from multiple anatomical views, the ATA mechanism captures complementary spatial information from orthogonal directions, effectively approximating the comprehensive context of full 3D analysis. This multi-view approach with adaptive temporal weighting maintains robust inter-slice consistency while offering significant computational efficiency and reduced memory requirements.

The result is then fused with the original input \mathbf{X} through a residual connection, producing \mathbf{F}_{out} , which is enriched in both spatial and temporal dimensions. The output generated by this module is then fed into the Synaptic Plasticity unit.

E. Synaptic Plasticity Mechanism

We designed a synaptic plasticity mechanism inspired by Hebbian learning to enable biologically driven adaptive feature modulation within the framework [32]. This mechanism is based on the fundamental principle of activity-dependent synaptic strengthening observed in neural systems, where connections between co-activated neurons are reinforced over time. While our implementation represents a computational approximation optimized for gradient-based learning rather than a direct model of neurophysiological processes, it incorporates the core concept of correlation-based weight adjustment into a differentiable framework that enhances feature adaptivity in medical image segmentation [33]. Given an input feature tensor \mathbf{X} , the module first computes a channel-wise activity vector for each sample. Specifically, for each sample b in batch B , we obtain a channel activity vector $\mathbf{a}_b \in \mathbb{R}^C$ by globally averaging the spatial responses of the feature tensor b^{th} . Next, for each sample, we computed a correlation matrix $\mathbf{C}_b \in \mathbb{R}^{C \times C}$ as the outer product of the activity vector \mathbf{a}_b with itself.

The batch-averaged correlation matrix $\mathbf{C} \in \mathbb{R}^{C \times C}$ is then obtained by averaging all \mathbf{C}_b over the batch dimension. The diagonal elements of \mathbf{C} , which represent the self-correlation of each channel, are used to update the synaptic trace vector $\boldsymbol{\eta}' \in \mathbb{R}^C$ following the Hebbian update rule, as shown in Equation (6):

$$\boldsymbol{\eta}' \leftarrow \boldsymbol{\eta} + \alpha \text{diag}(\mathbf{C}) - \boldsymbol{\eta} \quad (6)$$

where $\boldsymbol{\eta}$ is a trace vector; α is the plasticity learning rate that regulates the update process; and $\text{diag}(\mathbf{C})$ is the vector formed by the diagonal entries of \mathbf{C} . This update mechanism mirrors the strengthening of synaptic weights in response to correlated neural activity, providing an inductive bias that encourages the network to adaptively emphasize informative feature channels during training. The updated trace vector $\boldsymbol{\eta}'$ modulates a set of learnable synaptic scaling factors $\boldsymbol{\omega} \in \mathbb{R}^C$, producing an adaptive channel-wise gain. This gain adjusts the input tensor \mathbf{X} element-wise. The modulated features are then linearly transformed by a learnable transformation matrix $\mathbf{T} \in \mathbb{R}^{C \times C}$ to capture cross-channel interactions before convolution. Finally, as shown in Equation (7), a convolution with kernel weights $\mathbf{W}_{\text{conv}} \in \mathbb{R}^{C_{\text{out}} \times C \times K \times K}$ refines the transformed features to produce the final output tensor \mathbf{F}_{out} :

$$\mathbf{F}_{\text{out}} = \mathbf{W}_{\text{conv}} * [(\mathbf{X} \odot (\boldsymbol{\omega} \odot (1 + \boldsymbol{\eta}))) \mathbf{T}] \quad (7)$$

where $*$ denotes the convolution operation, \odot represents element-wise multiplication, and \mathbf{T} represents the channel-wise modulated features. This formulation ensures that the output feature map \mathbf{F}_{out} captures not only the information coming from the input, but also the adaptive changes from synaptic plasticity, making the network more realistic and better adapted to learn. The output feature map

\mathbf{F}_{out} from this stage now acts as the input to the Implicit Neural Representation module.

F. Implicit Neural Representation

We developed a continuous coordinate-based neural representation module to support high-resolution segmentation predictions [19]. For each normalized spatial coordinate $\mathbf{p}_{b,n} \in \mathbb{R}^2$ from the b th image in the batch, where values lie within the range $[-1, 1]^2$, we apply a positional encoding function $\gamma(\cdot)$ to effectively capture spatial information across multiple scales. This function encodes each coordinate dimension using sine and cosine functions with progressively increasing frequencies, as described in Equation (8):

$$\gamma(\mathbf{p}) = [\sin(2^l \pi p_j), \cos(2^l \pi p_j)] \quad \text{for } j = 1, 2; l = 0, \dots, L-1 \quad (8)$$

In parallel, for each coordinate \mathbf{p} , a latent feature vector $\varphi(\mathbf{p}, \mathbf{F}) \in \mathbb{R}^C$ is extracted by bilinearly sampling from the global feature map $\mathbf{F}_b \in \mathbb{R}^{C \times H \times W}$ using grid sampling. The multi-layer perceptron f_θ maps the concatenation of the coordinate, its positional encoding, and the sampled feature vector to a 2D logit, as specified in Equation (9):

$$f_\theta : \mathbb{R}^{2+4L+C} \longrightarrow \mathbb{R}^2 \quad (9)$$

Finally, the network output for the logit at each spatial coordinate is computed using this MLP, as described in Equation (10):

$$\mathbf{y}_{b,n} = f_\theta(\mathbf{p}_{b,n}, \gamma(\mathbf{p}_{b,n}), \varphi(\mathbf{p}_{b,n}, \mathbf{F}_b)) \quad (10)$$

This design enables precise, continuous spatial prediction without relying on a fixed-resolution grid. The resulting tensor produced by this module is subsequently utilized as the input for the Uncertainty Estimator. Figure 4 demonstrates the working process of Implicit Neural Representation.

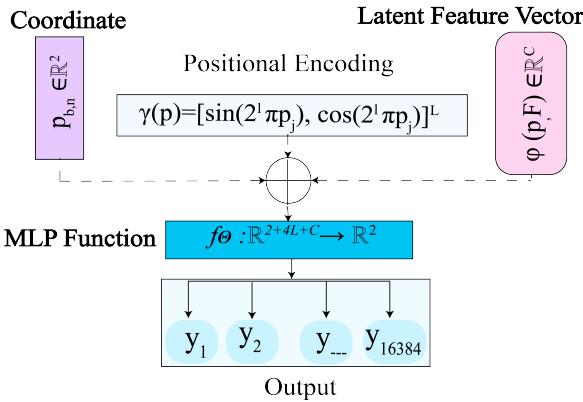


Fig. 4. The process of implicit neural representation, comprised of encoding coordinates with positional encoding, combining them with a latent feature vector, and passing the result through an MLP function to generate the output.

G. Uncertainty Estimator

We have constructed an uncertainty-aware prediction module (see Figure 5) based on Monte Carlo dropout to estimate both

predictive output and its associated confidence [34]. Given an input feature map \mathbf{X} , we first apply a spatial dropout with probability p to simulate stochastic behavior at the test time. During training, a single forward pass is performed applying dropout followed by a lightweight prediction head $f_\theta(\cdot)$, consisting of two convolutional layers and a sigmoid activation, which outputs a confidence map in pixels. During inference, to quantify epistemic uncertainty, we perform stochastic forward passes T by applying dropout independently at each iteration, these passes are indexed by a small variable $t \in \{1, 2, \dots, T\}$, where each $u^{(t)}$ in $f_\theta(\text{Dropout}(\mathbf{X})^{(t)})$ denotes the model prediction from the t th forward pass with a distinct dropout mask. The mean prediction $\mu(\mathbf{u})$ is obtained by averaging these outputs, while the uncertainty is quantified by the variance $\sigma^2(\mathbf{u})$, calculated as the pixel variance across the T samples. The formulations are provided in Equations (11) and (12):

$$\mu(\mathbf{u}) = \frac{1}{T} \sum_{t=1}^T \sigma(f_\theta(\text{Dropout}(\mathbf{X})^{(t)})) \quad (11)$$

$$\sigma^2(\mathbf{u}) = \frac{1}{T} \sum_{t=1}^T (\sigma(f_\theta(\text{Dropout}(\mathbf{X})^{(t)})) - \mu(\mathbf{u}))^2 \quad (12)$$

This approach enables the model to express uncertainty in a principled, data-driven manner without requiring architectural changes or additional supervision.

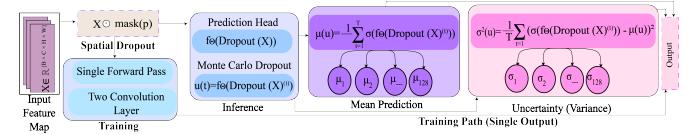


Fig. 5. The process of uncertainty estimator using Monte Carlo Dropout, where the input feature map undergoes spatial dropout, followed by inference, mean prediction, and uncertainty estimator

IV. EXPERIMENTAL DETAILS

In this section, we elaborate on the dataset preparation tasks along with the model's preprocessing and implementation setup.

A. Data Preparation

1) *Datasets:* In our study, we used two public datasets, including LiTS and AMOS 2022, to evaluate our model's performance and generalization capability. It is worth noting that since the official LiTS test set (70 samples) was not publicly available, we limited our analysis to the 131 publicly released training cases to ensure full reproducibility for all researchers. We randomly split these available 131 samples into an 80:20 ratio, resulting in 105 cases for training and 26 cases for testing. To demonstrate the cross-dataset generalization and modality robustness of our model, we also evaluated our model on the AMOS 2022 dataset, which contains both CT and MRI scans with 15 annotated abdominal organs. Although the AMOS 2022 dataset has a larger number of samples, we randomly selected 70 samples (unseen

by the model), containing both CT and MRI scans, for cross-dataset validation to assess the model's robustness across different imaging modalities and anatomical variations.

Liver Tumor Segmentation Challenge (LiTS). The LiTS dataset [35] comprises 131 (publicly available) volumetric abdominal CT scans, each with a pixel resolution of 512×512 and varying voxel spacings, with in-plane resolutions between 0.60 and 0.98 mm and slice thicknesses ranging from 0.45 to 5.0 mm. It includes a variety of liver-related abnormalities, such as tumors, metastases, and cysts. Tumor volumes vary from 38 mm^3 to 1231 mm^3 , and the number of tumors per scan ranges from 0 to 12. From these scans, we extract a total of 15,654 2D tumor-containing slices from all three anatomical views (4,108 axial, 5,730 coronal, and 5,816 sagittal). These selected slices are then divided into training (10,957), validation (2,348), and test (2,349) sets, ensuring that both liver parenchyma and tumor regions are present for effective segmentation learning.

AMOS 2022. The AMOS dataset [36] is a large-scale abdominal multi-organ benchmark containing 500 CT and 100 MRI scans collected from multi-center, multi-scanner clinical environments. Each scan provides voxel-level annotations for 15 abdominal organs, including the spleen, kidneys, gallbladder, esophagus, liver, stomach, aorta, postcava, pancreas, adrenal glands, duodenum, bladder, and prostate. The dataset features multi-modality, multi-phase, and multi-disease cases, capturing a broad spectrum of imaging conditions and pathological variations.

2) *Data Preprocessing:* We preprocess the 3D medical images to standardize their format and intensity ranges. Area-based resampling is applied to CT volumes and nearest-neighbor interpolation to segmentation masks to preserve anatomical boundaries, while retaining only slices with visible tumors, including subtle lesions. Windowing is applied by clipping each voxel intensity to the range $[-200, 400]$ and linearly rescaling it to $[0, 1]$, enhancing contrast around soft tissues. Intensity normalization is applied after windowing by dividing the clipped value range into equal intervals across all slices. This preprocessing is performed across axial, coronal, and sagittal views, with each 2D slice extracted from the original 3D volume at a resolution of 512×512 . Our framework processes these slices at a resized in-plane resolution of 128×128 due to computational resource constraints. This resizing is applied uniformly to all slices to maintain depth and ensure uniform spatial dimensions [37].

B. Implementation Setup

This research is conducted using an AMD Ryzen 5 5600X 6-core Central Processing Unit (CPU) and 16 GB of RAM for all experiments. For graphical processing, a ZOTAC GAMING GeForce RTX 3060 Twin Edge OC with 12 GB of video RAM (VRAM) is utilized, while Jupyter Notebook version 7.0.8 serves as the integrated development environment (IDE).

V. RESULTS

A. Performance Assessment

Segmentation performance of the proposed model. To evaluate the segmentation performance of our proposed HANS-Net frame-

work, we conducted comprehensive experiments on the LiTS dataset, focusing on accurate liver and tumor delineation. The architecture integrates hyperbolic convolutions, wavelet-inspired decomposition, synaptic plasticity mechanisms, and implicit neural representations to capture both global organ structure and local tumor details within a continuous and geometrically informed space. All CT volumes were resampled to a uniform resolution of 128×128 pixels, and the model was trained for 100 epochs using the Adam optimizer with a learning rate of 0.001 and a batch size of 16. As reported in Table I, HANS-Net achieves a Dice

TABLE I
SEGMENTATION PERFORMANCE OF THE PROPOSED HANS-NET MODEL ON THE LiTS DATASET FOR LIVER AND TUMOR STRUCTURES.

Organ	Dice (%)	IoU (%)	ASSD (mm)	VOE (%)
Liver	96.67	93.79	0.54	6.21
Tumor	89.84	82.39	0.89	17.61
Mean	93.26	88.09	0.72	11.91

score of 96.67% for liver and 89.84% for tumors, demonstrating high segmentation precision across both large organ regions and smaller tumor structures. The corresponding IoU values of 93.79% (liver) and 82.39% (tumor) confirm strong spatial overlap between predictions and ground truth. Furthermore, in terms of boundary alignment, the model attains a low Average Symmetric Surface Distance (ASSD) of 0.54 mm for the liver and 0.89 mm for tumors, while the Volume Overlap Error (VOE) remains limited to 6.21% and 17.61%, respectively. The observed performance gap between liver and tumor segmentation reflects the inherent difficulty of tumor delineation due to substantial heterogeneity in size, shape, contrast, and diffuse boundaries. The synaptic plasticity module and implicit neural representation branch specifically address these challenges by adaptively refining tumor-relevant features and modeling continuous boundaries. These results validate that HANS-Net effectively integrates geometric and biologically inspired components to produce accurate and anatomically consistent liver tumor segmentation.

Training Dynamics and Convergence. The training and validation curves for our proposed HANS-Net model demonstrate effective convergence and stable performance in liver and tumor segmentation tasks. As shown in Figure 6A, both training and validation losses decreased steadily from initial values around 0.7 to approximately 0.1, with training loss converging slightly faster than validation loss, indicating good model generalization without significant overfitting. The progression of the Dice score reveals distinct learning patterns between organ types. Training Dice scores for liver segmentation in Figure 6B rapidly increased from 10% to more than 90% within the first 20 epochs, demonstrating the model's ability to quickly capture large anatomical structures.

In contrast, tumor Dice scores showed a more gradual improvement, reaching approximately 80% by epoch 80, reflecting the inherent difficulty of segmenting smaller, more heterogeneous tumor regions. Validation Dice scores in Figure 6D followed similar trends, with liver achieving over 90% and tumor reaching approx-

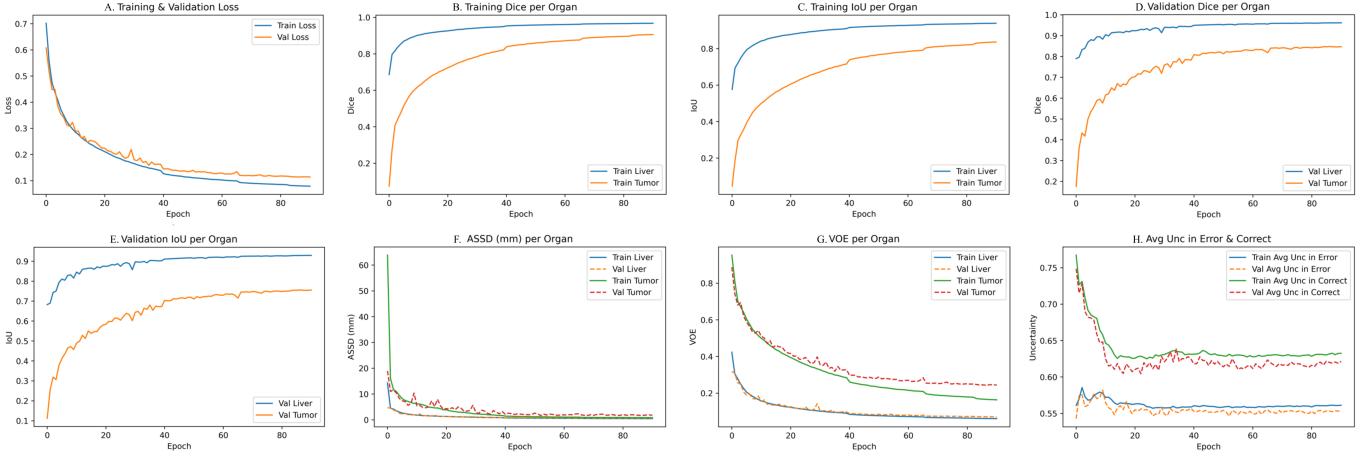


Fig. 6. Accuracy and loss curves of our proposed HANS-Net model.

imately 80%, confirming robust generalization performance. IoU metrics in Figure 6C and 6E mirrored the Dice score patterns, with liver IoU rapidly increasing to 90% during training and tumor IoU gradually improving to 75%. The faster convergence of liver metrics compared to tumor metrics highlights the model's hierarchical learning approach, where larger anatomical structures are learned before fine-grained tumor boundaries. The ASSD values in Figure 6F decreased to minimal levels below 5mm for both organs, indicating a precise boundary delimitation. The VOE metrics in Figure 6G decreased substantially during early training, stabilizing around 0.2 for the liver and 0.3 for the tumor, while the analysis in Figure 6H showed decreasing uncertainty levels for both correct and incorrect predictions, and the model became more confident in its predictions as training progressed. These comprehensive training dynamics validate that our HANS-Net framework effectively integrates hyperbolic convolutions, wavelet decomposition, and synaptic plasticity mechanisms to achieve robust liver tumor segmentation with distinct learning trajectories optimized for different anatomical structures.

B. Comparison With State-of-the-Art Methods

We compared the proposed HANS-Net model with several SOTA liver and tumor segmentation methods, including U-Net [38], U-Net++ [39], Attention U-Net [24], H-DenseUNet [40], TransUNet [41], and nnU-Net [42]. To ensure fair comparison with architectures originally designed for 3D or hybrid 2D/3D processing, we adapted them to our slice-based evaluation by processing CT slices individually through their 2D encoder-decoder pathways while retaining original components such as attention gates and dense blocks. Attention U-Net's spatial attention was applied at each decoder level, and H-DenseUNet's hybrid dense blocks were implemented in 2D convolutional layers with consistent hyperparameters. All baselines used the same preprocessing, multi-view slice extraction, Adam optimizer with a learning rate of 1e-4, batch size of 16, and training for 100 epochs, ensuring a consistent evaluation framework across all models. The proposed HANS-Net framework consistently surpasses all SOTA methods

in most metrics. Although TransUNet [41] achieved a marginally lower ASSD of 0.51 mm for liver, where H-DenseUNet [40] recorded a slightly lower VOE of 17.54% for tumors. The nnU-Net [42], on the other hand, showed competitive liver segmentation performance and achieved a second-highest Dice score of 96.20% in the liver segmentation task [35]. Compared to the existing models, our model achieved the highest Dice scores of 96.67% for liver and 89.84% for tumor, which is +0.27%, and +2.92% Dice score improvement compared to nnU-Net and TransUNet, respectively. The model also recorded the highest IoU scores of 93.79% for liver and 82.39% for tumor, and the lowest average ASSD of 0.72 mm and VOE of 11.91%.

TABLE II
COMPARATIVE PERFORMANCE OF HANS-NET AND SOTA MODELS ON THE LITS DATASET FOR LIVER AND TUMOR SEGMENTATION. METRICS INCLUDE DICE (%), IOU (%), ASSD, AND VOE (%). THE BEST AND SECOND-BEST RESULTS FOR EACH DATASET ARE **BOLDED** AND UNDERLINED, RESPECTIVELY.

Model	Organ	Dice \uparrow	IoU \uparrow	ASSD(mm) \downarrow	VOE \downarrow
U-Net [38]	Liver	92.18	86.21	1.32	13.79
	Tumor	82.36	74.52	1.76	25.48
	Mean	87.27	80.37	1.54	19.64
U-Net++ [39]	Liver	93.01	87.63	1.12	12.37
	Tumor	83.49	76.05	1.62	23.95
	Mean	88.25	81.84	1.37	18.16
Attention U-Net [24]	Liver	94.07	89.24	0.98	10.76
	Tumor	85.34	78.18	1.42	21.82
	Mean	89.71	83.71	1.20	16.29
H-DenseUNet [40]	Liver	93.63	88.72	1.04	11.28
	Tumor	84.89	77.03	1.51	17.54
	Mean	89.26	82.88	1.28	17.41
TransUNet [41]	Liver	94.82	<u>90.13</u>	0.51	<u>9.87</u>
	Tumor	<u>86.92</u>	<u>79.88</u>	1.33	20.12
	Mean	90.87	85.01	0.92	15.00
nnU-Net [42]	Liver	<u>96.20</u>	89.36	2.56	10.44
	Tumor	73.90	78.93	<u>0.90</u>	19.96
	Mean	85.05	85.22	1.73	15.20
HANS-Net (ours)	Liver	96.67	93.79	<u>0.54</u>	6.21
	Tumor	89.84	82.39	<u>0.89</u>	<u>17.61</u>
	Mean	93.26	88.09	0.72	11.91

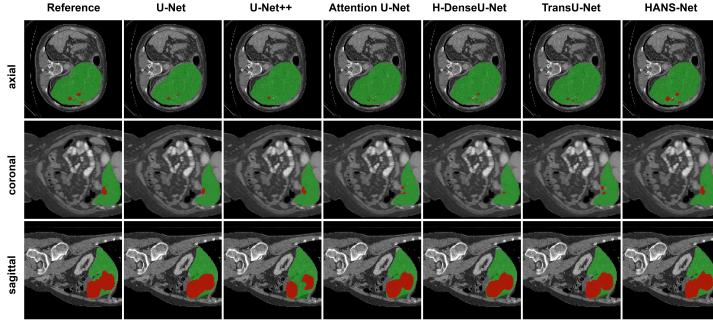


Fig. 7. Comparative performance of the proposed HANS-Net model and baseline methods (U-Net, U-Net++, Attention U-Net, H-DenseUNet, and TransUNet) for liver and tumor segmentation on the LiTS dataset.

Furthermore, the mean segmentation quality between both organs is the most consistent and accurate, confirming the superiority of HANS-Net in anatomical boundary preservation and tumor detection accuracy. Table II presents detailed comparative performance in Dice, IoU, ASSD, and VOE, demonstrating the superiority of our HANS-Net framework in accurate, robust, and anatomically consistent liver and tumor segmentation. Figure 7 visualizes the superior segmentation capabilities of our approach compared to the SOTA methods tested.

C. Comparison With Existing Literature

Our proposed method, HANS-Net, demonstrates superiority over existing models by achieving the highest Dice score of 93.26% on the LiTS dataset. This outcome proves HANS-Net's enhanced capability in the segmentation of liver tumors compared to recent SOTA methods.

Previous models, such as MSML-AttUNet [16], attained performances of 87.74% on combined datasets like LiTS and 3DIRCADb. Although models like EG-UNETR [12] exhibited commendable performance at 85.01% and S2DA-Net [11] demonstrated moderate results with 80.0%, HANS-Net presents a marked improvement. Models such as MCDA [17] and G-Unet [13] recorded lower scores, suggesting constraints in generalization or segmentation precision. Even methodologies employing private datasets, such as MCFMFNet [8], did not exceed the 90% threshold. The comparative analysis highlights the strength and practical effectiveness of HANS-Net, showing that it has a distinct advantage over the existing SOTA models. Table III provides a detailed comparison between the proposed approach and the SOTA techniques.

D. Cross-Dataset Evaluation

To further assess the model's cross-dataset generalization capabilities, we evaluated the proposed HANS-Net model on the AMOS 2022 dataset, which contains 15 annotated abdominal organs.

Figure 8 and Table IV present the overall segmentation performance of our proposed model on the AMOS 2022 test set. The results demonstrate robust performance across all anatomical

TABLE III
COMPARISON OF DIFFERENT MODELS AND THEIR DICE SCORES ON VARIOUS DATASETS. THE BEST AND SECOND-BEST RESULTS FOR EACH DATASET ARE **BOLDED** AND UNDERLINED, RESPECTIVELY.

Ref.	Model	Dataset	Dice
[5]	SBCNet	LiTS	82.50%
[8]	MCFMFNet	LiTS	84.13%
[11]	S2DA-Net	LiTS	80.00%
[12]	EG-UNETR	LiTS	<u>84.45%</u>
[13]	G-Unet	LiTS	72.90%
[15]	PGC-Net	LiTS	73.63%
[23]	DPC-Net	LiTS	76.40%
[12]	EG-UNETR	3D-IRCADb	85.01%
[15]	PGC-Net	3D-IRCADb	74.16%
[8]	MCFMFNet	Private	73.65%
[17]	MCDA	Private	79.22%
[26]	TMPLITS	Private	81.07%
Ours	HANS-Net	LiTS	93.26%
		AMOS 2022	<u>85.09%</u>

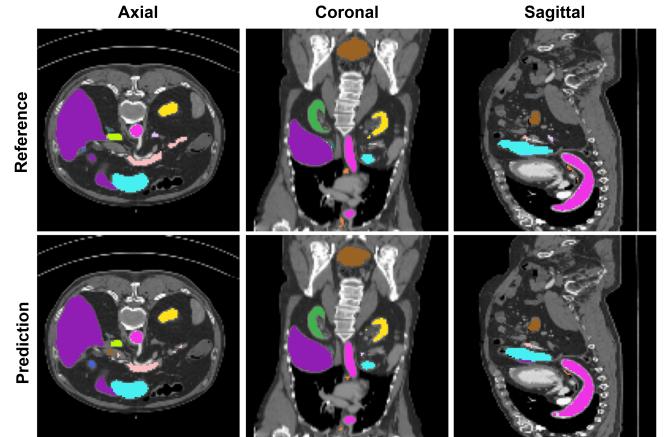


Fig. 8. Cross-dataset prediction visualizations from our HANS-Net model on AMOS 2022 dataset.

structures, with our model achieving an overall average Dice score of 85.09%, IoU of 76.66%, ASSD of 19.49 mm, and VOE of 23.34%. Notably, the model performed exceptionally well on large solid organs such as the liver (94.17% Dice) and the postcava (94.13% Dice), as well as the gallbladder (88.56% Dice). Even for challenging structures like the duodenum and prostate/uterus, which are typically difficult to segment due to their complex anatomical variations and low contrast, our model achieved Dice scores of 80.72% and 79.89%, respectively. The low ASSD values for liver (8.35 mm) and postcava (9.65 mm) further validate the effectiveness of our implicit neural representation in maintaining precise boundary delineation.

E. Uncertainty-Aware Performance Evaluation

The proposed HANS-Net model integrates Monte Carlo dropout-based uncertainty estimation to provide clinically significant reliability measures alongside segmentation predictions [43].

TABLE IV

SEGMENTATION PERFORMANCE OF HANS-NET ON THE AMOS 2022 DATASET ACROSS 15 ABDOMINAL ORGANS. THE ABBREVIATED ORGANS ARE: RIGHT KIDNEY (R. KID.); LEFT KIDNEY (L. KID.); STOMACH (STOM.); PANCREAS (PANC.); RIGHT ADRENAL GLAND (R. ADR.); LEFT ADRENAL GLAND (L. ADR.); DUODENUM (DUOD.); URINARY BLADDER (BLAD.); PROSTATE (PROST.).

Metric	Spleen	R. Kid.	L. Kid.	Gallbladder	Esophagus	Liver	Stom.	Aorta	Postcava	Panc.	R. ADR.	L. ADR.	Duod.	Blad.	Prost.	Mean
Dice (%)	81.37	83.36	80.85	88.56	81.51	94.17	85.52	80.96	94.13	86.23	87.65	87.90	80.72	83.53	79.89	85.09
IoU (%)	75.93	78.84	73.16	75.58	75.06	80.12	74.42	69.83	88.94	79.48	84.32	75.24	71.05	77.90	70.08	76.66
ASSD (mm)	20.63	26.31	20.37	15.96	24.94	8.35	16.72	25.89	9.65	18.88	17.75	13.53	21.39	19.44	32.49	19.49
VOE (%)	24.07	21.16	26.84	24.42	24.94	19.88	25.58	30.17	11.06	20.52	15.68	24.76	28.95	22.10	29.92	23.34

TABLE V

STATISTICAL CONSISTENCY AND AGREEMENT METRICS BETWEEN HANS-NET PREDICTED AND GROUND-TRUTH LIVER AND TUMOR VOLUMES, INCLUDING CORRELATION COEFFICIENTS, BIAS, MEAN ABSOLUTE ERRORS, NORMALIZED ERRORS, 95% LIMITS OF AGREEMENT, AND UNCERTAINTY RANGES.

Metric	Liver	Tumor	Avg.
Pearson Correlation (r)	0.999	0.995	0.997
Spearman Correlation (ρ)	0.998	0.993	0.995
Coefficient of Determination (R^2)	0.998	0.990	0.994
Relative Volume Difference (%)	2.0	6.6	4.3
Mean Ground Truth Volume (mL)	1359.9	252.0	N/A
Mean Predicted Volume (mL)	1362.6	252.3	N/A
Mean Bias Error (mL)	+2.7	+0.3	+1.5
Mean Absolute Error (mL)	25.8	13.6	19.7
Normalized MAE (% of GT)	1.9	5.4	3.65
95% Limits of Agreement (mL)	[−47.87, +53.27]	[−26.36, +26.9]	N/A
Uncertainty Range	0.595 – 0.637	0.619	

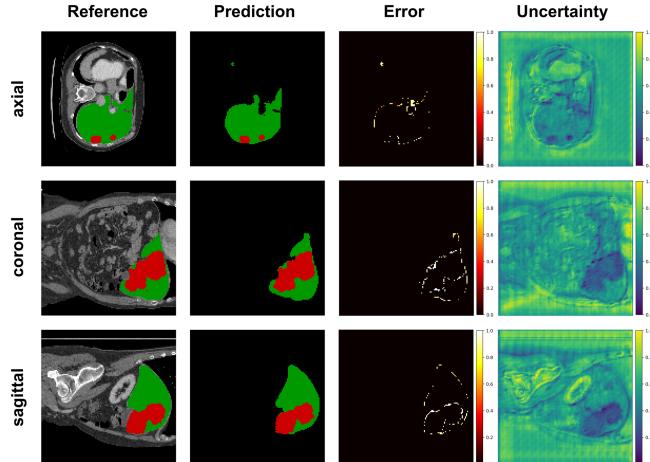


Fig. 9. Multi-view uncertainty analysis of HANS-Net model on LiTS dataset.

As shown in Figure 9, the uncertainty heatmaps highlight regions of lower model confidence, particularly at organ boundaries and areas of tumor heterogeneity. Quantitative evaluation in all test samples provides consistent uncertainty measures with mean values of 0.619 for liver and tumor predictions; and for both predictions, the uncertainty range was observed between 0.595 and 0.637 (see Table V). The narrow range of 0.042 indicates stable model behavior across different cases. These values represent the uncertainty characteristics observed in our liver and tumor segmentation task, where the model successfully distinguishes between confident and uncertain predictions. While

our model demonstrates effective uncertainty quantification for liver and tumor segmentation, the determination of acceptable uncertainty thresholds is application-specific and would require domain expert validation for different medical imaging tasks and clinical contexts.

F. Volumetric Correlation Analysis

As shown in Table V, the HANS-Net model demonstrates robust clinical validity through its volumetric measurements, with near-perfect Pearson’s correlation coefficients of 0.999 for liver and 0.995 for tumor volumes alongside Spearman’s correlations of 0.998 and 0.993, respectively. All correlations show statistical significance with p-values below 1×10^{-100} .

These strong correlations are supported by the Coefficients of Determination, R^2 (0.998 for liver and 0.990 for tumor) and show that over 99% of the variance in the ground-truth volumes is captured by the model predictions. The mean predicted volumes of 1362.6 mL for liver and 252.3 mL for tumors closely match the ground-truth means of 1359.9 mL and 252.0 mL, and this results in negligible Mean Bias Errors of +2.7 mL and +0.3 mL. The Mean Absolute Errors are 25.8 mL for liver and 13.6 mL for tumor, and they correspond to Normalized Errors of 1.9% and 5.4%. In addition to these findings, the reported 95% limits of agreement are [−47.9, +53.3] mL for liver and [−26.4, +27.0] mL for tumor, and most predictions remain within a narrow and interpretable range of the true volumes. These results collectively demonstrate that HANS-Net provides statistically reliable and clinically interpretable volumetric estimations. Figure 10 visualizes the model’s measurement reliability through the close alignment of data points with the identity line.

G. Ablation Study

We conducted a systematic ablation study to analyze the effect of progressive enhancement achieved by HANS-Net. In this section, we explore the main components, including hyperbolic convolution (HC), wavelet decomposition module (WDM), adaptive temporal attention (ATA), synaptic plasticity module (SPM), implicit neural representation (INR), and uncertainty estimator (UE). The configuration details are reported in Table VI.

Initially, we experimented with a standard 3D U-Net architecture (A0) that operates within Euclidean space, where feature correlations are defined through linear kernel operations and assumed flat spatial geometry. We observed that while the configuration was effective for local intensity modeling, this formulation

TABLE VI

ABLATION STUDY ON THE LiTS DATASET WITH AVERAGED LIVER-TUMOR METRICS. THE CONFIGURATIONS INCLUDE COMPONENTS SUCH AS HYPERBOLIC CONVOLUTION (HC), WAVELET DECOMPOSITION MODULE (WDM), ADAPTIVE TEMPORAL ATTENTION (ATA), SYNAPTIC PLASTICITY MODULE (SPM), IMPLICIT NEURAL REPRESENTATION (INR), AND UNCERTAINTY ESTIMATOR (UE). THE CONFIGURATIONS ARE EXPERIMENTED WITH FOR 100 EPOCHS, AND THE INFERENCE SPEED FOR THE CONFIGURATIONS IS PROVIDED IN SECONDS IN THE INF. TIME COLUMN.

Exp.	Components						Dice(%) ↑		IoU(%) ↑		ASSD (mm) ↓		VOE(%) ↓		Inf. Time ↓	Params (M)
	HC	WDM	ATA	SPM	INR	UE	Train	Validation	Train	Validation	Train	Validation	Train	Validation		
A0	✗	✗	✗	✗	✗	✗	87.62	83.22	81.84	75.65	1.65	1.69	18.05	25.42	100×238s	0.93M
A1	✗	✗	✗	✗	✗	✗	87.55	84.84	79.45	76.20	1.45	1.75	20.55	23.80	100×230s	0.89M
A2	✓	✗	✗	✗	✗	✗	89.60	87.23	82.58	79.18	1.24	1.46	17.43	20.83	100×235s	0.98M
A3	✓	✓	✗	✗	✗	✗	90.79	88.36	84.35	81.13	1.05	1.21	15.65	18.88	100×238s	1.05M
A4	✓	✓	✓	✗	✗	✗	91.89	89.63	86.03	82.78	0.94	1.07	13.98	17.23	100×240s	1.12M
A5	✓	✓	✓	✓	✗	✗	92.60	90.38	87.33	84.00	0.83	0.97	12.68	16.00	100×242s	1.15M
A6	✓	✓	✓	✓	✓	✗	93.18	91.03	88.20	84.85	0.77	0.91	11.80	15.15	100×243s	1.17M
A7	✓	✓	✓	✓	✓	✓	93.86	91.47	89.17	85.86	0.66	0.89	10.83	14.77	100×245s	1.17M

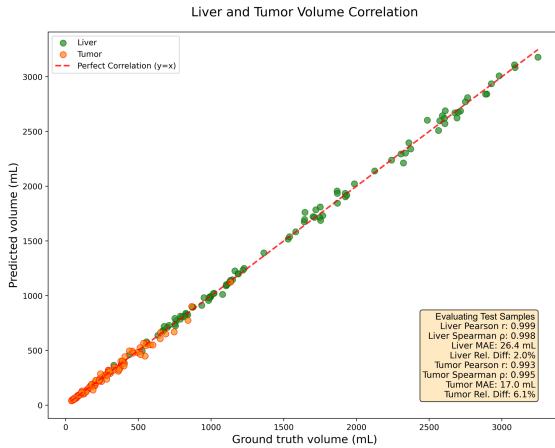


Fig. 10. Volumetric correlation analysis of HANS-Net model on LiTS dataset.

could not accurately represent the curved anatomical surfaces commonly present in volumetric data. As a result, the Euclidean baseline exhibited limited geometric expressiveness in validation across the defined metrics despite having higher training performance. These outcomes indicate that Euclidean convolutions primarily learn low-level texture dependencies (which requires more time) rather than topological or structural consistency across slices. Thus, in the next configuration, our designed baseline encoder-decoder architecture (A1) adopts a hierarchical architecture consisting of four encoding and decoding stages. Each stage includes two 3×3 convolutional layers followed by batch normalization and ReLU activation, starting with 32 channels and doubling at each subsequent level (32, 64, 128, 256). Downsampling is achieved using 2×2 max-pooling, while upsampling utilizes transposed convolutions with concatenated skip connections. The output layer applies a 1×1 convolution with sigmoid activation to generate pixel-wise predictions. Training was performed using the Adam optimizer (learning rate = $1e^{-4}$) with a combined Dice and Cross-Entropy loss and ReduceLROnPlateau scheduler for 100 epochs and a batch size of 16. This baseline establishes the

foundation with conventional convolutions and shows a notable train-validation gap that motivates our architectural innovations. Then, instead of traditional Euclidean convolutions, we employ hyperbolic convolutions (A2) to capture hierarchical anatomical structures in curved space, delivering a 2.39% Dice improvement and 16.57% ASSD reduction while enhancing IoU by approximately 3%.

Next, rather than relying on single-scale processing, we introduce the wavelet decomposition module (A3), which enables multifrequency analysis for fine-detail detection. This addition provides an additional 1.13% dice gain and an improvement of 17.12% ASSD with minimal computational overhead. Furthermore, in place of standard feature aggregation, the adaptive temporal attention module (A4) is incorporated to model slice-wise correlations, achieving a 1.27% Dice enhancement and an 11.57% reduction in ASSD, thus improving boundary precision.

Furthermore, unlike static feature learning, we integrate a synaptic plasticity module (A5), introducing bio-inspired adaptation mechanisms. This module contributes an additional 0.75% Dice improvement and 9.35% ASSD reduction while helping minimize train-validation discrepancies. Furthermore, instead of using discrete feature representations, we employ implicit neural representations (A6), which enable continuous coordinate-based learning. This technique delivers a 0.65% dice gain and a 6.19% ASSD improvement, all while maintaining exceptional parameter efficiency.

Finally, rather than producing deterministic results, we integrate uncertainty quantification (A7) to provide confidence estimation, contributing to another 0.44% Dice improvement. The complete HANS-Net framework demonstrates strong component synergy, outperforming individual contributions by 6.63% in Dice and achieving a 49.14% overall ASSD improvement, all while preserving computational efficiency.

To verify that performance gains are not artifacts of integration order, we conducted permutation experiments by altering the module sequence, such as incorporating WDM before HC or introducing INR earlier. Across five randomized permutations, the average validation Dice varied by only 0.31%, with ASSD

fluctuations under 0.08 mm, indicating order-independent contributions. This stability arises because each component targets orthogonal aspects: HC refines geometric encoding, WDM captures multiscale textures, ATA enforces temporal coherence, SPM adapts features dynamically, INR provides continuous refinement, and UE quantifies confidence. Their modular design ensures minimal interference, confirming that improvements reflect genuine architectural benefits rather than sequence-dependent interactions.

Analyzing the computational parameter and inference time, we notice a gradual increase across the configurations that reflects the cumulative computational impact of each added module, corresponding to deeper feature hierarchies and denser inter-module communication. The introduction of hyperbolic convolutions and the wavelet decomposition module slightly extends forward-pass operations due to additional manifold projections and frequency decompositions, yet these remain linear in parameter growth. Subsequent modules, such as adaptive temporal attention and synaptic plasticity, add minor tensor-matrix interactions that marginally increase runtime but significantly enhance representational adaptability. The near-constant epoch increment from 230s to 245s indicates that the additional parameters (0.28M overall) primarily improve weight expressiveness rather than computational depth, meaning the model achieves richer spatiotemporal encoding without a proportional rise in latency.

VI. DISCUSSION

Our proposed model comprises several advanced modules that enhance its robustness and overall performance. The incorporation of hyperbolic convolutions [31] for hierarchical geometric representation, wavelet-based [30] decomposition for multiscale segmentation, synaptic plasticity [32] for adaptive feature enhancement, and an implicit neural representation [19] branch for precise boundary modeling collectively contributed to enhanced segmentation accuracy. To further quantify the confidence of prediction, an uncertainty-aware Monte Carlo dropout is used [34]. We further integrated adaptive temporal attention to enhance the consistency of inter-slice without decreasing efficiency [20]. While our method operates on 2D slices for computational efficiency, volumetric continuity is preserved through multi-view extraction and adaptive temporal attention that explicitly models sequential relationships between adjacent slices. Since the liver channel encompasses both parenchyma and tumor, all liver-containing slices are evaluated regardless of tumor presence, maintaining sensitivity to potential false positives in tumor-free regions.

The integration of implicit neural representation contributed to the generation of smooth and precise organ boundaries, which significantly improved the quality of segmentation [19]. Moreover, our approach effectively captures both global structures and fine local details, which played a crucial role in boosting the Dice scores. To further improve performance, we incorporated a lightweight self-attention mechanism with reduced channel dimensions, utilized a multilayer perceptron for learning adaptive temporal weighting, and optimized the query-key-value attention

architecture to enhance consistency across slices [20]. The addition of such other advanced modules, such as hyperbolic convolution, uncertainty estimator, and implicit neural representation, greatly increased the robustness of the model, leading to superior segmentation performance.

The proposed HANS-Net model demonstrates significant segmentation performance in the LiTS dataset [35] for both liver and tumor structures. With respect to liver segmentation, the model achieved a Dice score of 96.67% and an IoU of 93.79%, which signifies an excellent score for segmenting the liver. The model also showed a low ASSD of 0.54 mm and a minimal VOE of 6.21%, indicating precise boundary definition and low volumetric error. Although the results for tumor segmentation are slightly lower due to increased structural complexity, they remain robust, with a Dice score of 89.84%, IoU of 82.39%, ASSD of 0.89 mm, and VOE of 17.61%. In summary, the model achieved a mean Dice score of 93.26%, IoU of 88.09%, ASSD of 0.72 mm, and VOE of 11. 91%, thus confirming its efficiency in segmenting large and small anatomical structures.

The robustness of HANS-Net is further evidenced by its stable convergence seen in both the training and validation curves, which show no major signs of overfitting. Interestingly, liver Dice scores rapidly surpassed 90% within the first 20 epochs, while tumor Dice scores increased more slowly, reaching about 80% in epoch 80. This trend highlights a hierarchical learning process that first captures broad anatomical structures before focusing on tumor boundary details. Metrics such as ASSD and VOE showed consistent decreases, pointing to precise segmentation and a growing confidence in the model. The evaluation of the cross-dataset on the AMOS 2022 [36] dataset confirms the impressive generalizability of the model, with a mean Dice score of 85.09%, an IoU of 76.66%, and low ASSD (19.49 mm) and VOE (23.34%).

Compared to recent SOTA models, HANS-Net achieves superior segmentation performance, attaining a mean Dice score of 93.26%, whereas most existing methods struggle to surpass the 90% threshold. For example, MSML-AttUNet [16] reached Dice scores of 87.74%, respectively, while models such as MCDA [17] and G-UNet [13] performed even lower. Among pre-trained architectures, only TransUNet [2] exceeded 90%, yet it remained approximately 3% lower than HANS-Net, highlighting the advantage of our framework in both accuracy and structural consistency.

As shown in Table VI, an ablation study reveals the progressive benefit of our proposed modules. The base configuration starts with a Dice score of 87.55%, which rises to 89.60% after integrating hyperbolic convolution. Incorporating all components, including synaptic plasticity, temporal attention, and uncertainty modeling, culminates in a final Dice score of 93.86%, demonstrating the additive effectiveness of each module. The inference time and model size analysis further confirm its suitability for clinical deployment. With only 0.28M additional parameters and a minimal increase of +15s in processing time, the model maintains practical efficiency and can be deployed in standard hospital GPU workstations without requiring specialized hardware.

While HANS-Net demonstrates promising results, it has several limitations. First, although our evaluation computes standard

segmentation metrics (Dice, IoU, ASSD, VOE) at the slice level, future work should report these metrics aggregated at the per-volume level to align with the original LiTS benchmark protocol. Additionally, while processing resized 2D slices, the model utilizes dynamic sequential weighting to maintain volumetric continuity, a mechanism that could be further adapted to achieve complete 3D reconstruction. Second, in cross-dataset evaluation on the AMOS 2022 dataset, the model's performance showed a moderate decline, with the Dice score dropping to 85.09% and the IoU to 76.66%. This decrease can be attributed to the high anatomical diversity of abdominal organs, variations in imaging modalities, and inter-organ boundary ambiguities present in the dataset. Third, although the implicit representation branch enhances boundary continuity, it introduces additional computational overhead and increases inference time.

VII. CONCLUSION

We develop and thoroughly evaluate HANS-Net, an innovative segmentation framework that achieves remarkable Dice scores for both liver and tumor segmentation. The key experimental findings and theoretical conclusions are summarized as follows:

- HANS-Net attains high segmentation accuracy on the LiTS dataset, with Dice scores of 96.67% for the liver and 89.84% for tumors, and IoU values of 93.79% and 82.39%, respectively.
- Cross-dataset evaluation on the AMOS 2022 dataset achieves Dice scores of 85.09% and IoU of 76.66% across 15 abdominal organs, demonstrating strong generalization ability across diverse anatomical structures and imaging variations.
- The significance of this work lies in its six integrated modules, which collectively define organ boundaries, extract multi-level features, and quantify prediction confidence, supporting the robust performance of the model.

Future work will focus on enhancing generalization across diverse imaging modalities while optimizing runtime efficiency for clinical deployment. While our current evaluation focuses on standard slice-level segmentation metrics, including Dice, IoU, ASSD, and VOE, future work should consider reporting these metrics at the volumetric level to maintain consistency with the original LiTS benchmark protocol. Additionally, while the model operates on resized 2D slices, it employs dynamic sequential weighting to preserve volumetric continuity, a mechanism that can be further adapted to explicitly enable 3D reconstruction capabilities.

ACKNOWLEDGMENTS

The authors declare that they have no financial conflicts of interest that could have influenced this work.

DATA AVAILABILITY

In this study, we used two publicly available datasets, including [Liver Tumor Segmentation Challenge \(LiTS\) dataset](#) [35], and [AMOS 2022 dataset](#) [36].

REFERENCES

- [1] F. Zhan, W. Wang, Q. Chen, Y. Guo, L. He, and L. Wang, "Three-direction fusion for accurate volumetric liver and tumor segmentation," *IEEE journal of biomedical and health informatics*, vol. 28, no. 4, pp. 2175–2186, 2023.
- [2] Y. Zhang, C. Peng, L. Peng, Y. Xu, L. Lin, R. Tong, Z. Peng, X. Mao, H. Hu, Y.-W. Chen *et al.*, "Deeprecs: from recist diameters to precise liver tumor segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 2, pp. 614–625, 2021.
- [3] H. Rahman, N. B. Aoun, T. F. N. Bukht, S. Ahmad, R. Tadeusiewicz, P. Pławiak, and M. Hammad, "Automatic liver tumor segmentation of ct and mri volumes using ensemble resunet-inceptionv4 model," *Information Sciences*, p. 121966, 2025.
- [4] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 7, pp. 3523–3542, 2021.
- [5] K.-N. Wang, S.-X. Li, Z. Bu, F.-X. Zhao, G.-Q. Zhou, S.-J. Zhou, and Y. Chen, "Sbcnet: Scale and boundary context attention dual-branch network for liver tumor segmentation," *IEEE Journal of Biomedical and Health Informatics*, 2024.
- [6] R. Azad, E. K. Aghdam, A. Rauland, Y. Jia, A. H. Avval, A. Bozorgpour, S. Karimijafarbigloo, J. P. Cohen, E. Adeli, and D. Merhof, "Medical image segmentation review: The success of u-net," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.
- [7] P.-H. Conze, G. Andrade-Miranda, V. K. Singh, V. Jaouen, and D. Visvikis, "Current and emerging trends in medical image segmentation with deep learning," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 7, no. 6, pp. 545–569, 2023.
- [8] L. Yang, J. Zhang, T. Wang, Q. Feng, S. Fu, and M. Huang, "Multi-scale camouflaged feature mining and fusion network for liver tumor segmentation," *Engineering Applications of Artificial Intelligence*, vol. 148, p. 110398, 2025.
- [9] H.-X. Liu, J.-X. Fang, Z.-J. Zhang, and Y.-C. Lin, "Localised edge-region-based active contour for medical image segmentation," *IET Image Processing*, vol. 15, no. 7, pp. 1567–1582, 2021.
- [10] S. Gul, M. S. Khan, A. Bibi, A. Khandakar, M. A. Ayari, and M. E. Chowdhury, "Deep learning techniques for liver and liver tumor segmentation: A review," *Computers in Biology and Medicine*, vol. 147, p. 105620, 2022.
- [11] H. Liu, J. Yang, C. Jiang, S. He, Y. Fu, S. Zhang, X. Hu, J. Fang, and W. Ji, "S2da-net: Spatial and spectral-learning double-branch aggregation network for liver tumor segmentation in ct images," *Computers in Biology and Medicine*, vol. 174, p. 108400, 2024.
- [12] D. Cheng, Z. Zhou, and J. Zhang, "Eg-unetr: An edge-guided liver tumor segmentation network based on cross-level interactive transformer," *Biomedical Signal Processing and Control*, vol. 97, p. 106739, 2024.
- [13] D. DJ and S. K. BS, "Liver tumor segmentation using g-unet and the impact of preprocessing and postprocessing methods," *Multimedia Tools and Applications*, vol. 83, no. 34, pp. 81383–81411, 2024.
- [14] G. Chen, H. Wang, Z. Lu, T.-H. Wu, K.-H. Lin, and G. S. Mok, "Convnext-2u: A 3d deep learning-based segmentation model for unified and automatic segmentation of lungs, normal liver and tumors in y-90 radioembolization dosimetry," *IEEE Transactions on Radiation and Plasma Medical Sciences*, 2024.
- [15] Y. You, Z. Bai, Y. Zhang, and Z. Li, "Contour-induced parallel graph reasoning for liver tumor segmentation," *Biomedical Signal Processing and Control*, vol. 92, p. 106111, 2024.
- [16] Z. Hu, H. Chen, L. Hua, X. Ren, and W. Mei, "Msml-attunet: A hierarchical attention network with multi-scale and multi-task for precision liver tumor segmentation," *Biomedical Signal Processing and Control*, vol. 99, p. 106861, 2025.
- [17] H. Kuang, X. Yang, H. Li, J. Wei, and L. Zhang, "Adaptive multiphase liver tumor segmentation with multiscale supervision," *IEEE Signal Processing Letters*, vol. 31, pp. 426–430, 2024.
- [18] J. Liu, X. Yi, and X. Huang, "Deepgcn: Toward deeper hyperbolic graph convolutional networks," *IEEE Transactions on Artificial Intelligence*, 2024.
- [19] K. Zhang, D. Zhu, X. Min, and G. Zhai, "Implicit neural representation learning for hyperspectral image super-resolution," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–12, 2022.
- [20] X. Shi, H. Qi, Y. Shen, G. Wu, and B. Yin, "A spatial-temporal attention approach for traffic prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 8, pp. 4909–4918, 2020.

- [21] M. Wan, Y.-C. Lin, M.-S. Chen, N.-F. Zeng, G.-C. Wu, and H.-J. Zhang, "Virus diffusion algorithm: a novel metaheuristic algorithm for segmenting the spun crack," *Measurement Science and Technology*, vol. 36, no. 4, p. 045601, 2025.
- [22] M. Wan, Y. Lin, S.-X. Li, G.-C. Wu, N.-F. Zeng, S. Zhang, M.-S. Chen, C. Li, X.-D. Zhan, and Y.-L. Qiu, "Mhaed-net: a lightweight multiscale hybrid attention encoder-decoder network for the efficient segmentation of industrial forging images," *The Journal of Supercomputing*, vol. 81, no. 8, p. 1001, 2025.
- [23] Y. Zhang, J. Yang, Y. Liu, J. Tian, S. Wang, C. Zhong, Z. Shi, Y. Zhang, and Z. He, "Decoupled pyramid correlation network for liver tumor segmentation from ct images," *Medical Physics*, vol. 49, no. 11, pp. 7207–7221, 2022.
- [24] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz *et al.*, "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.
- [25] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. R. Roth, and D. Xu, "Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images," in *International MICCAI brainlesion workshop*. Springer, 2021, pp. 272–284.
- [26] C. Hu, T. Xia, Y. Cui, Q. Zou, Y. Wang, W. Xiao, S. Ju, and X. Li, "Trustworthy multi-phase liver tumor segmentation via evidence-based uncertainty," *Engineering Applications of Artificial Intelligence*, vol. 133, p. 108289, 2024.
- [27] C. Huang, H. Han, Q. Yao, S. Zhu, and S. K. Zhou, "3d u 2-net: A 3d universal u-net for multi-domain medical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2019, pp. 291–299.
- [28] J. Yi, C. Chen, Q. Wei, D. Ding, and G. Yang, "Mmf-net: A novel multimodal multiscale fusion network for artery/vein segmentation in retinal fundus," in *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2022, pp. 1192–1197.
- [29] D. Jha, P. H. Smedsrød, M. A. Riegler, D. Johansen, T. De Lange, P. Halvorsen, and H. D. Johansen, "Resunet++: An advanced architecture for medical image segmentation," in *2019 IEEE international symposium on multimedia (ISM)*. IEEE, 2019, pp. 225–225.
- [30] C. Liu, F. Zhuo, and F. Wang, "Fault diagnosis of commutation failure using wavelet transform and wavelet neural network in hvdc transmission system," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–8, 2021.
- [31] A. Bdeir, K. Schwethelm, and N. Landwehr, "Fully hyperbolic convolutional neural networks for computer vision," *arXiv preprint arXiv:2303.15919*, 2023.
- [32] X. Wang, Y. Jin, and K. Hao, "Computational modeling of structural synaptic plasticity in echo state networks," *IEEE transactions on cybernetics*, vol. 52, no. 10, pp. 11 254–11 266, 2021.
- [33] C. Hansel and R. Yuste, "Neural ensembles: role of intrinsic excitability and its plasticity," *Frontiers in cellular neuroscience*, vol. 18, p. 1440588, 2024.
- [34] X. Tang, K. Yang, H. Wang, J. Wu, Y. Qin, W. Yu, and D. Cao, "Prediction-uncertainty-aware decision-making for autonomous vehicles," *IEEE Transactions on Intelligent Vehicles*, vol. 7, no. 4, pp. 849–862, 2022.
- [35] P. Bilic, P. Christ, H. B. Li, E. Vorontsov, A. Ben-Cohen, G. Kaassis, A. Szeskin, C. Jacobs, G. E. H. Mamani, G. Chartrand *et al.*, "The liver tumor segmentation benchmark (lits)," *Medical image analysis*, vol. 84, p. 102680, 2023.
- [36] Y. Ji, H. Bai, C. Ge, J. Yang, Y. Zhu, R. Zhang, Z. Li, L. Zhanng, W. Ma, X. Wan *et al.*, "Amos: A large-scale abdominal multi-organ benchmark for versatile medical image segmentation," *Advances in neural information processing systems*, vol. 35, pp. 36 722–36 732, 2022.
- [37] M. Salvi, U. R. Acharya, F. Molinari, and K. M. Meiburger, "The impact of pre-and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis," *Computers in Biology and Medicine*, vol. 128, p. 104129, 2021.
- [38] 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: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer, 2015, pp. 234–241.
- [39] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep learning in medical image analysis and multimodal learning for clinical decision support: 4th international workshop, DLMIA 2018, and 8th international workshop, ML-CDS 2018, held in conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, proceedings 4*. Springer, 2018, pp. 3–11.
- [40] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes," *IEEE transactions on medical imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.
- [41] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou, "Transunet: Transformers make strong encoders for medical image segmentation," *arXiv preprint arXiv:2102.04306*, 2021.
- [42] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnunet: a self-configuring method for deep learning-based biomedical image segmentation," *Nature methods*, vol. 18, no. 2, pp. 203–211, 2021.
- [43] O. Sikha, M. Riera-Marin, A. Galdran, J. G. Lopez, J. Rodriguez-Comas, G. Piella, and M. A. G. Ballester, "Uncertainty-aware segmentation quality prediction via deep learning bayesian modeling: Comprehensive evaluation and interpretation on skin cancer and liver segmentation," *Computerized Medical Imaging and Graphics*, vol. 123, p. 102547, 2025.