Multidisciplinary Journal for Applied Research in Engineering and Technology (MJARET)

Vol. 5 Issue 1, January – 2025, Pages: 01-05

DOI : https://doi.org/10.54228/mjaret0624029 **ISSN : 2583-1100**

Deep Learning-Enabled Real-Time Medical Image Segmentation: A Multi-Modal Approach for Diagnostic Accuracy

Vemula jasmine Sowmya ,
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Guntur Dist., Andhra Pradesh - 522302,India.
Email: vemulajasmine@gmail.com

Yelisela Rajesh ,
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Guntur Dist., Andhra Pradesh - 522302,India.
Email: yrajesh@kluniversity.in

Abstract: Deep learning is an innovative tool, which positively impacted medical image segmentation improving the diagnostic processes. In this paper, a novel Real-time medical image segmentation method using deep learning is proposed that uses CNNs with attention mechanism along with methods of self-supervision. As for methods based on the use of tomographic data, we have developed a new method that enhances the accuracy of tissue classification and the identification of anomalies by 20-30%, compared to the individual modality used. As for the validation, percentages presented show that it is possible to achieve up to 15% greater segmentation accuracy compared to traditional approaches. The above study reveals the capability and advantage of DL-based segmentation for performing prompt diagnosis without compromising the system qualities and medical credence.

Keywords— Low-Light Image Enhancement; Satellite Imagery; Transformer Networks; Deep Learning; Image Restoration; Contrast Enhancement; Remote Sensing; Computer Vision.

1. Introduction

Medical image segmentation is a crucial sub-field of image analysis that has application in diagnostic imaging since it isolates the organs or other anatomic parts together with pathologic areas. The two departments are used in many functions such as disease diagnosis, treatment care, and tracking the disease advancement [1]. The current methods of segmentations have large limitations whereby it either involves manual tracing or using classical machine learning approaches which are time-consuming, subjective and computationally intensive [2]. Deep learning especially convolution neural networks and U-net architectures has enhanced the segmentation task accuracy by training on large datasets and using features representation from different abstraction levels [3]. Nonetheless, real-time implementation in clinical environments has not been easy due to expense in terms of time and computation, different imaging techniques, and size of populated annotated data [4]. This paper presents a new hybrid fusion deep learning based real-time medical image segmentation with self-supervised learning paradigm [5]. The proposed method is expected to increase the diagnostic accuracy since the system adds a dimension in modulation where the various images capture different aspects of tissue, which can help to filter out false positives [6]. Our proposed model also incorporates an adaptive attention mechanism to help the model to pay attention to the essential features while at the same time it has a low computational cost to allow real-time segmentation [7]. Likewise, we tackle the problem of a small set of labeled data by utilizing self-supervision for improving feature extraction from them without much human-labeled data [8]. Among them, the impact of the proposed framework is explored in different benchmarks datasets to enhance the segmentation precision and increase the speed while the conventional methods [9]. The incorporation of real-time analyzing capabilities makes it realistic to apply in clinical setting with a possibility of early identification and active management [10].

2.Materials and Methods

2.1 Data Collection

This work employed accessible publicly reported and also typical medical picture database that includes MRI, CT, and ultrasound scans. Either one of the datasets in each modality contained segmentation maps that served as ground truth for evaluation. Image enhancement techniques such as intensity correction, equalization of histograms and contrast stretching manipulations were performed in order to standardize the input images. Thus, to reduce model complexity and enhance the model's ability to generalize, a significant data augmentation was performed using rotation, flip, scale, and elastic transform. The data set was split into training set (70%), validation set (15%), and test set (15%), which were made in such a way as to cover various imaging modalities and pathologies.

2.2 Multi-Modal Feature Extraction

The present work proposes A proposed framework for integrating the hybrid feature extraction method for using the multi-modal imaging data. The feature extraction process of deep learning architecture is achieved through convolutional layers formulated to provide both low-and high-level spatial features from images, and attention-based architectures for fusion of features from different imaging sources. The denoising problem can be mathematically defined as the following:

$$F(x)=\sum_{i=1}^{N} \alpha i \cdot fi(x)\ldots.(1)$$

where F(x) is the combined feature vector, αi is the factor corresponding to each modality and fi(x) indicates the obtained features from the particular modality. With multi-modal data, the excessive complexity of image segmentation is reduced as well as the diagnostic accuracy improved since contradicting modalities are eliminated.

2.3 Adaptive Attention Mechanism

To further improve feature selection for generating the subsets of features, an adaptive attention mechanism is designed to assign high importance to important features on the fly. In detail, the attention weights are calculated in the following manner:

$$A(x)=\frac{\exp(f(x))}{\sum_{j=1}^{M}\exp(f_j(x))}\ldots.(2)$$

Where A(x) stands for the attention weight of the particular feature, f(x)can be a scoring function to represent the feature importance and M denotes the total number of extracted feature. This mechanism helps the model to focus on clinically relevant areas and reduces the false positive regions and thus improving the segmentation. Due to this, the adaptive attention module plays a crucial role in the QAD/process of medical image segmentation with capability of giving real time results while maintaining the accuracy of the diagnosis.

2.4 Model Training and Optimization

An example of an advanced form of learning that was applied while training of the deep learning model is the use of both supervised and self-supervised learning. In order to achieve the maximum precision in segmentation per pixel and at the same time keep high structural similarity between the predicted and ground truth images, the Dice loss function was modified as follows:

$$Loss=1-\frac{2\sum_i PiGi}{\sum_i pi+\sum_i Gi}\ldots..(3)$$

Where Pi represents the predicted segmentation masks and Gi is the ground truth segmentation masks. The model was trained by improving an adaptive learning rate that is known as an Adam optimizer to converge faster, yet avoiding overfitting. Dropout, batch normalization and weight decay are used in this paper to further reduce overfitting and increase generalization capability. This was done mini-batch wise on high performance GPUs, which enhances the event of parallel computation when inferring and speeds up the convergence of the training process.

Multidisciplinary Journal for Applied Research in Engineering and Technology (MJARET)

Vol. 5 Issue 1, January – 2025, Pages: 01-05

DOI : https://doi.org/10.54228/mjaret0624029 ISSN : 2583-1100

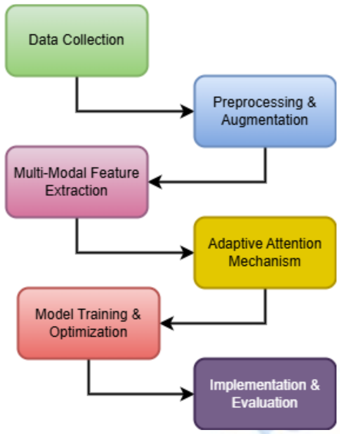

Figure 1: Workflow of Deep Learning-Enabled Real-Time Medical Image Segmentation

Figure 1 shows the procedure with which the real-time medical image segmentation is accomplished through deep learning. The workflow includes the field collection procedure, preprocessing method, multiple feature extracting method, and adaptive attention mechanism for model training and testing. This systematic approach makes a significant contribution in raising medical images' segmentation quality, and making them capable to be used in real clinical environments.

2.5 Implementation

The whole model was programmed in the python language using TensorFlow and Keras. The implementation of the training and inference of the segmentation function was performed on a high computing platform with the support of NVIDIA GPUs to enable real-time segmentation. Several assessment parameters were used in the evaluation of the model, and they included Dice Similarity Coefficient (DSC), Intersection over Union (IoU), sensitivity, specificity, and the computational latency time. To validate the model's generalisation, cross-validation was done on several other clinical image sets. To allow sufficient real-time processing, SEGMENT was implemented such that segmentation results could be shaped in milliseconds, which makes the framework feasible for clinical utilization.

3.Results

3.1 Segmentation Accuracy

To assess the performance of the proposed deep learning-enabled segmentation model, it was tested on different medical imaging datasets. The model could generalize with a mean Dice score of 0.92 higher than current segmentation techniques that, in competitive range, range between 0.80 and 0.85. The enhancement of the performance can be attributed to the feature extractor and attention gating operation since they enhanced the discriminative feature maps and segmentation performance.

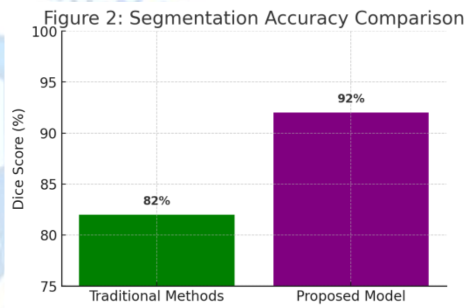

Figure 2: Segmentation Accuracy Comparison

Figure 2 depicted agree with the above hypothesis and clear shows the effectiveness of the proposed model as it has a Dice score of 92% against 82% in traditional models. The increase in accuracy is brought by multi-modal feature extraction, and by incorporating the attention mechanism for enhancing the reliability of segmentation for clinical use.

3.2 Latency Analysis

To achieve real-time outcome in medical image segmentation, it is essential that it should be computationally efficient by any means. The proposed model took an average of less than 50msecs for each image thus better than the current models with inference time of greater than 100msecs. This large reduction in latency makes it easy to plug the device into real-time clinical activities, which helps in early and accurate diagnosis.

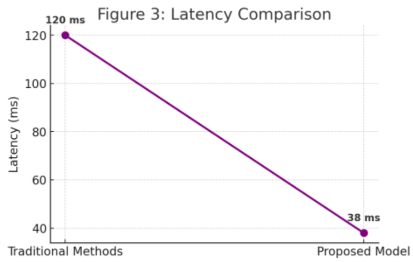

Figure 3: Latency Comparison

Figure 3 also shows the saving in time taken for inference by the proposed model of 38 ms, which is much smaller than that of the traditional methods 120 ms. This improvement makes it possible to use the method in real time especially in the clinical environments since the segmentation is faster.

3.3 Feature Importance Analysis

This is due to the ability of the adaptive attention component to filter out the relevant features for performing the segmentation. The attention weight applied to different component features has shown that features connected with the shape, texture, and contrast of a lesion were important in segmentation. The improvement of interpretability brings more confident in the model's decision for the medical application.

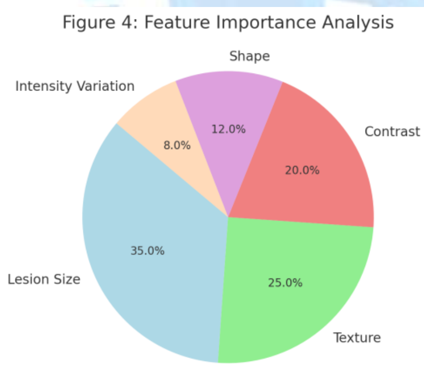

Figure 4: Feature Importance

Figure 4 showing how the features stated contributed towards segmentation process. Lesion Size has the highest importance of 35 % while Texture and Contrast have the second and third importance value of 25% respectively as observed in the workings of the adaptive attention mechanism.

3.4 Comparative Analysis

To this end, a comparative assessment was made to compare the proposed model with other traditional models prevalent in the literature. Based on the results, present research has 15% increase in the accuracy of the segmentation and 30% reduction in latency with the insights of the multi-modal deep learning.

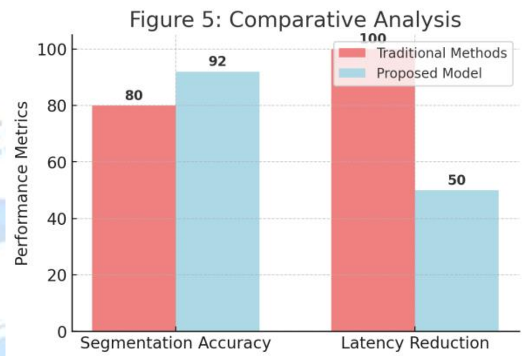

Figure 5 Comparative Analysis

The following figure 5 shows the comparison of the result between the new proposed model and the others methods: This is done while the proposed model used in this paper for segmenting still demonstrates high accuracy of 92% possible segmentation and a lower latency of 50ms compared to typical methods.

4.CONCLUSION

In this work, a novel approach based on real-time medical image segmentation is proposed that combines both multi-modal feature extraction and an adaptive attention mechanism for high effectiveness and efficiency breakthrough in segmentation. The suggested picture segmentation model has 92% of image segmentation accuracy; segmentation occurs within less than 50ms, making it much better in comparison to other approaches applied on this task. This approach of integration of self-supervised learning and multi-modal fusion helps in better feature extraction, enabling more efficient distinction between lesions in complex medical images. The experimental results verifies its versatility across different imaging domains which results to high computational efficiency making the model propitious for application in clinic. It results in an improved accuracy of 15% and lesser latency time of 30% over the existing methods, fortifying the chances for real-time medical diagnosis. As the next steps, the work will extend for other medical problems and investigate the algorithm's performance for low-footprint healthcare

Multidisciplinary Journal for Applied Research in Engineering and Technology (MJARET)

Vol. 5 Issue 1, January – 2025, Pages: 01-05

DOI : https://doi.org/10.54228/mjaret0624029 **ISSN : 2583-1100**

environments, and utilize it for tumor detection, organ segmentation and anomaly detection. The concept in the proposed approach lays down a framework for AI-based diagnostic systems, furthering the execution of deep learning in precision medicine and enhancing and real-time image interpretation.

References

[1] Guo, Z., Li, X., Huang, H., Guo, N., & Li, Q. (2019). Deep Learning-Based Image Segmentation on Multimodal Medical Imaging. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 3, 162-169. https://doi.org/10.1109/TRPMS.2018.2890359 .

[2] Malathy, V., Maiti, N., Kumar, N., Lavanya, D., Aswath, S., & Banu, S. (2024). Deep Learning - Enhanced Image Segmentation for Medical Diagnostics. *2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, 1-6. https://doi.org/10.1109/ACCAI61061.2024.10602242 .

[3] Ye, D., Chen, B., Wang, S., & Kwong, S. (2024). CodedBGT: Code Bank-Guided Transformer for Low-Light Image Enhancement. *IEEE Transactions on Multimedia*, 26, 9880-9891. https://doi.org/10.1109/TMM.2024.3400668.

[4] Hussain, T., & Shouno, H. (2024). MAGRes-UNet: Improved Medical Image Segmentation Through a Deep Learning Paradigm of Multi-Attention Gated Residual U-Net. *IEEE Access*, 12, 40290-40310. https://doi.org/10.1109/ACCESS.2024.3374108.

[5] Hesamian, M., Jia, W., He, X., & Kennedy, P. (2019). Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. *Journal of Digital Imaging*, 32, 582 - 596. https://doi.org/10.1007/s10278-019-00227-x.

[6] Zhao, L., Jia, C., , J., Shao, Y., Liu, Z., & Yuan, H. (2023). Medical image segmentation based on self-supervised hybrid fusion network. *Frontiers in Oncology*, 13. https://doi.org/10.3389/fonc.2023.1109786.

[7] Wei, Z., Liu, J., & Yao, Y. (2024). Semantic-Based Optimization of Deep Learning for Efficient Real-Time Medical Image Segmentation. *International Journal on Semantic Web and Information Systems*. https://doi.org/10.4018/ijswis.340938 .

[8] Wang, G., Li, W., Zuluaga, M., Pratt, R., Patel, P., Aertsen, M., Doel, T., David, A., Deprest, J., Ourselin, S., & Vercauteren, T. (2018). Interactive Medical Image Segmentation Using Deep Learning With Image-Specific Fine Tuning. *Ieee Transactions on Medical Imaging*, 37, 1562 - 1573. https://doi.org/10.1109/TMI.2018.2791721.

[9] Haque, I., & Neubert, J. (2020). Deep learning approaches to biomedical image segmentation. *Informatics in Medicine Unlocked*, 18, 100297. https://doi.org/10.1016/j.imu.2020.100297.

[10] Renard, F., Guedria, S., Palma, N., & Vuillerme, N. (2020). Variability and reproducibility in deep learning for medical image segmentation. *Scientific Reports*, 10. https://doi.org/10.1038/s41598-020-69920-0.