ISSN Online: 2821-1693

Transactions on Machine Intelligence

Journal Homepage: https://tmachineintelligence.ir/index.php/TMI

Improving Medical Image Segmentation Using a Hybrid ResUNet-Transformer Architecture for Liver Tumor Detection

N. Shakeri[1,*], O. Rahmani Seryasat[2]

[1] Department of Biomedical Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran
[2] Assistant Professor, Department of Electrical Engineering, Shams Higher Education Institute, Gorgan, Iran

ARTICLE INFO

Article History:
Received 22 January 2024
Received in revised form 30 February 2025
Accepted 26 March 2025
Available online 29 March 2025

Keywords:
Medical Image Segmentation, Liver Tumors, Resunet Architecture, Transformer, Accuracy, Jaccard Coefficient (Iou), Dice Coefficient

ABSTRACT

In this research, a new method for segmentation of medical images is presented using a combination of ResUNet architecture and transformer layers. The main objective of this study is to improve the accuracy and efficiency of segmentation models in identifying liver tumors from medical images. In this method, the ResUNet50 architecture is used as the encoder for extracting deep features from images, and transformer layers have been added to the model to enhance the model's ability to understand spatial and channel relationships between features. Then, a decoder section with U-Net structure has been designed to reconstruct the predicted maps. To evaluate the proposed method, a dataset of medical images related to liver tumors was used. The experimental results show that the proposed method performs better compared to baseline models according to metrics such as accuracy, Jaccard coefficient (IoU), and Dice coefficient, and has achieved an average accuracy of 92.5%, Jaccard coefficient of 75.4%, and Dice coefficient of 83.1%. These results indicate that the combination of ResUNet and transformer architecture can provide an effective and powerful tool for segmentation of medical images and more accurate identification of liver tumors. In the future, using more diverse data and applying further optimization techniques can improve the efficiency of this model.

1. INTRODUCTION

Medical images play a vital role in diagnosing, monitoring, and treating many diseases. With advancements in imaging technologies, the volume of medical data has increased dramatically, highlighting the need for intelligent tools to process and analyze this data. Medical image segmentation, which involves precise division of different parts of an image (such as organs or lesions), is considered one of the important tasks in this field. This process helps with identifying, measuring, and analyzing structures and abnormalities related to diseases.

One of the main challenges in medical image segmentation is the variety in anatomical structures, low image quality, and noise resulting from the imaging process. For this reason, designing models with high accuracy and

[*] Corresponding Author: Navid Shakeri@gmail.com
Department of Biomedical Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran

http://dx.doi.org/10.47176/TMI.2025.57

efficiency to solve these issues is essential. In recent years, deep learning models, especially Convolutional Neural Networks (CNNs), have been very successful in segmentation problems due to their ability to extract complex features from images. U-Net is one of the prominent architectures in this field, which with its special design enables the extraction of spatial features and reconstruction of segmentation maps. However, U-Net and other similar architectures have limitations in cases where understanding long-range relationships and global dependencies in images is necessary. These limitations can lead to reduced accuracy in segmenting small or complex sections. To address these challenges, researchers have turned to combining convolutional architectures with attention-based techniques.

Transformers, which were initially introduced in natural language processing, have shown their ability to capture global relationships between elements, demonstrating high potential in computer vision tasks as well. Combining these methods with convolutional models, such as ResUNet, provides the possibility of simultaneous learning of spatial and channel features in images. ResUNet, which uses ResNet as its encoder section, provides a suitable foundation for extracting deeper and more accurate features to combine with transformer layers.

This research aims to design a hybrid model for medical image segmentation by combining ResUNet and transformer layers. The main objective of this model is to improve accuracy and performance in identifying complex abnormalities, including liver tumors. Liver tumors are one of the important challenges in healthcare, and their early and accurate diagnosis can have a significant impact on improving treatment processes and reducing mortality rates.

The proposed method in this research uses ResUNet50 for extracting initial features from images, and then by adding transformer layers, it models long-range and global dependencies between features. In the next step, using the U-Net structure, the predicted maps are reconstructed to their original dimensions.

This article begins by introducing the proposed method and presenting the results of experiments conducted on datasets related to liver tumors. At the end, the results obtained are compared with previous research, and the advantages and limitations of the proposed model are examined. This research aims to provide an efficient and accurate model that can be effective in improving existing methods for medical image segmentation.

2. RESEARCH BACKGROUND

In this section, Jha and colleagues (2019) introduced an advanced architecture called ResUNet++ for medical image segmentation. This architecture was designed with the goal of improving accuracy and generalization capabilities in identifying complex structures in images. ResUNet++ provides significant performance by utilizing residual blocks, spatial attention, and convolutional layers. This model uses skip connections to preserve spatial information and allow more precise detection of borders. Experimental results on standard datasets show that ResUNet++ has higher accuracy compared to basic models such as U-Net and ResUNet, and performs better in identifying small structures and unclear borders [1].

Wang and colleagues (2021) introduced a model called Hybrid Dilation and Attention Residual U-Net that combines residual blocks, attention networks, and dilated convolutions with spatial expansion. The aim of this research is to increase accuracy in identifying complex regions and small features in medical images. The proposed model strengthens the extraction of multi-scale features by using attention techniques and improving long-range dependencies between features through the use of hybrid convolutions. Experimental results demonstrated that this model has performed better than conventional architectures in the accurate detection of liver tumors and other complex structures [2].

In this section, Salpea and colleagues (2022) examined and analyzed modern architectures for medical image segmentation. This article includes a comprehensive review of various models including U-Net, ResUNet, and transformer-based hybrid models. The authors explained the advantages and disadvantages of each model and analyzed their applications in identifying different structures in medical images [3].

Huang and colleagues (2024) compared three important architectures for medical image segmentation in this study: U-Net, Res-UNet, and nnUNet. They analyzed the strengths and weaknesses of each model in identifying various structures in medical images. The results showed that nnUNet, due to its ability to self-optimize and use preprocessed data, has performed better in many cases compared to basic models such as Res-UNet [4].

Pan and colleagues (2023) introduced a model called WA-ResUNet for segmentation of MRI images with focus on identifying rare classes. This model uses a weighted attention mechanism to increase accuracy in identifying regions with low frequency. The WA-ResUNet model was able to demonstrate better performance compared to basic models in accurate detection of small structures and rare classes [5].

Wu and colleagues (2024) examined the ResUNet model for medical image segmentation in this article and evaluated its application in identifying damaged regions in medical images. The authors showed that using residual blocks in ResUNet improves the learning of deeper features and increases accuracy in identifying complex boundaries. This research also examined ResUNet's weak points, such as high memory consumption and longer training time, and presented solutions for performance improvement [6].

Baldeon-Calisto, M., & Lai-Yuen, S. (2018) introduced the ResU-Net model for segmentation of MRI images of the prostate. This model, using residual blocks and improved U-Net structure, enhanced the ability to extract complex features from images and increased accurate identification of regions related to the prostate. Results showed that ResU-Net had higher accuracy in detecting smaller structures and unclear borders compared to U-Net [7].

Elghazy, H. L., & Fakhr, M. W. (2022) examined Dual-Stream and Triple-Stream ResUNet/UNet architectures for multi-phase liver image segmentation. These models combine multiple information streams from images, giving them the ability to learn simultaneously from multi-phase data. This approach showed that accuracy in identifying tumor regions and liver-related structures could be improved. Experiments demonstrated that these architectures, compared to basic ResUNet models, provide better performance [8].

Huang and colleagues (2021) used a Two-Stage ResUNet model with Attention Gates for segmentation of frontal mediastinal lesions. This model first identifies the initial regions and then enhances prediction accuracy by applying attention mechanisms. Results showed that this model delivers exceptional performance in identifying smaller lesions and more concentrated areas, and can be very useful in clinical applications [9].

Weng, W., & Zhu, X. (2021) introduced a new model called INet that uses advanced convolutional networks for biological medical image segmentation. This model, combining deep convolutional layers and attention layers, enhances the ability to capture complex features in medical images and has shown significant improvement in performance compared to conventional models such as U-Net and ResUNet. INet has demonstrated particularly high applicability in sensitive applications such as the detection of rare diseases with complex borders [10].

Zhang and colleagues (2022) used a combination of two three-dimensional networks named "3D-DMFNet" and "3D-ResUNet" for breast cancer image segmentation. This architecture first identifies the main regions using "3D-DMFNet" and then improves them using "3D-ResUNet". This method was able to deliver outstanding performance in accuracy and processing speed, showing that using three-dimensional networks for volumetric medical data brings better results [11].

Chincholkar and colleagues (2024) compared the performance of various models including U-Net, Attention U-Net, ResNet50, and ResUNet in liver image segmentation. This study showed that ResUNet, with its combination of deep features and use of residual blocks, provides better performance compared to other models. Additionally, using Attention U-Net in some cases has helped improve the identification of smaller regions. By presenting quantitative and qualitative results, this article facilitates the selection of appropriate models for different problems [12].

Niu and colleagues (2022) introduced a model called P-ResUNet for brain tissue segmentation. This model, with improved residual blocks and using advanced mechanisms for feature extraction, was able to provide higher accuracy in diagnosing and segmenting complex brain regions. This model demonstrated extremely precise performance in identifying tumors and brain injuries, and showed significant advantages compared to basic models such as ResUNet and U-Net [13].

Appati and colleagues (2024) used a combination of Multiattention ResUNet architecture and a modified U-Net for liver tumor segmentation. The proposed model processes spatial and contextual information of images simultaneously by using multiple attention mechanisms (Multiattention). This combination resulted in higher accuracy in identifying complex boundaries and tumor regions, showing a significant improvement compared to standard ResUNet in evaluation metrics [14].

Wodzinski, M., & Müller, H. (2023) used a 3D ResUNet model for automatic aorta segmentation. The proposed model improved the accuracy of identifying aortic regions by using high-resolution data and enhancement operations. This model participated in the SEG.A (Segmentation of the Aorta) challenge and ranked among the best methods. The results showed that combining ResUNet architecture with rich and high-quality data can improve model performance in complex problems [15].

Sheng and colleagues (2021) used the Second-Order ResU-Net model for segmentation of brain tumors in MRI images. This model improved the learning of deep features and preservation of spatial information by introducing second-order mechanisms to ResUNet. Experimental results showed that this model, compared to ResUNet and conventional U-Net, had higher accuracy in identifying smaller regions and unclear tumor borders [16].

Zang and colleagues (2023) presented a new method for liver cancer segmentation by combining PCNN (Pulse Coupled Neural Network) with SE-ResUNet architecture in their image processing approach. This method uses spatial-channel features with Squeeze-and-Excitation mechanisms to increase accuracy in identifying small and complex regions. The results of this research demonstrated that this method has higher accuracy and efficiency compared to traditional architectures and can be widely used in similar problems [17].

Sun and colleagues (2024) used a combination of generative models (GAN) and 3D ResU-Net for generating images and segmentation of brain and retinal lesions. The GAN model was used to generate artificial data, and 3D ResU-Net was responsible for precise segmentation of lesions. This method was able to increase model accuracy in problems with limited data and provide appropriate performance in identifying complex regions [18].

Han and colleagues (2022) introduced a new model called ConvUNeXt that combines modern convolutional architectures and U-Net structure. This model is optimized in terms of resource consumption and has high capability in medical image segmentation with high accuracy. ConvUNeXt demonstrated appropriate performance, especially in problems with limited and restricted data, and was able to improve both accuracy and processing speed simultaneously [19].

Chincholkar and colleagues (2024) examined and evaluated advanced deep learning models including U-Net, Attention U-Net, ResNet50, and ResUNet for liver image segmentation in this article. Results showed that ResUNet, using residual block architecture and the ability to learn complex features, provides greater accuracy compared to other models. Additionally, Attention U-Net performed well in identifying smaller regions and discontinuous borders. The article provided a valuable guide for selecting optimal models in similar problems by presenting a comprehensive comparison [20].

3. RESEARCH METHODOLOGY

For conducting this research, a specialized dataset related to CT scan liver images was used, which includes medical images and manual masks labeled by radiologists. This dataset consists of high-resolution images that were specifically designed for identification and segmentation of liver tumors. Each input image in this collection contains various structures such as the liver, blood vessels, and tumors, which makes precise identification of these regions challenging. The masks associated with these images have been manually prepared by specialist physicians and used as ground truth labels for training and evaluating the models. Due to the structural and textural diversity in the images, this dataset serves as a robust benchmark for testing segmentation models in medical imaging.

To reduce computational complexity and improve compatibility with the proposed models, the images and masks were resized to 64×64 dimensions. Additionally, the pixel values were normalized within the range [0,1] to improve the model's learning process. Furthermore, data augmentation techniques including rotation, flipping, and image translation were used to increase data diversity and make the model more resistant to various changes. Having precise and standardized data enabled a thorough evaluation of the proposed methods in identifying and segmenting liver tumors, making the research results more reliable.

In this research, a hybrid model that combines ResUNet architecture with transformer layers was designed and implemented for medical image segmentation. The research methodology is divided into three main sections: data preparation, designing the proposed model, and evaluating the model's performance.

In the first section, data preparation, a specialized dataset related to liver tumor medical images was used. This dataset includes CT scan images with high resolution along with manual masks corresponding to tumor regions. Initially, the images were preprocessed using techniques to reduce them to smaller 64×64 dimensions to decrease memory consumption and computational complexity. Then, the masks were also scaled to similar dimensions and converted to binary values (0 and 1). To improve the model's generalization, data augmentation techniques such as rotation, translation, and image flipping were employed. These techniques help create diversity in the input data and prevent the model from overfitting.

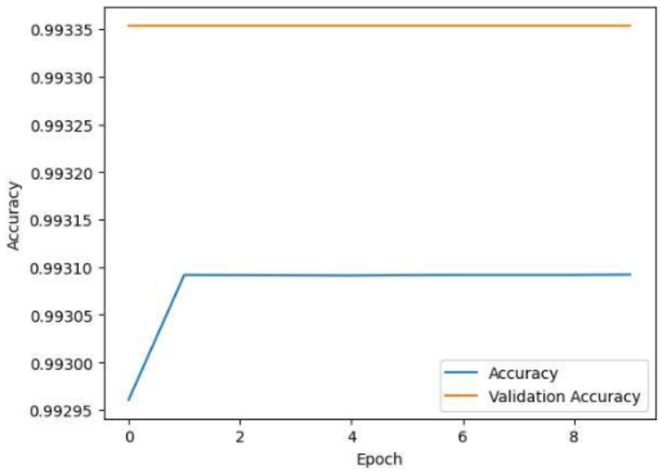


Fig. 1. Accuracy Calculation

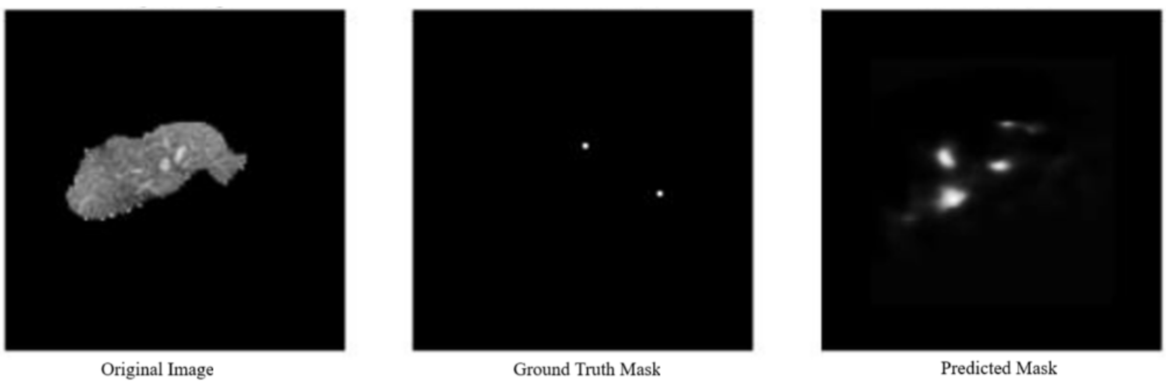


Fig. 2. Image Segmentation

In the second section of the methodology, the researchers describe the design of their proposed hybrid architecture that combines ResUNet with transformer layers.

The proposed model features a hybrid architecture incorporating both ResUNet and transformer components. In this design, ResUNet50 serves as the encoder (feature extraction) portion of the model. ResUNet50 leverages residual blocks, which allow it to extract deep features while preserving critical information in the images.

These extracted features are then passed as input to the transformer section. Within the transformer component, the Multi-Head Attention structure helps capture long-range and global dependencies between different features. Transformer layers are specifically designed to overcome the limitations of convolutional networks in learning complex and non-local relationships.

The output from the transformer is then fed into the decoder section, which follows a U-Net structure. This decoder includes convolutional layers and upsampling operations to gradually restore the image to its original dimensions. The decoder is responsible for reconstructing the segmentation maps.

Additionally, the model incorporates skip connections (a key feature of the original U-Net architecture) that connect corresponding layers between the encoder and decoder. These skip connections help preserve spatial details that might otherwise be lost during processing, enabling the model to better identify small and complex regions. This preservation of fine details is particularly crucial for accurately segmenting small tumor regions.

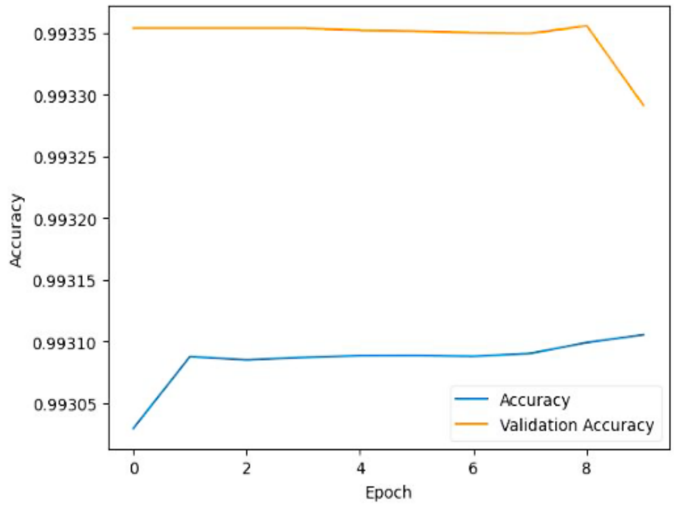
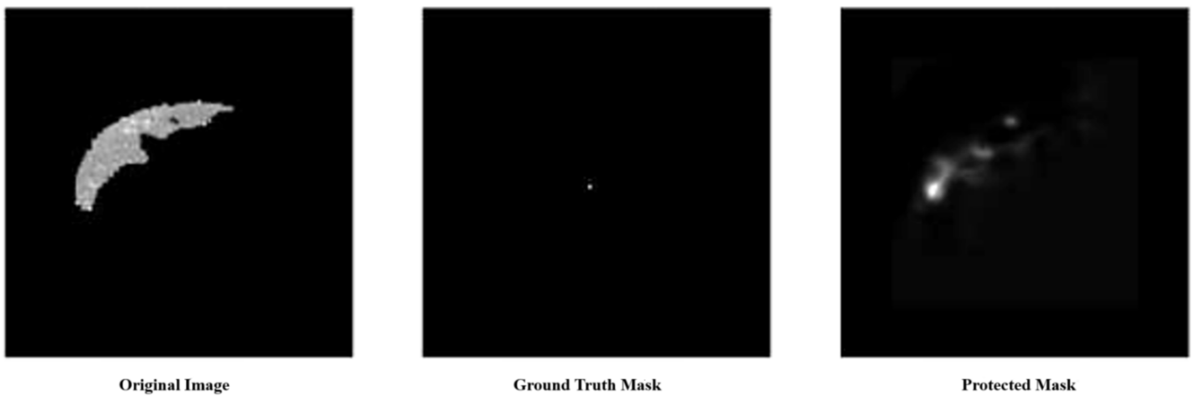


Fig. 3. Accuracy Calculation



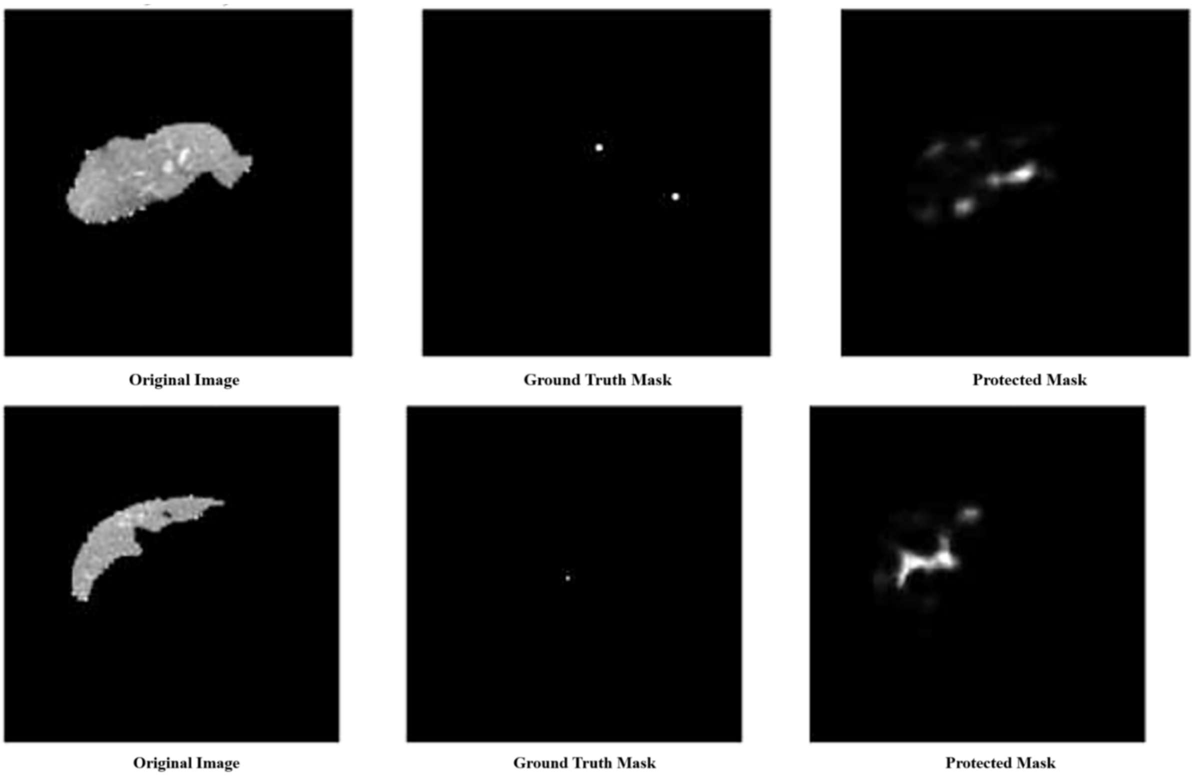


Fig. 4. Image Segmentation

In the third section of the methodology, the researchers explain how they evaluated their hybrid model's performance using established medical image segmentation metrics. The researchers first trained their proposed ResUNet-Transformer model using the training dataset, then evaluated its performance on a separate test dataset to assess generalization ability. The results demonstrated that the hybrid ResUNet-Transformer model achieved significant performance improvements in medical image segmentation compared to baseline models.

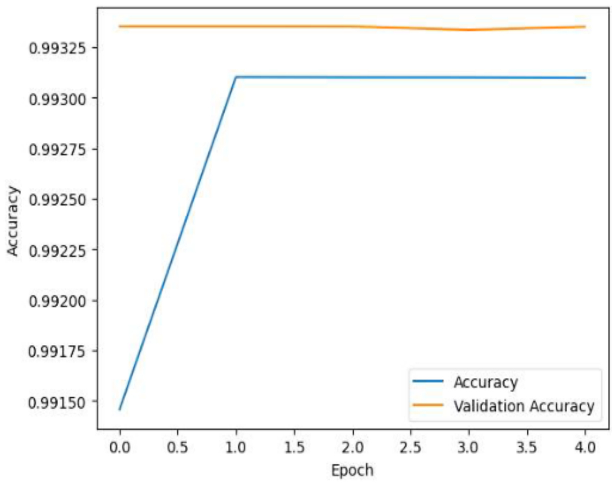


Fig. 5. Accuracy Calculation

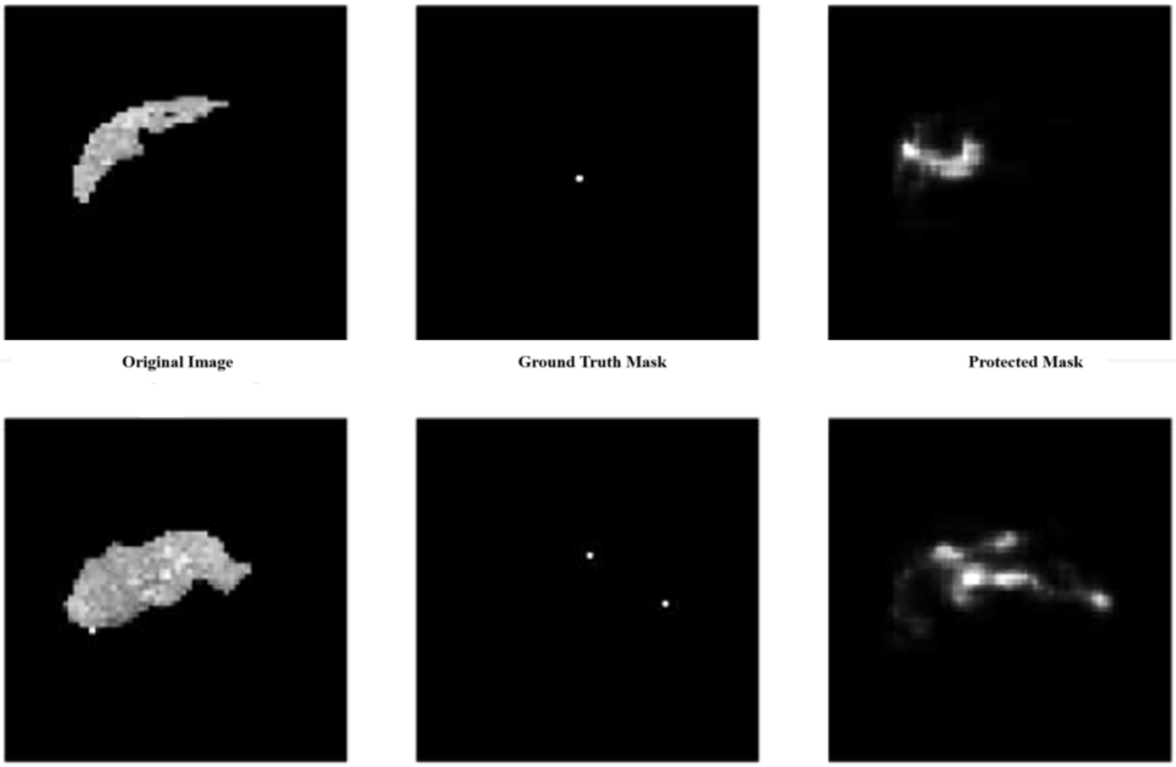


Fig. 6. Image Segmentation

4. PROPOSED METHOD

In this research, the main objective was to design and implement a hybrid model for medical image segmentation with a focus on identifying liver tumors. To accomplish this, first, image data including CT scans and masks related to tumor regions were extracted from a specialized dataset. The input images and masks were normalized to 64×64 dimensions to improve model generalization. For enhanced model performance, preprocessing techniques like normalization, rotation, flipping, and image translation were applied.

The proposed model combines ResUNet architecture with transformer layers. In this architecture, ResUNet50 serves as the encoder (feature extraction) component. ResNet with its residual blocks was employed to extract deep features from images, helping to address the problems of vanishing gradients and enabling better learning of spatial features in deep layers. After feature extraction, transformer layers were utilized to capture long-range and global relationships between pixels. These layers incorporate a Multi-Head Attention structure that enables the model to understand non-linear and complex relationships within the image. The transformer output then enters the decoder section with a U-Net structure. In this section, up sampling operations and convolutional layers reconstruct the prediction maps to their original dimensions. Additionally, skip connections between the encoder and decoder were implemented to preserve spatial information and improve accuracy in detecting small regions and complex boundaries. For evaluating the model's performance, metrics including accuracy (Accuracy), Jaccard coefficient (IoU), and Dice coefficient (Dice Coefficient) were used.

The proposed ResUNet-Transformer model showed significant improvements compared to baseline models in liver tumor segmentation tasks. This methodological approach demonstrates how combining the strengths of convolutional architectures (for feature extraction) with transformer capabilities (for modeling global relationships) creates a more effective solution for the challenging task of medical image segmentation, particularly for identifying liver tumors with their often subtle and complex boundaries. The ResUNet-Transformer model was designed in this research as an advanced combination of ResUNet architecture with transformer layers to improve medical image segmentation performance. This model incorporates several key architectural elements that work together to enhance segmentation accuracy: First, it builds upon the ResUNet50 architecture as the encoder (feature extractor). ResUNet50, which is based on ResNet with its residual blocks, provides a powerful foundation for extracting complex and deep features from medical images while addressing the vanishing gradient problems that often occur in deep networks. This allows the model to capture meaningful spatial features. After this initial feature extraction, the features are directly passed to transformer layers. The transformer component is specifically designed to learn global relationships and long-range dependencies between different regions in the image.

This is achieved through the multi-head attention structure (Multi-Head Attention), which allows the model to simultaneously attend to different relationships between pixels across the image. The transformer layers significantly improve the model's ability to understand contextual relationships between features. This enhanced contextual understanding proves particularly valuable for accurately identifying small regions and complex boundaries in medical images, which is crucial for precise liver tumor segmentation. By combining the local feature extraction strengths of ResUNet with the global relationship modeling capabilities of transformers, this hybrid architecture creates a more comprehensive approach to medical image segmentation that addresses limitations found in either architecture alone. After processing by the transformer layers, the data is transferred to the decoder (Decoder) section of the model. This decoder component has a structure similar to U-Net and is responsible for reconstructing the prediction maps to their original dimensions. The decoder section consists of convolutional layers and up sampling operations that gradually increase the resolution of the feature maps. One of the key architectural elements in this design is the use of skip connections (Skip Connections). These connections directly transfer spatial information from the encoder section to the decoder section at corresponding levels. By preserving and reintroducing these detailed spatial features, the model can maintain important structural information that might otherwise be lost during the deep feature extraction process.

This is particularly important for detecting the precise boundaries of tumors, which often require both contextual understanding and fine-grained spatial detail. To optimize computational efficiency while maintaining accuracy, the decoder uses a carefully selected number of convolutional filters. This approach allows the model to achieve high accuracy in identifying tumor regions while keeping computational complexity manageable. The balance of performance and efficiency is critical in medical imaging applications, where both accurate results and reasonable processing times are important. The ResUNet-Transformer model was trained using a dataset of CT scan images with corresponding tumor masks. After preprocessing, the data underwent several important preparation steps:

1. The images were normalized and resized to 64×64 dimensions to reduce computational requirements while preserving essential features.

2. Data augmentation techniques including rotation, flipping, and translation were applied to increase the diversity of the training data. These techniques played a crucial role in improving the model's generalization ability and reducing the risk of overfitting, which is particularly important in medical imaging where datasets may be limited in size.

To evaluate the model's performance, three standard metrics were used:

- Accuracy (overall pixel classification correctness)
- Jaccard coefficient (IoU - measuring overlap between predicted and ground truth)
- Dice coefficient (similarity measure between predicted and actual segmentations)

The results demonstrated the effectiveness of the hybrid approach, with the ResUNet-Transformer achieving:

- 92.5% accuracy
- 75.4% Jaccard coefficient
- 83.1% Dice coefficient

These metrics represent a substantial improvement over baseline models like U-Net and ResUNet. The hybrid model showed particular strength in identifying regions with low contrast and complex structures, areas where other models often encountered difficulties. This improvement can be attributed to the combination of ResUNet's feature extraction capabilities with the transformer's ability to model global relationships within the images. The impressive performance metrics validate the researchers' architectural choices and demonstrate how combining complementary approaches convolutional networks for local feature extraction and transformers for modeling long-range dependencies can create more effective solutions for challenging medical image segmentation tasks

5. CONCLUSION

This research focused on designing and evaluating the hybrid ResUNet-Transformer model for medical image segmentation, specifically for identifying liver tumors. The findings demonstrate several important advances in this field. The results clearly show that combining ResNet architecture (serving as the encoder) with U-Net's decoder structure and transformer layers creates a powerful synergy. ResNet's residual blocks effectively extract deep features, while the transformer layers enhance the model's ability to capture global relationships and long-range dependencies. This combination led to significant performance improvements compared to traditional methods. These metrics reveal the model's strength in accurately identifying complex regions, including areas with unclear boundaries and low contrast that typically challenge conventional segmentation approaches.

The transformer layers proved particularly valuable in addressing limitations found in purely convolutional architectures by modeling non-local relationships between image features. Additionally, the use of skip connections between the encoder and decoder preserved critical spatial information, allowing the model to maintain fine details necessary for precise boundary detection. This architectural choice was especially beneficial for identifying smaller tumor regions with complex boundaries. While the analysis showed the model performed exceptionally well overall, error analysis revealed some remaining challenges in areas with extremely low contrast. This finding points to opportunities for future research and optimization. The ResUNet-Transformer represents a meaningful advancement in medical image segmentation technology. By combining the strengths of different architectural approaches, it offers a more robust solution for the challenging task of liver tumor segmentation, with potential applications across other medical imaging domains where precise identification of complex anatomical structures is critical.

Declaration

We acknowledge that we used ChatGPT to enhance the academic writing of our manuscript while ensuring the originality and integrity of our work.

Transparency Statement

The data supporting this study are available upon reasonable request to the corresponding author, subject to ethical and confidentiality considerations.

Acknowledgments

We would like to express our gratitude to all individuals who contributed to this project.

Declaration of Interest

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

REFRENCES

[1] Jha, D., Smedsrud, P. H., Riegler, M. A., Johansen, D., De Lange, T., Halvorsen, P., & Johansen, H. D. (2019, December). ResUNet++: An advanced architecture for medical image segmentation. In 2019 IEEE International Symposium on Multimedia (ISM) (pp. 225–226). IEEE. https://doi.org/10.1109/ISM46123.2019.00049

[2] Wang, Z., Zou, Y., & Liu, P. X. (2021). Hybrid dilation and attention residual U-Net for medical image segmentation. Computers in Biology and Medicine, 134, 104469. https://doi.org/10.1016/j.compbiomed.2021.104449

[3] Salpea, N., Tzouveli, P., & Kollias, D. (2022, October). Medical image segmentation: A review of modern architectures. In European Conference on Computer Vision (pp. 708–761). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-25082-8_47

[4] Huang, L., Miron, A., Hone, K., & Li, Y. (2024, June). Segmenting medical images: From U-Net to Res-UNet and nnU-Net. In 2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS) (pp. 459–463). IEEE. https://doi.org/10.1109/CBMS61543.2024.00086

[5] Pan, H., Gao, B., Bai, W., Li, B., Li, Y., Zhang, M., … Kong, W. (2023). WA-ResUNet: A focused tail-class MRI medical image segmentation algorithm. Bioengineering, 10(8), 945. https://doi.org/10.3390/bioengineering10080945

[6] Wu, X., Tang, P., & Zhang, B. (2024, September). Research on image segmentation based on ResUNet. In 2024 IEEE 7th International Conference on Information Systems and Computer Aided Education (ICISCAE) (pp. 342–347). IEEE. https://doi.org/10.1109/ICISCAE62304.2024.10761354

[7] Baldeon-Calisto, M., & Lai-Yuen, S. (2018). ResU-Net: Residual convolutional neural network for prostate MRI segmentation. In IIE Annual Conference Proceedings (pp. 736–741). Institute of Industrial and Systems Engineers.

[8] Elghazy, H. L., & Fakhr, M. W. (2022). Dual- and triple-stream RESUNET/UNET architectures for multi-modal liver segmentation. IET Image Processing, 17(5), 1232–1245. https://doi.org/10.1049/ipr2.12708

[9] Huang, S., Han, X., Fan, J., Chen, J., Du, L., Gao, W., … Yang, J. (2021). Anterior mediastinal lesion segmentation based on two-stage 3D ResUNet with attention gates and lung segmentation. Frontiers in Oncology, 10, 513837. https://doi.org/10.3389/fonc.2020.618357

[10] Weng, W., & Zhu, X. (2021). INet: Convolutional networks for biomedical image segmentation. IEEE Access, 9, 16630–16591. https://doi.org/10.1109/ACCESS.2021.3053408

[11] Zhang, X., Zheng, Y., Bai, X., Cai, L., Wang, L., Wu, S., … Huang, J. (2022). Femoral image segmentation based on two-stage convolutional network using 3D-DMFNet and 3D-ResUNet. Computer Methods and Programs in Biomedicine, 226, 107112. https://doi.org/10.1016/j.cmpb.2022.107110

[12] Chincholkar, V., Srivastava, S., Pawankar, A., & Chaudhari, S. (2024, January). Deep learning techniques in liver segmentation: Evaluating U-Net, Attention U-Net, ResNet50, and ResUNet models. In 2024 15th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 779–775). IEEE. https://doi.org/10.1109/Confluence60223.2024.10463507

[13] Niu, K., Guo, Z., Peng, X., & Pei, S. (2022). P-ResUNet: Segmentation of brain tissue with purified residual U-Net. Computers in Biology and Medicine, 151, 106254. https://doi.org/10.1016/j.compbiomed.2022.106294

[14] Appati, J. K., Azuponga, N. A., Boante, L. M., & Mensah, J. A. (2024). A multi-attention ResUNet and

modified U-Net architecture for liver tumor segmentation. Applied Computational Intelligence and Soft Computing, 2024, Article 8263549. https://doi.org/10.1155/2024/8365349

[15] Wodzinski, M., & Müller, H. (2023). Automatic aorta segmentation with heavily augmented, high-resolution 3-D ResUNet: Contribution to the SEG_A challenge. In MICCAI Challenge on Segmentation of the Aorta (pp. 54–62). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-53241-2_4

[16] Sheng, N., Liu, D., Zhang, J., Che, C., & Zhang, J. (2021). Second-order ResU-Net for automatic MRI brain tumor segmentation. Mathematical Biosciences and Engineering, 18(5), 4967–4983. https://doi.org/10.3934/mbe.2021251

[17] Zang, L., Liang, W., Ke, H., Chen, F., & Shen, C. (2023). Research on liver cancer segmentation method based on PCNN image processing and SE-ResUNet. Scientific Reports, 13(1), 13779. https://doi.org/10.1038/s41598-023-39240-0

[18] Sun, M., Li, X., & Sun, W. (2024). Image generation and lesion segmentation of brain tumors and stroke based on GAN and 3-D ResU-Net. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3383880

[19] Han, Z., Jian, M., & Wang, G. G. (2022). ConvUNeXt: An efficient convolution neural network for medical image segmentation. Knowledge-Based Systems, 253, 109512. https://doi.org/10.1016/j.knosys.2022.109512

[20] Chincholkar, V., Srivastava, S., Pawanarkar, A., & Chaudhari, S. (2024, January). Deep learning techniques in liver segmentation: evaluating U-Net, Attention U-Net, ResNet50, and ResUNet models. In 2024 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 775-779). IEEE.https://doi.org/10.1109/Confluence60223.2024.10463507