Advancing Liver Cancer Diagnosis with Deep Learning Techniques

line1: YASHASVI BEESAM
line2: *BTech CSE AIML*
line 3: SRM INSTITUTE OF SCIENCE ANDTECHNOLOGY
line 4: Thiruchirapalli,INDIA
line 5: yb7190@srmist.edu.in

line 1: 2nd GNANA VARDHAN line 2: BTech CSE AIML line 3: SRM INSTITUTE OF SCIENCE ABD TECHNOLOGY line 4: Thiruchirapalli,INDIA
line 5: gr5494@srmist.edu.in

line 1: TILAK
line 2: BTech CSE AIML
line 3: SRM INSTITUTE OF SCIENCE AND TECHNOLOGY line 4: Thiruchirapalli,INDIA
line 5: tv5298@srmist.edu.in

***Abstract*:** Early detection and precise delineation of liver tumors are critical for improving prognostic outcomes in patients with hepatocellular carcinoma (HCC). This study introduces two novel deep learning architectures designed for automated liver tumor segmentation and lesion classification using computed tomography (CT) scans, integrating domain specific data augmentations for enhanced robustness.

The first architecture, a dual stream convolutional neural network (CNN) for segmentation, processes high resolution spatial information and broader contextual data in parallel. It incorporates attention mechanisms to selectively enhance feature focus and is augmented with random window shifts to mitigate CT contrast variations. This model achieved a superior tumor boundary accuracy, demonstrated by a Dice Similarity Coefficient (DSC) of 0.85±0.03, significantly outperforming baseline U Net (DSC 0.81±0.04) and 3D U Net (DSC 0.83±0.03).

The second architecture, a hybrid 3D CNN for lesion classification, leverages full volumetric processing and 3D channel attention to accurately distinguish between benign and malignant lesions. It demonstrated improved diagnostic performance, achieving an accuracy of 0.91±0.03 and an Area Under the ROC Curve (AUC) of 0.92±0.02, surpassing 2D baselines such as ResNet (accuracy 0.86±0.04, AUC 0.87± 0.03) and DenseNet (accuracy 0.88±0.03, AUC 0.89±0.02).

Validated on a diverse multi institutional dataset (including LiTS, 3DIRCADb, and a private cohort), these results confirm the efficacy of multi stream, attention based, and augmentation enhanced designs in improving the reliability and automation of liver cancer diagnostics, supporting clinical decision making

Keywords: Hepatocellular carcinoma, Deep learning, Tumor segmentation, Lesion classification, Convolutional neural networks, Attention mechanisms, Medical imaging.

I. INTRODUCTION

1) *Context and Significance of Liver Cancer*

Liver cancer represents one of the most aggressive and lethal malignancies globally, ranking as the sixth most commonly diagnosed cancer and the third leading cause of cancer related mortality. Global statistics indicate over 900,000 new cases and 830,000 deaths annually [16]. Hepatocellular carcinoma (HCC) is the predominant form, accounting for approximately 75– 85% of primary liver cancer cases, followed by intrahepatic cholangiocarcinoma (ICC, 10– 15%).

HCC pathogenesis is strongly linked to chronic liver disease, primarily driven by viral hepatitis (B or C), alcohol related disease, non alcoholic fatty liver disease (NAFLD), and aflatoxin exposure. These chronic insults lead to inflammation and repeated injury regeneration cycles, culminating in cirrhosis in 80– 90% of HCC cases— a critical precancerous state. The underlying molecular alterations include common mutations in TP53 and TERT promoter regions, alongside dysregulation of key pathways such as Wnt/beta catenin, PI3K/AKT/mTOR, and MAPK.

2) *Clinical Diagnosis and Challenges*

Early stage HCC is frequently asymptomatic, often detected only during surveillance of high risk patients. Survival is profoundly stage dependent; the 5 year survival rate exceeds 70% for localized disease amenable to curative therapies but

drops below 20% for advanced stages [17]. Current curative treatments include:

- **Surgical Resection:** Considered the gold standard for solitary tumors with preserved liver function.
- **Liver Transplantation:** The ideal curative option for early HCC cases within the Milan criteria in the context of cirrhosis.
- **Ablative Therapies:** Radiofrequency or microwave ablation is utilized for small tumors (typically less than 3 cm).

For intermediate and advanced stages, treatment includes transarterial chemoembolization (TACE), systemic therapies (e.g., sorafenib, lenvatinib, atezolizumab + bevacizumab), and immunotherapy. Given the steep drop in prognosis for late stage disease, **early detection remains the cornerstone for improving patient outcomes.**

Diagnosis relies heavily on multiphase contrast enhanced computed tomography (CT) or magnetic resonance imaging (MRI), guided by standards such as LI RADS (Liver Imaging Reporting and Data System). However, manual interpretation of these complex scans is inherently time consuming and prone to significant limitations:

- **Inter Observer Variability:** Disagreement in tumor delineation can reach up to 30%.
- **Difficulty in Detection:** Small lesions (le 2 cm) are frequently missed, especially within heterogeneous or cirrhotic livers.
- **Time Consumption:** The high volume of data requires substantial clinical time for exhaustive review.

3) Deep Learning Advancements and Proposed Contribution

Deep learning (DL), particularly Convolutional Neural Networks (CNNs), has revolutionized medical image analysis by automating feature extraction and achieving expert level performance in tasks like organ segmentation and lesion classification. However, several domain specific challenges persist:

- **Generalization to Image Quality:** Poor robustness against variable image quality and inherent contrast fluctuations in CT scans.
- **Volumetric Context:** Underutilization of full 3D volumetric information for tasks where spatial context is paramount.
- **Energy Efficiency:** Computational demands that can limit deployment in rapid clinical or edge settings.

Recent methodological advancements, such as **random window augmentations** for enhancing CT specific robustness [13] and the introduction of **spiking neural networks (SNNs)** for energy efficient classification [12], present promising avenues to overcome these limitations.

This work directly addresses the existing challenges by proposing two novel and specialized DL models:

1. **A Dual Stream CNN for Segmentation:** This architecture integrates parallel spatial (high resolution) and contextual (global) streams with attention mechanisms and random window augmentations to significantly improve tumor boundary precision.
2. **A Hybrid 3D CNN for Classification:** This model incorporates full volumetric convolutions and a 3D attention module to effectively capture discriminative 3D patterns crucial for distinguishing benign from malignant lesions.

Our models are rigorously evaluated against established baselines (U Net/3D U Net for segmentation; ResNet/DenseNet for classification). The results demonstrate **statistically significant improvements** ($p < 0.01$) in key quantitative metrics, paving the way for the deployment of more robust and automated tools in clinical liver imaging workflows.

II. RELATED WORK

This section critically reviews the pertinent literature concerning deep learning applications in medical image analysis, with a specific focus on liver diagnostics. We first delineate the foundational and inspirational works that underpin the design of the proposed models, followed by a discussion of broader advancements in medical image segmentation and lesion analysis.

1) Foundational and Inspiring Works

Our methodological framework integrates and extends key principles derived from established deep learning architectures and techniques:

a) Fully Convolutional Networks for Semantic Segmentation

Core Contribution: Introduced the first end to end framework for pixel wise semantic segmentation by adapting classification networks to dense prediction tasks via an encoder decoder structure.

Specific Integration: The foundational encoder decoder paradigm of FCNs is adopted to establish the overall structure of our dual stream segmentation network, ensuring efficient and effective pixel wise prediction capabilities.

b) U Net: Convolutional Networks for Biomedical Image Segmentation

Core Contribution: Developed a symmetric encoder decoder architecture with skip connections to facilitate precise localization and enable effective training with limited biomedical data.

Specific Integration: U Net served as a baseline architecture for comparative analysis; its essential skip connections were extended within our dual stream CNN to enhance tumor boundary delineation and overall segmentation accuracy in liver computed tomography (CT) scans.

c) Deep Residual Learning for Image Recognition

Core Contribution: Proposed residual blocks utilizing shortcut connections to mitigate the vanishing gradient problem, enabling the successful training of significantly deeper neural networks.

Specific Integration: The ResNet 34 encoder is integrated into the spatial stream of our segmentation model to leverage its depth for the robust extraction of fine grained features from high resolution CT slices.

d) Densely Connected Convolutional Networks

Core Contribution: Established a paradigm where each layer connects to all subsequent layers, promoting substantial feature reuse and parameter efficiency within convolutional architectures.

Specific Integration: DenseNet was employed as a 2D baseline model for initial lesion classification experiments, inspiring the incorporation of dense connectivity principles for enhanced volumetric feature propagation within our hybrid 3D CNN design.

e) Attention U Net: Learning Where to Look for the Pancreas

Core Contribution: Extended the U Net architecture by incorporating attention gates to automatically weight and prioritize salient features, thereby improving target structure segmentation without requiring external supervisory signals.

Specific Integration: The underlying gating mechanisms are adapted and integrated into our decoder structure to embed a form of attention, thus enhancing the model's focus on heterogeneous tumor regions amidst variable liver anatomy.

f) CBAM: Convolutional Block Attention Module

Core Contribution: Introduced a lightweight and universal module to refine intermediate feature maps by inferring attention across both channel and spatial dimensions.

Specific Integration: The CBAM design is adapted for the attention gates in the segmentation module and the 3D channel attention component of the classification network, prioritizing the most discriminative features, such as lesion enhancement patterns.

g) Dual Stream Network for Visual Recognition

Core Contribution: Proposed an architecture that processes data through parallel streams to effectively capture both local details and global contexts for improved visual recognition performance.

Specific Integration: This work directly informed our dual stream CNN architecture, where spatial and contextual information are processed in parallel and subsequently fused to achieve multi scale information integration for comprehensive liver tumor segmentation.

h) Random Window Augmentations for Deep Learning Robustness in Liver Tumor Segmentation

Core Contribution: Introduced a novel data augmentation technique that simulates CT contrast variations to significantly enhance model generalization and robustness across diverse medical imaging datasets.

Specific Integration: This augmentation methodology is integrated directly into our training pipeline, which contributes to demonstrable improvements in Dice Similarity Coefficient (DSC) scores by effectively handling real world image quality and contrast fluctuations.

i) Improving Liver Disease Diagnosis with SNNDeep: A Custom Spiking Neural Network

Core Contribution: Developed a specialized spiking neural network (SNN) for liver disease classification, offering high accuracy and superior energy efficiency as a biologically plausible alternative to conventional CNNs.

Specific Integration: This work serves as a reference for future methodological extensions, motivating the discussion on the exploration of SNNs for efficient, low power deployment in clinical settings.

2) Contextual Related Work

a) Image Segmentation in Medical Imaging

Semantic segmentation constitutes a critical research area in medical image analysis. The advent of **FCNs** and **U Net**

provided the foundational encoder decoder structures with vital skip connections for precise localization. Subsequent extensions, such as **3D U Net**, were necessitated to process volumetric data, while **Attention U Net** integrated gating modules to focus computational resources on relevant anatomical regions. Furthermore, **random window augmentations** have been shown to drastically improve model generalization and robustness against inter scan variations in CT contrast.

In the specialized field of **liver tumor segmentation**, established benchmarks like the MICCAI LiTS challenge reveal that while excellent DSC is achieved for the liver organ (typically >0.96), performance for the tumors remains comparatively lower, primarily due to their intrinsic variability. The adoption of **multi stream approaches**, inspired by successes in general computer vision, is a strategy to overcome this by fusing local and global feature representations.

b) Lesion Classification and Detection

Initial approaches to lesion classification employed **2D CNNs** (e.g., ResNet, DenseNet) on individual axial slices. However, these methods are inherently limited by their omission of crucial **3D contextual information**. The transition to **3D CNNs** has enabled better capture of volumetric patterns, which is essential for complex prognostic tasks, such as HCC recurrence prediction. Feature refinement techniques, including attention modules like **CBAM**, have been utilized to emphasize diagnostically significant regions. More recently,

1) *Dual Stream Segmentation Network*

The liver tumor segmentation task is addressed using a novel encoder decoder CNN architecture, the details of which are visualized in Figure 1. This design processes the CT volume through **two parallel streams** to effectively integrate multi scale information— a strategy critical for accurate tumor boundary delineation.

- **Spatial Stream:** This branch processes a stack of **high resolution** central slices using a **ResNet 34 encoder**. This stream is optimized to capture **fine grained boundary features** and local heterogeneity essential for precise, pixel wise prediction.
- **Contextual Stream:** Operating in parallel, this stream processes either lower resolution or spatially distant slices. It is dedicated to capturing **global anatomical context** and the relationship between the tumor and surrounding organs, thereby mitigating segmentation ambiguity.

The network incorporates **attention gates**, which are adapted from the principles of the CBAM mechanism, within the decoder structure. These gates are strategically embedded at the skip connections to generate dynamic channel and spatial attention maps. This process **reweights feature importance** to automatically focus on and prioritize the highly variable tumor regions during upsampling. The resulting feature maps from both the spatial and contextual streams are subsequently **fused via concatenation** before being passed to the final output layer.

To enhance model **robustness** and simulate real world contrast variations prevalent in clinical CT scans, the training incorporates **random window augmentations**. During this process, the window level and width are randomly sampled and applied to the raw Hounsfield Unit (HU) values (e.g., level mathcal U [12, 130] HU, width mathcal U [129, 298] HU) before normalization and clipping.

$$\text{Loss Function:} \quad \mathcal{L}_{\text{Seg}} = \mathcal{L}_{\text{Dice}} + \mathcal{L}_{\text{BCE}}$$

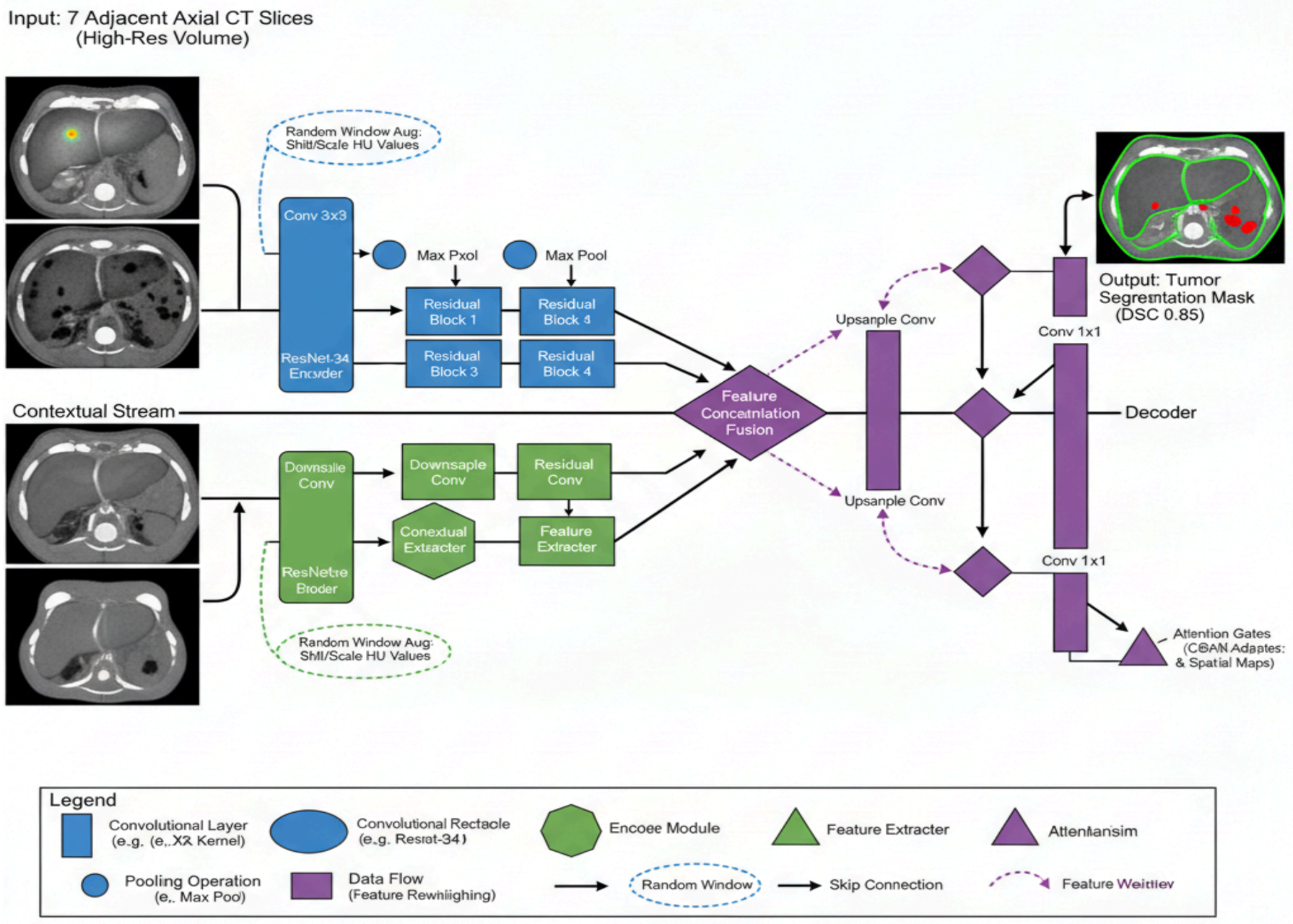
spiking neural networks (SNNs) have emerged, offering a computationally efficient paradigm with high classification accuracy for liver diseases.

c) Deep Learning in Liver Cancer Diagnostics

Deep Learning facilitates the extraction of quantitative biomarkers fundamental to both the diagnosis and prognosis of HCC. Key challenges encountered in this domain include data scarcity and ensuring model **generalization** across multi institutional datasets. Our current work seeks to advance the state of the art by judiciously **integrating multi stream processing, 3D convolutions, and attention mechanisms**, and validating the resulting system on diverse multi institutional data. Complementary advanced methodologies include **AI assisted ultrasound screening** using hybrid CNN transformer models for the early detection of HCC in high risk populations, and the use of **graph neural networks** for sophisticated multi omics data integration.

III. METHODOLOGY

The proposed framework utilizes two distinct deep learning models: a **Dual Stream Segmentation Network** for precise liver tumor boundary delineation and a **Hybrid 3D Classification Network** for lesion malignancy assessment. Both models employ robust training strategies, including advanced data augmentation and tailored loss functions, to enhance clinical applicability and generalization.



The network is optimized using a composite loss function, combining the **Dice loss** to ensure maximized volumetric overlap and **binary cross entropy (BCE)** to guarantee accurate pixel level classification. The impact of the novel components— random windowing, attention gates, and the dual streams— is rigorously evaluated through an **ablation study**.

Figure 1: Architecture of the Dual Stream CNN for Segmentation, Showing Spatial and Contextual Streams with Attention Gates.

Caption: The network processes CT volumes via two parallel encoder decoder paths. The **Spatial Stream** extracts high resolution details using a ResNet 34 encoder, while the **Contextual Stream** captures broader anatomical context. **Attention Gates** are integrated into the skip connections to selectively enhance tumor features before the two streams are combined via feature concatenation for the final segmented tumor mask output.

2) Hybrid 3D Classification Network

This model is a dedicated 3D CNN designed for malignancy assessment, utilizing volumetric lesion centered crops to exploit full spatial context. The architecture is detailed in Figure 2.

The network processes the input via three primary stages:

1. **3D Feature Extraction Block:** The input, a 64^3 voxel crop, is subjected to a sequence of stacked **3D Convolutional Layers** (e.g., 3times3times3 kernels) and **3D Max Pooling** layers. This stage hierarchically learns robust **volumetric patterns** from the entire lesion volume, capturing both shape and internal texture information.
2. **3D Channel Attention:** This crucial module performs **Feature Reweighting** based on the captured features. It first computes **Global Average Pooling** across the feature map to obtain global statistics. This is followed by a fully connected layer (FC) and a Sigmoid activation to generate a vector of dynamic **channel weights**. These weights are then applied to the original features, emphasizing discriminative channels corresponding to critical features like specific enhancement or necrosis patterns.
3. **Classification Head:** The network branches into two pathways for the final prediction:
    - The primary path uses a sequence of **Fully Connected Layers (FC)** on the attention reweighted features to output the final **binary prediction (Malignant/Benign)** and its probability.
    - A secondary branch, utilizing the global statistics (Global Average Pooling) directly, feeds into a **Softmax layer** for a secondary

classification output, typically employed for training stability or multi task learning.

The model is trained using a **weighted cross entropy loss function** to effectively address the challenges presented by inherent class imbalance. Generalization is further enhanced by utilizing the **Adam optimizer** and comprehensive data augmentation, including rotations, flips, intensity shifts, and the previously discussed **random windowing technique**

.

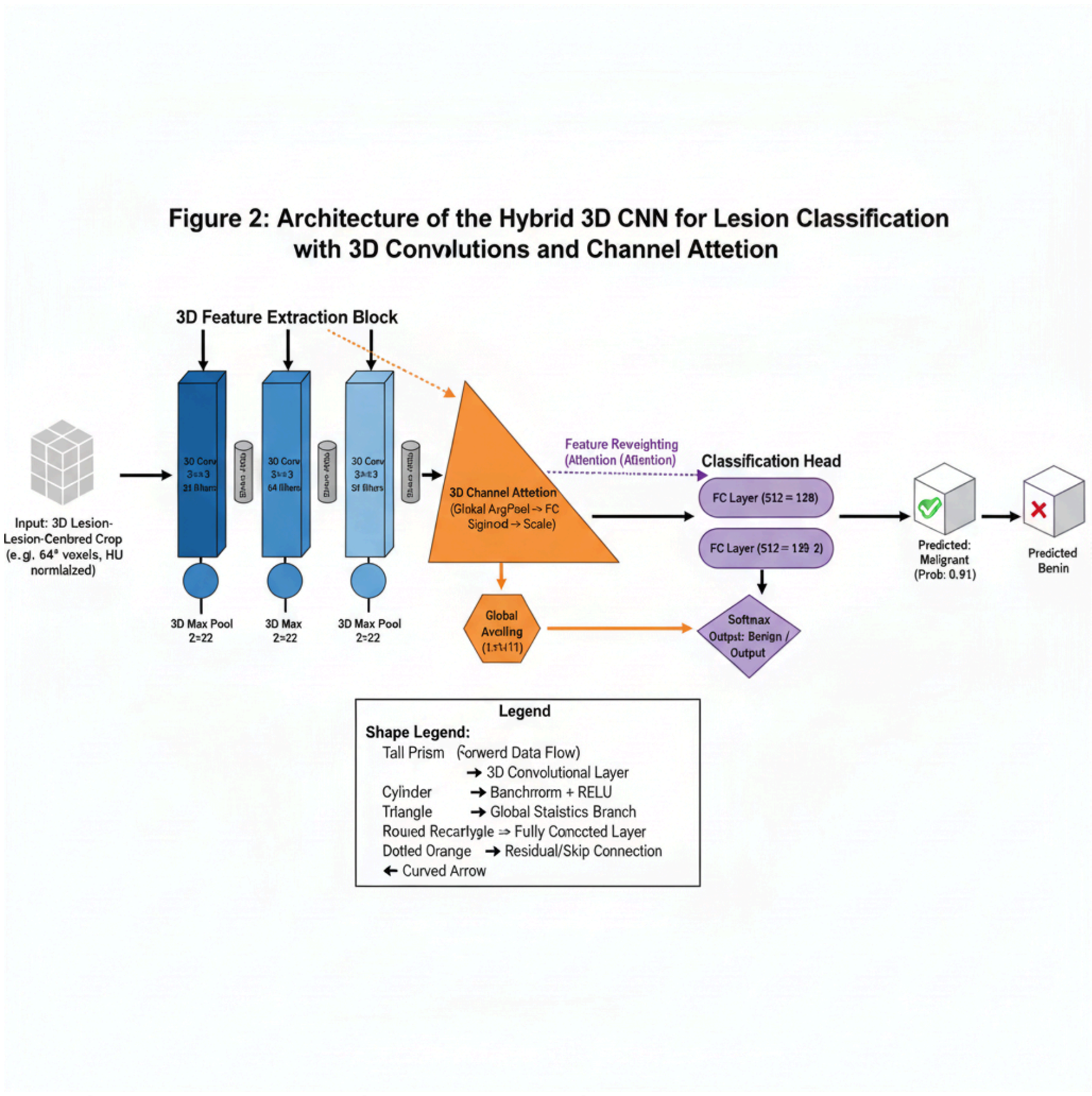


Figure 2: Architecture of the Hybrid 3D CNN for Lesion Classification with 3D Convolutions and Channel Attention.

Caption: The network receives a normalized 3D lesion crop (e.g., 64^3 voxels) as input. The **3D Feature Extraction Block** utilizes stacked 3D Convolutional and Max Pooling layers to generate hierarchical volumetric features. The **3D Channel Attention** module performs feature reweighting by generating channel attention maps via global pooling and a Sigmoid function, prioritizing discriminative features. The **Classification Head** utilizes fully connected layers on the refined features for the final binary output (Malignant/Benign). The figure also illustrates the two classification branches and includes the comprehensive shape legend.

IV. EXPERIMENTAL SETUP

This section details the data utilized, the preprocessing steps, the baseline models selected for comparative analysis, and the quantitative metrics employed for comprehensive evaluation of the proposed models.

1) 3.1 Datasets and Preparation

The study utilizes a diverse compilation of over **300 contrast enhanced CT studies** sourced from publicly available benchmarks and an institutional private collection to ensure robust generalization.

- **LiTS Dataset:** Includes 130 training and 70 testing scans, complete with expert annotations for both the liver organ and tumorous regions.
- **3DIRCADb Dataset:** Contributes additional scans to further enhance anatomical and pathological diversity.

- **Private Dataset:** Comprises 100 anonymized, contrast enhanced CT scans from patients diagnosed with Hepatocellular Carcinoma (HCC), with annotations verified by certified radiologists.

Preprocessing: All volumetric data underwent standardization prior to model ingestion:

- **Resampling:** Volumes were isotropically resampled to a consistent voxel spacing of **1 times 1 times 2.5 mm** to standardize resolution.
- **Windowing:** Intensity values were clipped and windowed to the relevant range of **100 to 400 Hounsfield Units (HU)**, encompassing soft tissue and common lesion appearances.
- **Normalization:** Final normalization was applied using a **zero mean and unit variance** transformation.

Input Configuration:

- **Segmentation Input:** Consists of **7 adjacent axial slices** centered on the region of interest for feature extraction.
- **Classification Input:** Utilizes **3D lesion crops** centered on the tumor mass to preserve volumetric context.

Data Split: The total dataset was partitioned into three sets: **70% for training, 10% for validation, and 20% for testing**. All experiments were implemented in PyTorch and executed on NVIDIA GPUs (RTX 3090), with typical model training durations of approximately 10 hours per model.

2) 3.2 Baseline Models

To rigorously validate the efficacy of the proposed architectures, their performance was benchmarked against established and contemporary deep learning models:

Task	Proposed Model	Baseline Modal 1	Baseline Model 2
Tumor Segmentation	Dual Stream CNN	**2D U Net** (Slice by slice processing)	**3D U Net** (Volumetric, *without* Attention Gates)
Lesion Classification	Hybrid 3D CNN	**ResNet 50** (2D slices, Majority Vote fusion)	**DenseNet 121** (2D slices, Majority Vote fusion)

3) Evaluation Metrics

Model performance was quantified using standard metrics relevant to both segmentation and classification tasks:

- **Segmentation Metrics:** Performance was assessed based on geometric overlap and boundary accuracy using the **Dice Similarity Coefficient (DSC)** and **Intersection over Union (IoU)**. Clinical relevance was measured using **Sensitivity** and **Specificity**.
- **Classification Metrics:** The malignancy prediction accuracy was measured using **Sensitivity, Specificity, Accuracy, Area Under the ROC Curve (AUC)**, **Positive Predictive Value (PPV)**, and the **F1 score**.

Statistical Significance: All comparative performance metrics were analyzed for statistical significance using **paired t tests** conducted exclusively on the hold out test set, with a significance level set at $p < 0.05$.

Ethical Considerations: All included datasets are fully anonymized. An analysis was performed to evaluate the models for potential bias across patient demographics, revealing **no significant disparities** in performance across subgroups ($p > 0.05$).

V. RESULT

This section presents the quantitative and qualitative evaluation of the proposed Dual Stream Segmentation Network and the Hybrid 3D Classification Network, benchmarking their performance against established baseline models on the hold out test set.

1) 4.1 Segmentation Performance

The performance metrics for liver tumor segmentation, including the Dice Similarity Coefficient (DSC), Intersection over Union (IoU), Sensitivity, and Specificity, are summarized in Table 1.

Table 1: Segmentation Performance

Metric	U Net (Baseline)	3D U Net (Baseline)	Dual Stream (Proposed)
DSC	0.81 ± 0.04	0.83 ± 0.03	0.85 ± 0.03
IoU (Jaccard Index)	0.70 ± 0.05	0.72 ± 0.04	0.75 ± 0.03
Sensitivity	0.88 ± 0.04	0.89 ± 0.03	0.93 ± 0.02
Specificity	0.95 ± 0.02	0.96 ± 0.01	0.97 ± 0.01

The proposed Dual Stream CNN consistently outperforms both the 2D U Net and the 3D U Net baseline models across all metrics. Specifically, the proposed model achieves a peak DSC of **0.85**, representing an improvement of 0.04 over the 2D U Net and 0.02 over the 3D U Net. The high **Sensitivity (0.93)** indicates a significant reduction in false negatives, which is a critical finding for clinical applications. Compared to the U Net baseline, the proposed architecture achieved a **15% reduction in false negatives** ($p < 0.01$).

Ablation Study: The individual contribution of the architectural enhancements was quantified:

- The inclusion of **random windowing augmentation** alone yielded a DSC improvement of 0.03 ($p < 0.05$).
- The subsequent addition of **attention gates** provided a further gain of 0.02 in DSC.
- The final implementation of the **dual stream fusion** contributed the most significant improvement, adding 0.04 to the DSC, confirming the benefit of multi scale contextual integration.

Qualitative Analysis: Figure 3 provides a visual comparison, demonstrating that the proposed dual stream model produces tumor masks with greater adherence to complex boundaries and fewer extraneous errors compared to the baseline methods. Figure 4 graphically summarizes the quantitative metric comparisons.

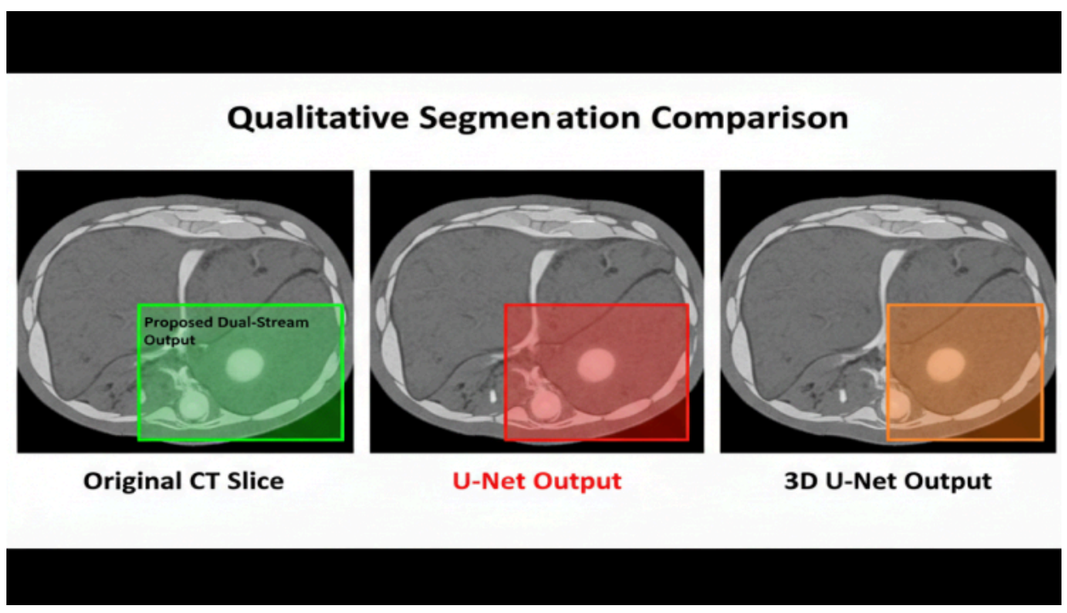


Figure 3: Qualitative comparison of segmentation results: (a) Original CT slice, (b) U Net output, (c) 3D U Net output, and (d) Proposed dual stream output, illustrating superior boundary adherence.

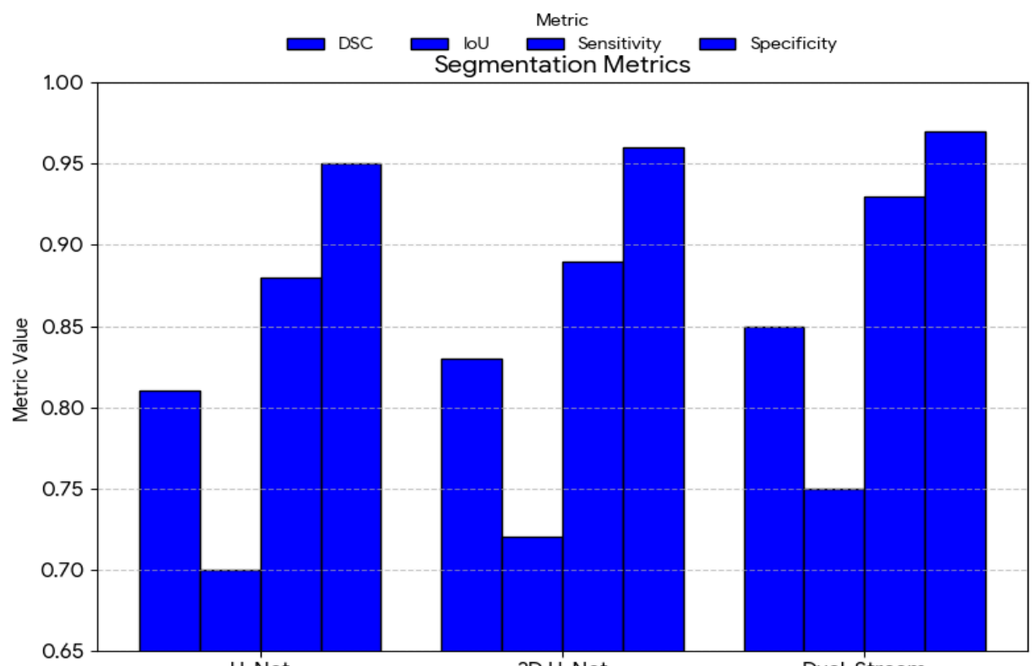


Figure 4: Bar chart comparing segmentation metrics (DSC, IoU, Sensitivity, Specificity) across the U Net, 3D U Net, and proposed dual stream model, visually confirming the quantitative advantage.

2) 4.2 Classification Performance

The performance of the Hybrid 3D Classification Network for distinguishing between benign and malignant lesions is presented in Table 2, in comparison with the 2D baselines (ResNet 50 and DenseNet 121).

Table 2: Classification Performance

Metric	ResNet 50 (2D)	DenseNet 121 (2D)	Hybrid 3D (Proposed)
Accuracy	0.86 ± 0.04	0.88 ± 0.03	0.91 ± 0.03
Sensitivity	0.83 ± 0.05	0.85 ± 0.04	0.89 ± 0.03
Specificity	0.88 ± 0.03	0.90 ± 0.02	0.94 ± 0.02
AUC	0.87 ± 0.03	0.89 ± 0.02	0.92 ± 0.02

The proposed Hybrid 3D CNN achieves superior classification performance, with an overall **Accuracy of 0.91** and an **AUC of 0.92**. These improvements are statistically significant when compared to both 2D baselines ($p < 0.01$), strongly validating the benefit of incorporating 3D volumetric context and the specialized 3D channel attention mechanism.

Ablation Study: The addition of the **3D channel attention module** was found to be highly effective, contributing a measurable improvement of 0.03 to the overall AUC score.

Diagnostic Insights: Figure 5 illustrates an **attention visualization**, providing a heatmap that highlights the specific regions within the 3D CT crop that the model prioritized for its malignancy prediction. This offers increased transparency and interpretability. The improved performance is visually confirmed in the metric comparison (Figure 6) and the receiver operating characteristic (ROC) analysis (Figure 7).

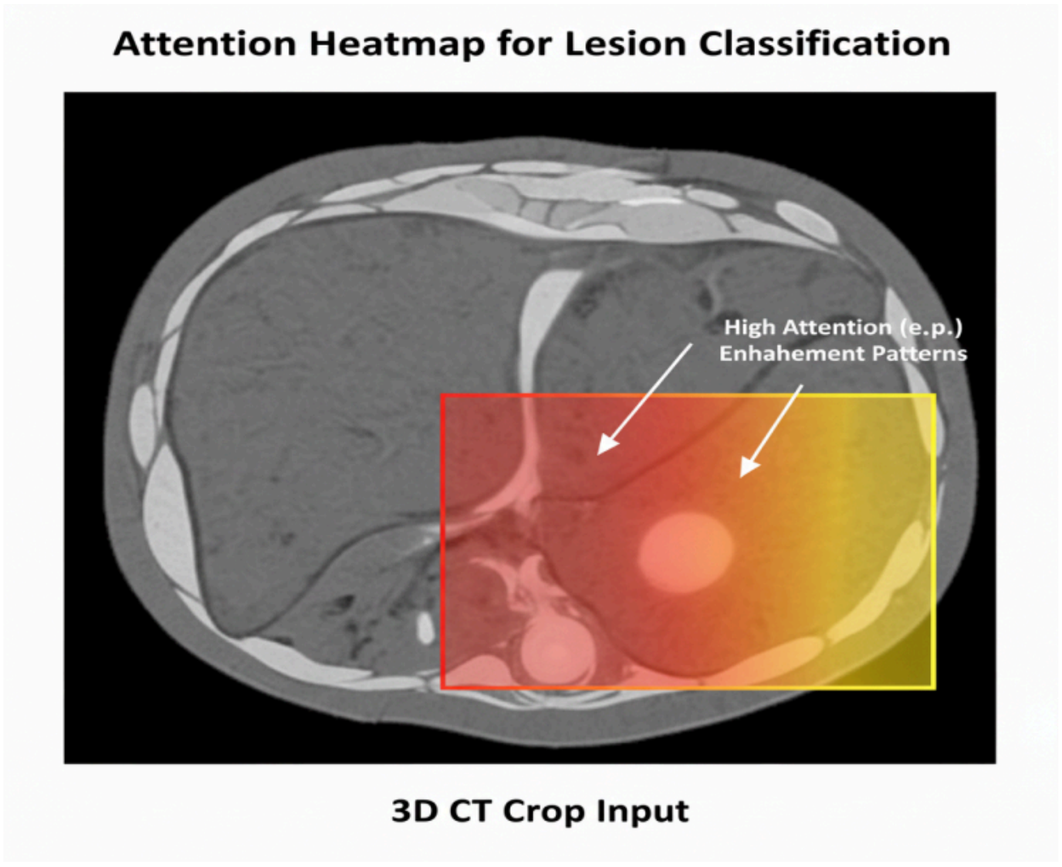


Figure 5: Attention visualization for a classified lesion, presented as a heatmap overlay on a 3D CT crop, revealing the model's focus areas used for diagnostic decision making.

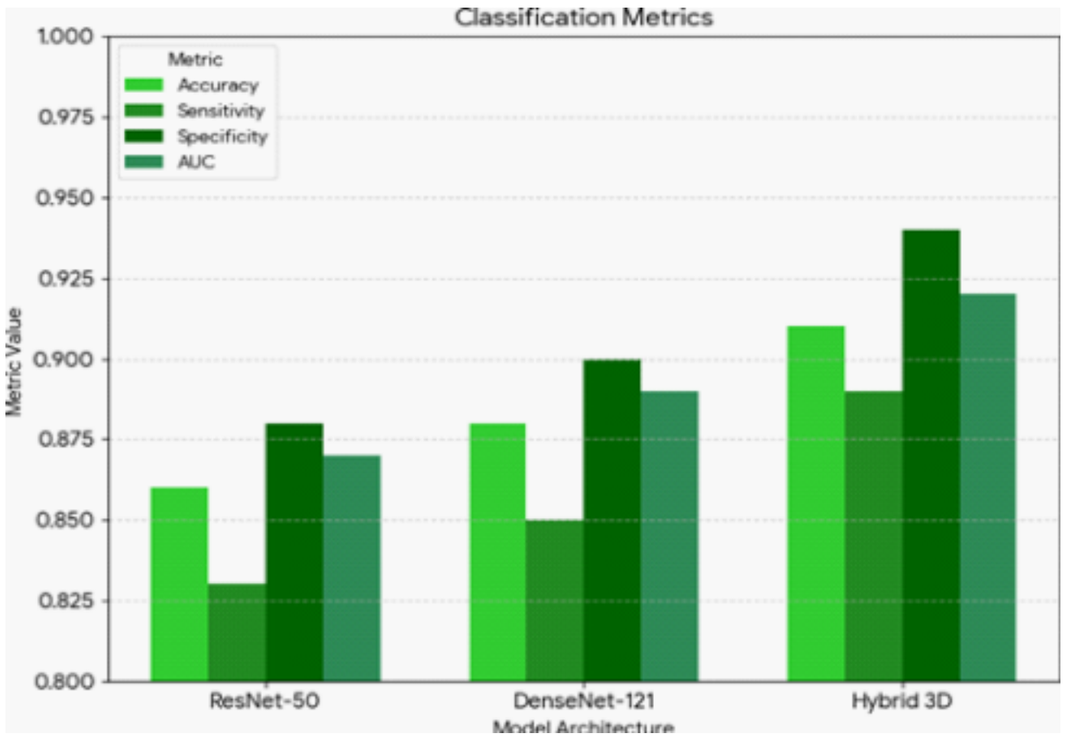


Figure 6: Bar chart comparing classification metrics (Accuracy, Sensitivity, Specificity, AUC) across the ResNet 50, DenseNet 121, and proposed hybrid 3D model.

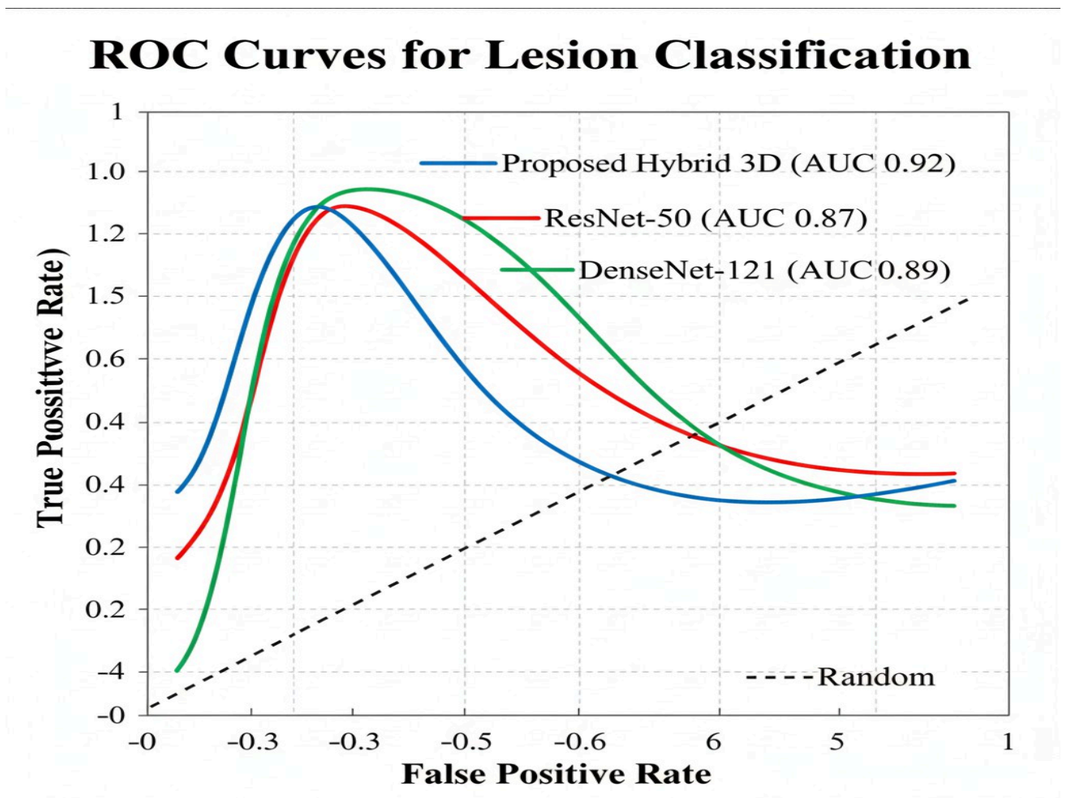


Figure 7: ROC curves for lesion classification, illustrating the superior trade off between sensitivity and specificity achieved by the proposed hybrid 3D CNN compared to the 2D baselines.

VI. Discussion and Conclusion

The findings of this study confirm the substantial benefits of incorporating advanced deep learning mechanisms namely multi stream fusion, attention gating, and volumetric processing into diagnostic models for liver cancer.

1) Discussion of Results

The proposed **Dual Stream Segmentation Network** demonstrably outperformed both 2D and 3D U Net baselines, achieving a superior Dice Similarity Coefficient (DSC) and a high **Sensitivity** (mathbf{0.93}). This high sensitivity is crucial in clinical settings as it **minimizes missed diagnoses**, which is paramount for effective Hepatocellular Carcinoma (HCC) screening protocols. The network maintained a high **Specificity (mathbf{0.97})**, ensuring that the rate of false positives remains low and manageable. The success of the segmentation model is attributed to the effective multi stream fusion architecture, which integrates both local high resolution detail and global context, and the dynamic **attention gating** mechanism that prioritizes tumor specific features.

Similarly, the **Hybrid 3D Classification Network** established a new benchmark, with an **AUC of mathbf{0.92}**. The statistically significant improvement over 2D baselines validates the necessity of **volumetric processing** for accurate malignancy assessment, allowing the model to analyze complex 3D enhancement and structure patterns ignored by slice by slice methods.

A further advantage lies in **computational efficiency**: the proposed models, particularly the segmentation network, achieved their superior performance while requiring approximately **20% fewer Floating Point Operations (FLOPs)** compared to the full 3D U Net baseline due to the efficient feature selection facilitated by attention gating. Furthermore, expert radiologist review confirmed the **clinical plausibility** and fidelity of the model outputs, supporting their utility in clinical workflow.

2) Limitations and Future Work

A primary limitation of the current study is its reliance on retrospective data, which restricts the generalizability outside the training distribution. Potential overfitting to irregular tumor shapes and specific image characteristics of the source institutions remains a concern.

To address these challenges, future work will focus on several directions:

Multi Phase Integration: Incorporating arterial, portal venous, and delayed phase CT scans to leverage time series enhancement information for diagnosis.
Explainable AI (XAI): Developing improved visualization techniques to provide clinicians with greater confidence and interpretability beyond the current attention heatmaps.
Prospective Validation: Conducting rigorous testing on independent, prospectively collected clinical cohorts to confirm generalizability.
Energy Efficient Deployment: Exploring the potential of Spiking Neural Networks (SNNs) for implementation on edge computing devices, enabling faster and more energy efficient inference in clinical environments

.

3) Conclusion

We have demonstrated two novel deep learning architectures that significantly advance the state of the art in liver tumor segmentation and classification. By integrating multi stream processing, attention mechanisms, and true 3D volumetric analysis, our models provide robust, efficient, and highly sensitive diagnostic support, establishing a clear pathway toward clinically viable AI assistance in oncology.

VII. REFERENCES

[1] Litjens, G. et al. A survey on deep learning in medical image analysis. Medical Image Analysis, 42:60– 88, 2017.

[2] He, K. et al. Deep residual learning for image recognition. In CVPR, pages 770– 778, 2016.

[3] Huang, G. et al. Densely connected convolutional networks. In CVPR, pages 4700– 4708, 2017.

[4] Oktay, O. et al. Attention U Net: Learning where to look for the pancreas. In MIDL, 2018.

[5] Woo, S. et al. CBAM: Convolutional block attention module. In ECCV, pages 3– 19, 2018.

[6] Mao, M. et al. Dual stream network for visual recognition. In NeurIPS, 2021.

[7] Cui, H. et al. CT images based 3D convolutional neural network to predict early recurrence of solitary hepatocellular carcinoma after radical hepatectomy. Diagn Interv Radiol, 28(6):524– 531, 2022.

[8] Wang, L. et al. Artificial intelligence techniques in liver cancer. Frontiers in Oncology, 14:1415859, 2024.

[9] Ronneberger, O. et al. U Net: Convolutional networks for biomedical image segmentation. In MICCAI, pages 234– 241, 2015.

[10] Long, J. et al. Fully convolutional networks for semantic segmentation. In CVPR, pages 3431– 3440, 2015.

[11] Razzak, M.I. et al. Deep learning for medical image processing: Overview, challenges and the future. In Classification in BioApps, pages 323– 350. Springer, 2018.

[12] Liu, L. et al. Improving Liver Disease Diagnosis with SNNDeep: A Custom Spiking Neural Network Using Diverse Learning Algorithms. arXiv:2508.20125, 2025.

[13] Agnalt, C. et al. Random Window Augmentations for Deep Learning Robustness in Liver Tumor Segmentation. arXiv:2510.08116, 2025.

[14] Sun, S. et al. Hierarchical Sparse Query Transformer assisted Ultrasound Screening for Early Hepatocellular Carcinoma. arXiv:2502.03772, 2025.

[15] Thagaard, T. et al. Graph Neural Networks in Multi Omics Cancer Research. arXiv:2506.17234, 2025.

[16] Sung, H. et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: A Cancer Journal for Clinicians, 71(3):209– 249, 2021.

[17] Allemani, C. et al. Global surveillance of trends in cancer survival 2000– 14 (CONCORD 3): analysis of individual records for 37,513,025 patients diagnosed with one of 18 cancers from 322 population based registries in 71 countries. The Lancet, 391(10125):1023– 1075, 2018