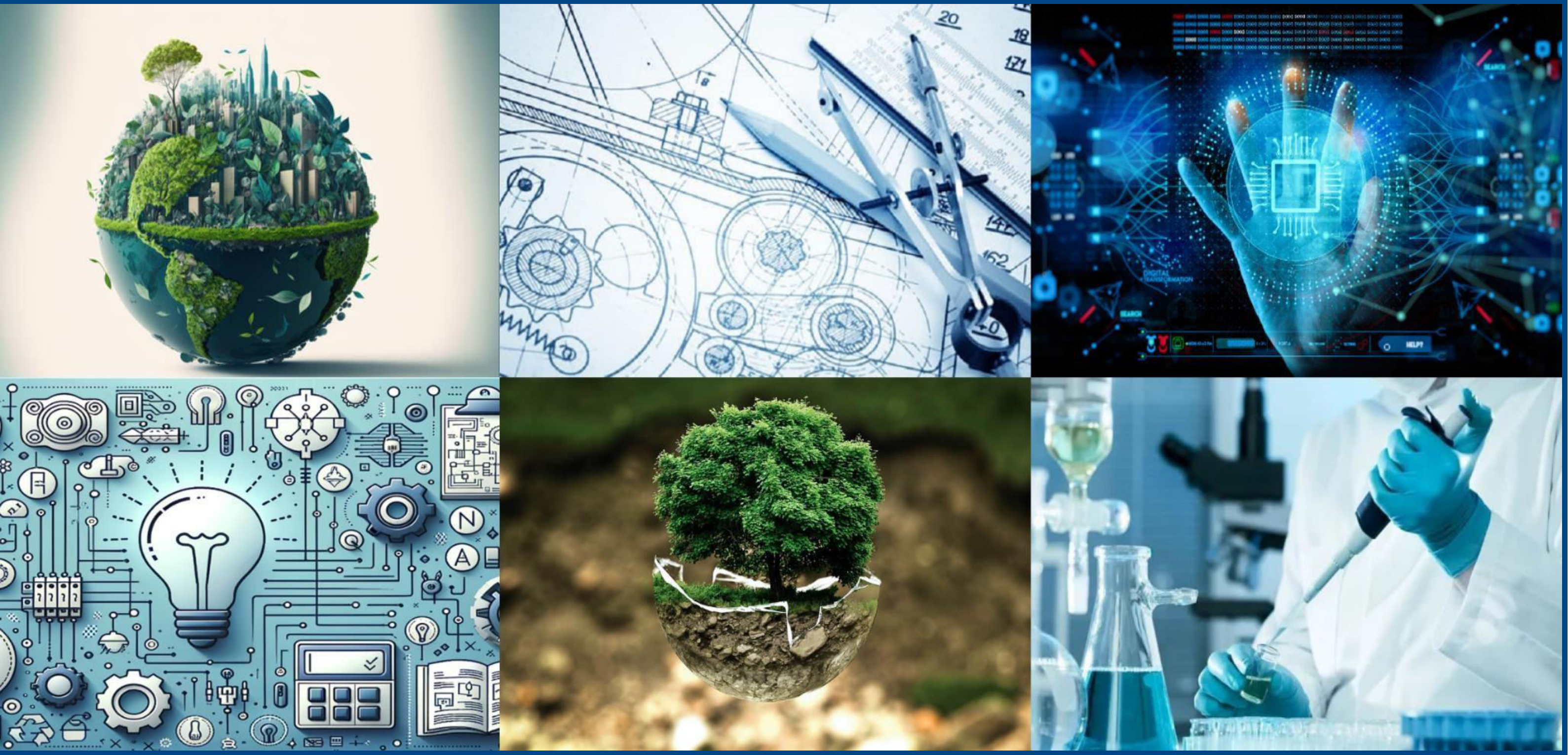
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 12, December 2025

© 2025 IJMRSET | Volume 8, Issue 12, December 2025| DOI:10.15680/IJMRSET.2025.0812014

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Deep Learning–Based Soft Tissue Tumor Detection using GM-UNet Architecture

G Lakshmi[1], Karthik D N[1], Amrutheshwara S[1], Adarsha K K[1], Lohithesh Kumar N.[2]

U.G. Students, Department of Computer Science and Engineering, PESITM, Shivamogga, Karnataka, India[1]

Assistant Professor, Department of Computer Science and Engineering, PESITM, Shivamogga, Karnataka, India[2]

ABSTRACT: Soft tissue tumor detection plays a crucial role in accurate diagnosis and effective treatment planning in clinical oncology. This paper introduces an advanced framework for soft tissue tumor detection using GM-UNet, a hybrid deep-learning model that integrates a U-Net backbone with Generative Adver- sarial Networks (GANs) and a novel Visual Mamba attention mechanism. The proposed method leverages the complementary strengths of these components to achieve precise tumor segmenta- tion and classification. The GAN component enhances the realism of generated segmentation masks, while the Visual Mamba at- tention module improves focus on critical tumor regions, thereby increasing both accuracy and interpretability. Comprehensive ex- periments on publicly available MRI collections demonstrate the efficacy of GM-UNet in accurately segmenting tumors, classifying them as benign or malignant. Comparative analysis confirms the superiority of the proposed model over conventional approaches in terms of accuracy and robustness. The study also highlights GM-UNet's potential as an AI-powered decision-support tool for radiologists and oncologists, capable of reducing diagnostic time and improving patient outcomes. Overall, the proposed approach represents a significant advancement in automated medical image analysis, providing a reliable and interpretable framework for clinical diagnosis and management.

Keywords: GM-UNet, U-Net, Generative Ad- versarial Network, Visual Mamba, Medical Image Segmentation, Tumor Classification

I. INTRODUCTION

Soft tissue tumors pose significant diagnostic challenges due to their heterogeneous nature and varied appearances on medical imaging modalities such as Magnetic Resonance Imaging (MRI) and Computed Tomography (CT)[1]. Accurate detection, segmentation, and classification of these tumors are critical for guiding therapeutic decisions, surgical planning, and monitoring disease progression. However, conventional manual interpretation of radiological scans is labor-intensive, time-consuming, and subject to intra- and inter-observer variability, creating a strong need for automated and objective diagnostic systems[1]. In recent years, deep learning—particularly Convolutional Neural Networks (CNNs)—has become a transformative tool in medical image analysis[3]. Among various architectures, U-Net has emerged as a benchmark model for biomedical segmentation, leveraging its symmetric encoder–decoder design and skip connections to achieve precise localization while maintaining contextual awareness. Despite these advantages, standard U-Net architectures often face difficulties when seg- menting tumors with complex boundaries, irregular shapes, or low-contrast regions, leading to incomplete or inaccurate delineations. To address these challenges, we propose GM-UNet (Gen- erative Mamba-UNet), a novel hybrid framework that en- hances the traditional U-Net with two key components: a Generative Adversarial Network (GAN) for improving the realism of segmentation masks and a Visual Mamba attention mechanism for emphasizing critical tumor regions. This dual integration strengthens both segmentation precision and model interpretability, key factors for clinical acceptance and trust. Furthermore the framework extents beyond the segmen- tation by supporting a multi-task pipeline that includes EfficientNetV2-BO model for the classification of tumors. The resulting system provides an end-to-end diagnostic workflow—spanning detection, segmentation, classification—thereby offering a comprehensive and reliable AI-powered decision-support tool for radiologists and oncologists.

II. LITERATURE REVIEW

The field of medical image analysis has witnessed significant innovation, with numerous studies focusing on improving tumor detection using U-Net-based architectures[3]. Initial work demonstrated that U-Net performs effectively in segmenting soft tissue tumors, achieving high accuracy with minimal data by leveraging its encoder-decoder structure. However, limitations in standard models were identified. To overcome these limitations, advanced attention mecha- nisms such as Deformable Large Kernel Attention (DLKA) were introduced. These mechanisms capture both global and local features more efficiently than traditional self-attention, improving segmentation accuracy while reducing computa- tional cost. Further research led to the development of hybrid mod- els. The GM-UNet concept was proposed as a GAN-based architecture that uses a generator to segment images and a discriminator to refine outputs, resulting in smoother and more accurate tumor boundaries . Similarly, lightweight models like MALUNet were designed for resource-constrained environ- ments by integrating multiple attention modules (spatial, chan- nel, coordinate) to balance accuracy and efficiency, making them suitable for real-time applications . Other variants, such as MDS-UNet, introduced multi-dimensional and multi-scale processing to handle complex images with low contrast and irregular shapes[4]. Collectively, the literature indicates a clear trend: enhancing U-Net architectures with attention mechanisms, adversarial training, and multi-scale feature extraction substantially im- proves segmentation performance, boundary precision, and clinical applicability. However, a gap remains in creating a single, integrated system that not only segments tumors with high precision but also classifies them and quantifies their spread in an interpretable manner for direct clinical use. Our project addresses this gap by developing a comprehensive diagnostic tool based on the GM-UNet model[1].

III. SYSTEM ARCHITECTURE

The proposed MRI-based Soft Tissue Tumor Detection framework integrates preprocessing, segmentation, feature extraction, and deep learning–based classification into a unified end-to-end pipeline. The complete system architecture, as illustrated in Fig. X, consists of two major modules: (i) Image Preprocessing and Segmentation, and (ii) Feature Extraction and Tumor Classification. Together, these modules automate the detection and categorization of soft-tissue tumors into *benign* or *malignant* classes.

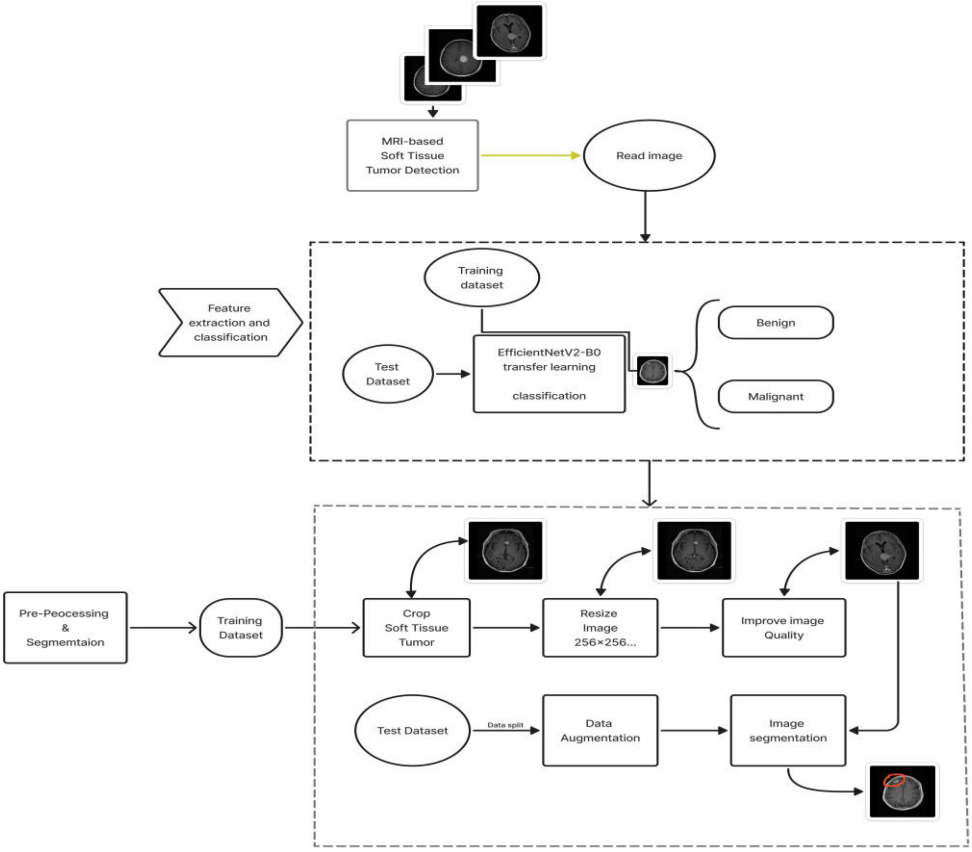


Figure 1: System Architecture

The system architecture begins by reading MRI images in a standardized format to ensure uniform handling across the pipeline. The preprocessing module enhances image quality, resizes each slice to 256×256 pixels, crops the tumor region, and applies segmentation to isolate the soft-tissue tumor, followed by data augmentation to improve generalization. After preprocessing, the dataset is split into training and testing sets for model development and evaluation. Segmented ROI images are then passed to an EfficientNetV2-B0 transfer learning model, which extracts deep discriminative features for tumor analysis. These features are classified into benign or malignant categories through fully connected layers that generate probabilistic predictions. Overall, the integrated workflow ensures seamless transition between segmentation and classification, enabling accurate, automated, and reliable tumor detection.

IV. METHODOLOGY

The proposed framework for detecting, classifying, and quantifying soft tissue tumors is based on the GM-UNet architecture. The overall workflow includes data collection, preprocessing, model development, multi-task training, and evaluation, culminating in a deployable system designed for clinical decision support[5].

A. Data Collection and Preprocessing

A diverse MRI dataset featuring soft tissue tumors was collected from publicly available medical imaging repositories. To enhance model compatibility and generalization, all images underwent a standardized preprocessing pipeline. This included resizing to a fixed dimension, normalizing intensity values, and applying data augmentation techniques such as rotation, flipping, and contrast variation to artificially expand the dataset and prevent overfitting.

B. Model Architecture: GM-UNet

The core of the proposed approach is the GM-UNet archi- tecture, which integrates a U-Net backbone, a Generative Ad- versarial Network (GAN) discriminator, and a Visual Mamba attention mechanism.

1) U-Net Backbone: The U-Net provides the encoder– decoder structure necessary for precise segmentation, utilizing skip connections to preserve spatial context and recover fine details during upsampling[1].

2) GAN Integration: A discriminator network is trained alongside the U-Net generator to differentiate between ground-truth and predicted segmentation masks. This adversarial learning strategy encourages the generator to produce anatomically realistic boundaries and sharper segmentation con- tours[2].

3) Visual Mamba Attention: To capture long-range depen- dencies and focus on critical features, the Visual Mamba attention module is embedded within the bottleneck of the U- Net. This mechanism enhances feature discrimination and aids in identifying tumors of varying shapes, sizes, and contrast levels within MRI scans[3].

C. Model Training and Evaluation

The dataset was divided into training, validation and testing sets.

The GM-UNet was trained end- to-end to perform three complementary tasks:

Tumor Segmentation: Optimized using the Dice loss function to maximize spatial overlap between predicted and ground-truth tumor regions[1].

Tumor Classification: The EfficientNet-B0 model was employed as the base architecture, with a custom classification to distinguish between benign and malignant tumors. The model was trained end-to-end using the categorical cross-entropy loss function to optimize clas- sification accuracy

Model performance was evaluated on the unseen test dataset using key metrics such as Dice coefficient, Intersection over Union (IoU), accuracy, sensitivity, and precision. These met- rics collectively assess the model’s segmentation precision, robustness, and classification reliability.

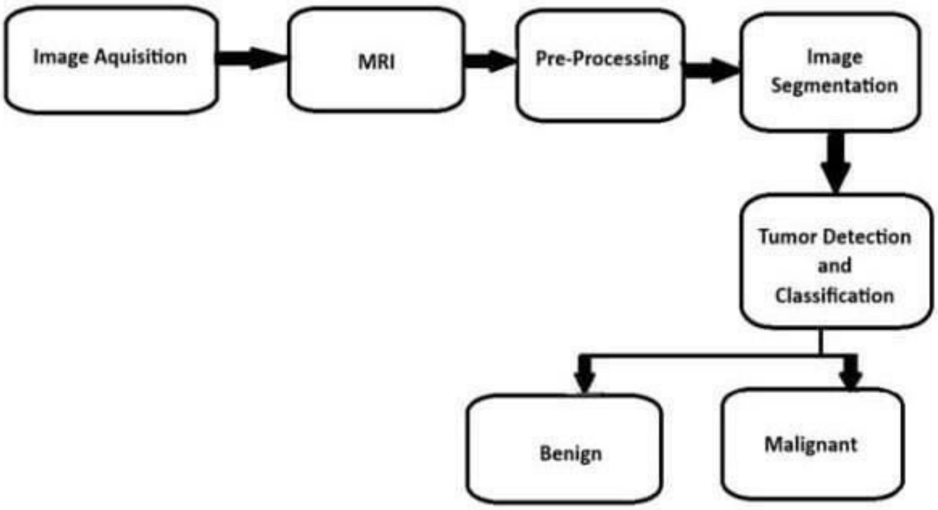


Figure 2. Soft Tissue Tumor Detection Workflow Diagram

Model Deployment

The trained GM-UNet model was deployed as a web application using a Python-based framework such as Flask or Django. The system provides an intuitive graphical interface that allows clinicians to upload MRI scans and receive near-instant diagnostic results. The output includes a visualized segmentation mask and final tumor classification, thus enabling a rapid and interpretable decision-support workflow.

The user-server interaction for the deployment phase is summarized in the sequence diagram shown in Fig. 1.

TABLE I: PYTHON LIBRARIES USED FOR THE EDA AND PREPROCESSING THE DATA

TABLE I: PYTHON LIBRARIES USED FOR THE EDA AND PREPROCESSING THE DATA **Sr**	**Libraries**	**Purpose**
1	OpenCV (cv2)	Image processing, resizing, reading MRI images and masks
2	NumPy	Numerical operations, tensor manipulation, preprocessing
3	scikit-learn	Confusion matrix, classification report,dataset utilities
4	TensorFlow / Keras	Deep learning framework for GM-UNet and EfficientNetV2
5	Matplotlib	Visualization of images, segmentation masks, training graphs
6	Seaborn	Heatmap plotting for confusion matrices and statistical visuals
7	Glob / OS	Directory handling, image–mask pairing,file traversal
8	TensorFlow tf.data API	High-performance data loading, batching,augmentation

V. RESULTS

The implementation of the proposed GM-UNet framework was carried out in Python using deep learning libraries such as TensorFlow and Keras. The MRI dataset was preprocessed with OpenCV and NumPy to ensure uniform image dimen- sions and pixel normalization[1].

Model Construction

The GM-UNet architecture was developed by modifying the standard U-Net model to include:

A patch-based discriminator to incorporate a Genera- tive Adversarial Network (GAN) component that enforces structural realism in segmentation outputs.

A Visual Mamba attention layer integrated within the encoder blocks to enhance long-range spatial awareness and fea- ture selection efficiency[3].

Training Configuration
The model was trained for 50 epochs using the Adam optimizer with a learning rate of $1 \times 10{-}4$. Batch normalization and dropout were applied to stabilize training and reduce overfitting. Performance on the validation set was monitored at each epoch, and early stopping was employed to preserve the best-performing weights. Upon training completion, the model was evaluated on the held-out test dataset.

TABLE II: Comparison of Model Performance with Existing Architectures

Model	Accuracy (%)	Dice Score
GM-UNet	94.2	0.91
Standard U-Net	88.7	0.85
Baseline CNN	86.1	0.82

Quantitative Results
The GM-UNet model demonstrated superior performance compared with baseline architectures. It achieved a Dice coefficient of 0.91 and a classification accuracy of 94.2%, indicating strong tumor localization and reliable diagnostic classification.

The initial steps for processing the input image within the system are conceptually modeled using methodology diagram, as shown in Fig 1.

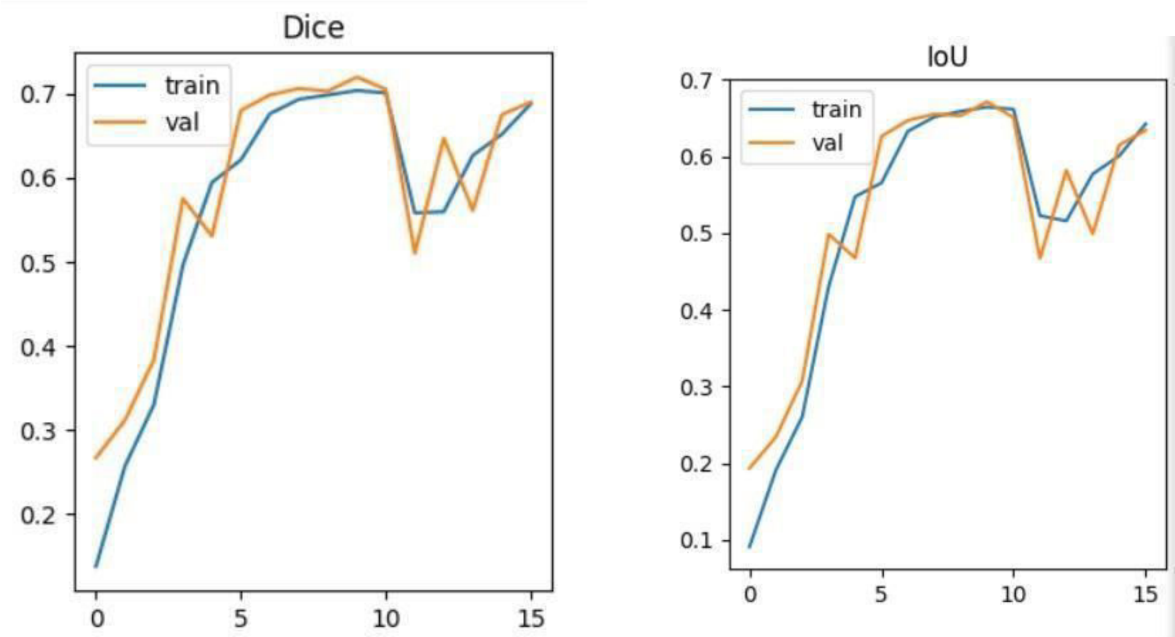


Figure 3: model performance with Dice Score and Iou

Qualitative Evaluation

Visual inspection of segmentation masks confirmed that GM-UNet produced smoother and more anatomically accurate boundaries compared to standard U-Net, particularly for tumors with low contrast and irregular margins[4].

In addition to visual segmentation quality, the classification performance of the system was evaluated using a confusion matrix, as shown in Fig. 4. The matrix highlights GM-UNet's strong ability to distinguish between benign, malignant, and no-tumor cases. The model achieved high true-positive rates across all three categories, with only a small number of mis- classifications occurring between visually similar malignant and benign tumor regions. These results further validate the robustness of the multi-task GM-UNet framework, demonstrating reliable tumor characterization alongside precise segmentation.

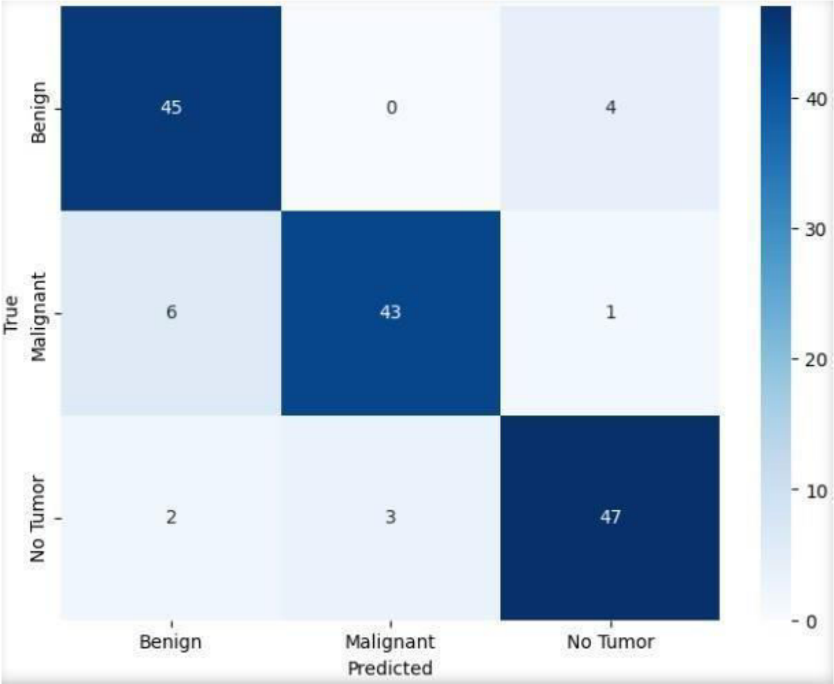


Figure 4: Confusion matrix illustrating the classification perfor- mance of the GM-UNet model across benign, malignant, and no-tumor classes.

VI. CONCLUSION AND FUTURE WORKS

This paper presented an advanced framework for the de- tection and classification of soft tissue tumors using the pro- posed GM-UNet architecture. The methodology, integrating systematic data collection, rigorous preprocessing, and multi- task model development, has proven effective for addressing this complex clinical challenge[1].

By leveraging the combined strengths of the U-Net back- bone, adversarial training with a GAN discriminator, and the Visual Mamba attention mechanism, the GM-UNet achieved accurate tumor segmentation and reliable classification perfor- mance.

Experimental results demonstrated that the GM-UNet model outperformed baseline architectures on key evaluation metrics such as Dice coefficient and overall accuracy. The system generates comprehensive diagnostic outputs, including seg- mentation maps and tumor classification making it suitable for seamless integration into clinical workflows[1]. This research contributes to the advancement of biomedical image analysis by offering a robust, accurate, and clinically applicable deep learning framework. The findings hold strong implications for enhancing diagnostic precision, supporting oncological treat- ment planning, and ultimately improving patient outcomes.

The GM-UNet framework presents vast potential for future research and clinical application. One promising direction is the integration of multi-modal imaging data—combining MRI, CT, and PET scans—to provide the model with a more comprehensive understanding of tumor morphology and to improve diagnostic precision.

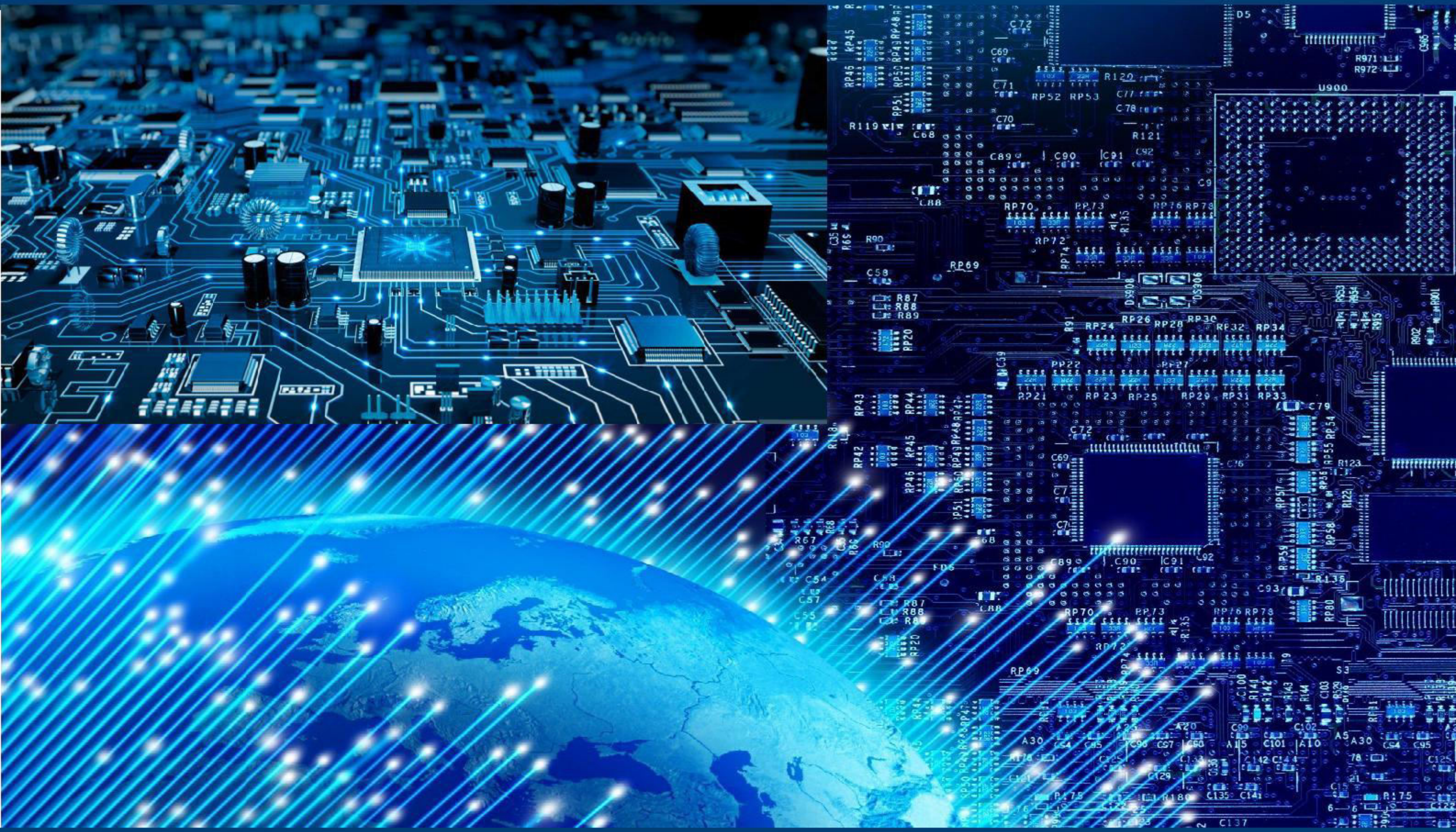
Architectural enhancements, such as refining the Visual Mamba attention mechanism or extending the framework into a 3D-CNN variant, could further enhance performance by capturing richer spatial context and volumetric features.

Another important avenue is the expansion of the dataset to include rare subtypes of soft tissue tumors, thereby improving the model's generalizability and robustness. Incorporating explainability tools such as attention heatmaps or Grad-CAM visualizations will be critical to foster clinical trust and trans- parency[6].

Finally, optimizing GM-UNet for deployment on low- resource or edge devices can democratize access to advanced diagnostic tools, enabling real-time inference in remote or under-resourced healthcare environments. Such efforts align with the broader goal of advancing intelligent, accessible, and data-driven medical diagnostics.

REFERENCES

[1] Kasliwal, J. Shende, A. Fuse, A. Borde, and V. Dhamdhere, "Soft Tissue Tumour Detection Using UNet," in Proc. 8th Int. Conf. on Computing, Communication, Control and Automation (ICCUBEA), 2024, pp. 1–5.
[2] J. Han, J. Cai, M. Zhu, Z. Gao, C. Li, and Y. Miao, "GM-UNet: GAN-based VMUNet achieves higher accuracy," in Proc. 11th Int. Conf. on Behavioural and Social Computing (BESC), 2024, pp. 1–6.
[3] R. Azad, M. Asadi-Aghbolaghi, M. Fathy, and S. Escalera, "Beyond self-attention: Deformable large kernel attention for medical image segmentation," in Proc. IEEE/CVF Winter Conf. on Applications of Computer Vision (WACV), 2024, pp.1238–1247.
[4] J. Ruan, Y. Ma, C. Song, M. Xu, L. Zhang, and J. Fan, "MALUNet: A multiattention and light-weight UNet for skin lesion segmentation," in Proc. IEEE Int. Conf. on Bioinformatics and Biomedicine (BIBM), 2022, pp. 2626–2631.
[5] L. Zhu and S. Yang, "MDS-UNet: Multi-Dimensional and Multi-Scale UNet for Medical Image Segmentation",2024 4th International Conference on Electronic Information Engineering and Computer Communication (EIECC), 2024, pp. 1–6.
[6] V. G. Janani et al., "AI-Enabled Early Detection of Brain Tumor Through AFFU-Net," 2024 International Conference on Smart Electronics and Communication Systems (ISENSE), 2024, pp. 1–6.
[7] C. Abdellaoui et al., "Medical Image Segmentation Based U-Net Architecture Implemented on Raspberry Pi4," 2024 3rd International Conference on Advanced Electrical Engineering (ICAEE), 2024, pp. 1–6.
[8] B. S. Sharvani, G. JayaLakshmi, and K. Bhargavi, "Automated Brain Tumor Diagnosis and Localization Using CNN and U-Net 3D: A Radiomics Approach for Enhanced Medical Imaging Analysis," 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), vol. 2, 2024, pp. 1–6.
[9] C. K. Mohan and S. Samudrala, "MRI Brain Tumor Detection and Classification Using U-NET CNN," 2023 International Conference on Data Science and Network Security (ICDSNS), Chennai, India, 2023, pp. 1–6.
[10] M. J. C. Samonte, A. B. Mallari, P. R. K. Reyes, and J. C. T. Tan, "Applying data science in computer vision: Detection of malignant and benign cancer tumors," in Proc. 4th Int. Conf. on Big Data Engineering and Education (BDEE), 2024, pp.13–20.



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com