

Segmentation of Medical MRI Images Using Nested U-Net with Attention Mechanism and Fuzzy Pooling

Noor M. Basheer
noor.23enp115@student.uomosul.edu.iq

Ali Al-Saegh
ali.alsaegh@uomosul.edu.iq

Computer Engineering Department, College of Engineering, University of Mosul, Mosul, Iraq.

Received: April 22th, 2025 Revised: June 27th, 2025 Accepted: July 28th, 2025

ABSTRACT

Segmentation of medical images is a crucial step in the diagnosis and treatment of various diseases, particularly when analyzing magnetic resonance imaging (MRI) scans. Despite significant progress, accurate segmentation remains challenging due to the complexity of anatomical structures, ambiguous boundaries, and variations in image quality. This paper proposes an enhanced Nested U-Net structure that incorporates an attention mechanism and fuzzy pooling to improve retail performance. Nested U-Net benefits from extensive skip connections to capture multi-level contextual information, while the attention mechanism enhances the model's ability to focus on relevant features and noise suppression. In addition, we replace the traditional extreme pooling layers with fuzzy pooling, which enables the network to handle spatial ambiguity more effectively and produce more accurate and refined retail maps. Experimental results on publicly available MRI datasets show that the proposed model achieves remarkable improvements, with a 4.97% increase in the dice coefficient, a 2.76% improvement in accuracy, and a 4.25% increase in recall compared to enhanced U-Net structures. Furthermore, significant improvements have been observed in Hausdorff's distance measures.

Keywords:

Attention Mechanism, Deep Learning in Medical Imaging, Fuzzy Pooling, Nested U-Net, MRI Image Segmentation, Semantic Segmentation.

This is an open access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://rengj.uomosul.edu.iq>

Email: alrafidain_engjournal3@uomosul.edu.iq

1. INTRODUCTION

Segmentation of medical images is a critical and institutional task in the field of medical imaging, as it supports a wide range of healthcare applications, including the identification of abnormalities in tissues, the detection of tumors, and surgical planning assistance. Among imaging methods, magnetic resonance imaging (MRI) has gained prominence due to its non-invasive nature and exceptional ability to distinguish between soft tissues, making it particularly suitable for segmenting complex anatomical structures such as the brain, heart, and tumors [1]. However, accurately segmented MRI images remain a major challenge. Factors such as image noise, low variability, and differences in anatomical structures across different patients often complicate the process [2].

To overcome these challenges, deep learning models have shown great promise, particularly the U-Net architecture [3], which has demonstrated strong performance in medical

image segmentation due to the decoding structure and use of skip connections that maintain spatial information [4]. U-Net features a shrinking path that captures context and an expanding path that restores spatial resolution, in addition to skip connections that help preserve essential features lost during sample reduction[5]. Its ability to perform well even with limited annotated data has made it particularly suitable for segmenting fine anatomical details [6]. Despite this success, traditional U-Net models can still fall short when it comes to modelling complex spatial relationships and detecting precise features, especially in low-contrast images or ambiguous boundaries [7].

To address these limitations, more advanced variants have been proposed, such as U-Net Mesh (also known as U-Net++) [8], which improves on the original structure by incorporating dense skip connections and deep supervision to reduce the semantic gap between the Encoder and decoder. This results in smoother transitions and improved segmentation accuracy [9]. In these

embodiments, the encoder captures multi-scale features through convolution and pooling operations, while the decoder reconstructs detailed masks using hierarchical intermediate layers that facilitate gradient flow and feature reuse.

In order to further enhance deep learning structures, attention mechanisms have been introduced. These mechanisms dynamically weigh features based on their importance, allowing the model to focus on vital areas of interest while suppressing irrelevant information. In medical image segmentation, this ability is particularly valuable for detecting small or poorly defined areas, thereby enhancing boundary resolution and overall accuracy.

In parallel, ambiguous aggregation has emerged as a strong reinforcement of traditional aggregation processes in conventional networks [10]. Unlike maximum or medium pooling, fuzzy pooling leverages ambiguous logic to assign an organic score to each element within the pooling window, reflecting the degree to which it belongs to specific, distinct groups [11]. This results in a more precise and flexible pooling process that can better handle uncertain or inaccurate standard features in medical images. Fuzzy pooling has shown remarkable benefits in tasks such as object detection and speech recognition, particularly in the analysis of medical images, where preserving accurate spatial information is crucial [12].

Modern segmentation models increasingly utilize attention and multi-scale learning to enhance focus and context, as seen in TransAttUNet and EnigmaNet. However, they often require high computational resources and large data sets, as they face difficulties in dealing with ambiguous boundaries and class imbalances, especially in cardiac structures. Converter-based models such as LeViT-UNet improve global context but face deployment challenges. CNN-based variants (e.g., Residual-Attention UNet++, ST-HarDNet) enhance spatial features but ignore boundary uncertainty [13]. The model addresses these gaps by incorporating Nested U-, spatial attention, and fuzzy pooling to improve boundary accuracy and robustness in cardiac MRI segmentation through efficient computation.

Based on these developments, this work proposes a new deep learning model for segmenting medical MRI images that integrates three major innovations: (1) the overlapping U-Net structure, which expands classic U-Net with heavy skip connections to enable deeper supervision and extraction of wealthier multi-scale features [8]; (2) a mechanism of interest to improve the model's focus on vital features, especially in complex or noisy areas; (3) fuzzy pooling to replace traditional

pooling layers, which allows the network to capture spatial ambiguity more effectively and produce more precise and more accurate retail maps.

Despite significant progress on the ground, many challenges remain. The scarcity of annotated MRI datasets remains a bottleneck, particularly in rare cases such as gliomas or specific cardiovascular diseases [14]. In addition, while the latest models of interest and vague mechanisms have shown impressive segmentation accuracy, issues of computational cost, model interpretability, and durability of invisible data still need to be addressed [15].

Future research in this field can explore the development of more efficient and interpretable hybrid structures [16] that combine the strengths of various advanced components, such as attention layers, vague logic-based units, and specially designed loss functions. Furthermore, leveraging knowledge and area-specific techniques such as semi-supervised learning or increased data can enhance performance in specific data scenarios [17]. The key contributions of this study are:

- Development of an enhanced Nested U-Net with integrated spatial attention and fuzzy pooling layers.
- Comprehensive evaluation of the model on the ACDC cardiac MRI dataset for both segmentation and classification.
- Analysis of computational efficiency ensures the model is suitable for clinical application deployment.

The remainder of the paper is organized as follows: Section 2 reviews several related works, Section 3 gives theory about the used techniques and methods, Section 4 illustrates the experimental setup achieved in the paper, Section 5 presents the experimental results with comparison to similar works, Section 6 disuses the achieved results, and finally Section 7 concludes the paper.

2. RELATED WORKS

Over the past decade, the U-Net's structure has emerged as a foundational model for segmenting medical images due to its encrypted structure, decoding, and skipping connections that preserve spatial information even in limited data environments [17]. Various modifications of U-Net have been developed to enhance segmentation outcomes in challenging medical imaging contexts, including the segmentation of brain tumors, heart structures, and abdominal organs. Working at [18] provides a comprehensive analysis of the U-Net family, detailing

improvements such as deep layers and overlapping structures to enhance performance in high-complexity areas. For example, U-Net++ (also known as Nested U-Net or Densely Connected U-Net) utilizes dense skipping pathways, which have proven particularly effective in segmenting white blood cells and other complex anatomical features [19]. These improvements highlight the effectiveness of overlapping and overlapping structures in capturing precise spatial and contextual details in biomedical images.

To further improve segmentation accuracy, especially in scenarios involving fuzzy tissue boundaries or class imbalances, researchers explored the mechanisms of attention. The study identifies in [20] the basic types of interest in deep learning - channel, place, and time - and their applications across different vision tasks. Integrating interest into U-Net variables has shown remarkable improvement. For example, the U-Net's multi-band overlapping network with integrated attention [21] showed an improvement in localizing features in biomedical segmentation. Similarly, U-Net++ [22] and the MCA-U-Net (Multi-level Cross Attention U-Net) network integrated into CBAM (Convolutional Block Attention Module) [23] expanded the concept by adding mutual interest across multiple scales, which led to more accurate feature maps. These methods highlight how attention can help focus the model's focus on areas most relevant to diagnosis, reduce false positives, and improve dice scores in segmentation tasks.

In parallel with attention, fuzzy pooling was investigated to overcome the constraints of maximum or medium conventional assembly, especially in dealing with noisy or low-variability MRI data. Unlike strict aggregation strategies, fuzzy aggregation leverages ambiguous logic to assign membership scores to features within the aggregation window, which allows the model to retain more accurate information [10][11]. This technique enhances the model's durability against spatial ambiguity and anatomical variation. For example, explore the work in [24] and [25] on extracting features deep into different networks using ambiguous or multilevel assembly methods to improve segmentation performance under challenging conditions. In liver and tumor segmentation, fuzzy pooling has demonstrated its ability to retain critical boundary information, especially when combined with attention mechanisms or lightweight structures such as SCU-Net (Semantic Context U-Net) [9]. These models have successfully addressed segmentation difficulties caused by unclear boundaries, anatomical complexity, and

similar grey values among neighboring structures.

3. MATERIALS AND METHODS

In this study, we propose an improved model of MRI medical image segmentation that integrates the overlapping U-Net structure with the attention mechanism and fuzzy pooling. This hybrid framework is designed to achieve superior segmentation accuracy by allowing the model to focus on critical anatomical structures while retaining accurate spatial information, as well as utilizing multi-scale to define boundaries accurately [25].

3.1. Architecture design

U-Net is a CNN architecture originally designed for biomedical image segmentation [5], featuring a decoding-encoding structure with skip connections that retain spatial detail [6]. To improve performance, Nested U-Net offers dense skip connections and deep supervision, which allows for better feature consolidation and more precise segmentation, as shown in Fig. 1 [8].

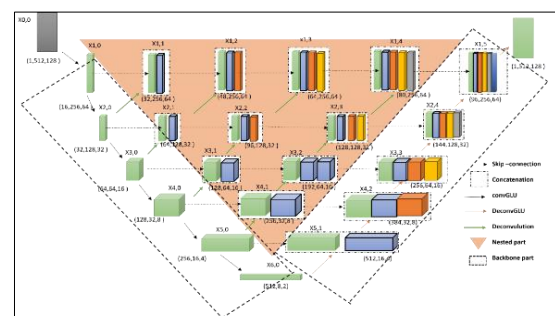


Fig. 1: The architecture of Nested U-Net.

Hybrid models that combine U-Net with attention and fuzzy pooling mechanisms are effective in medical image segmentation, providing improved feature extraction, focus, and boundary preservation [26][25]. In the model, fuzzy pooling is used in the encoder to retain fine details, while attention gates in skip connections enhance related features.

The fuzzy pooling structure replaces the standard maximum pooling in the encoder to better preserve uncertain boundary information. As spatial attention units are added before each decoding block, this allows the model to focus on key anatomical structures. Each decoding stage benefits from enhanced spatial cues via attention-rich skip connections. This integration clearly defines the roles of fuzzy logic and attention within the Nested U-Net framework, which distinguishes it from previous hybrid models.

Fig. shows the proposed model pipeline, which consists of the following stages:

Image acquisition and pre-processing: where the MRI heart images are resized to $256 \times 256 \times 3$. Preprocessing includes density normalization, Gaussian filtering, and TorchIO-based augmentation (rotation, translation, and morphological changes), which increases data diversity fourfold.

Feature extraction: In an encoder, fuzzy pooling replaces maximal pooling to preserve spatial detail. A mesh U-Net with dense skip connections extracts hierarchical features, facilitating robust learning.

Attention-enhanced decoding: Decoding blocks reconstruct spatial resolution. Attention gates improve skip communications by focusing on relevant structures (LV, MYO, RV).

Output generation: A 1×1 Convolution with sigmoid/softmax activation generates probability maps, which are then converted to binary values using a threshold (e.g., 0.5).

Post-processing: Morphological processes (opening, closing, and smoothing) remove noise and false positives, while also improving segmentation results.

Feature-based classification: From final masks, where anatomical features are extracted and entered into a classification unit, each patient is assigned to one of five diagnostic categories (NOR, MINF, DCM, HCM, RV) using rule-based or learned models [27].

3.2. Fuzzy pooling

Conventional pooling processes, such as maximum pooling, can overlook significant spatial features, particularly in complex medical images. Fuzzy pooling [28], introduced as an enhancement, integrates uncertainty modelling into the pooling process. This approach maintains more accurate features by applying vague logical principles to assess local neighborhood information [29], thereby improving durability and accuracy.

Mathematically, fuzzy pooling leverages fuzzy logic theory to map pixel values within a pooling window to membership degrees. Given an input patch $X = \{x_1, x_2, \dots, x_n\}$, the fuzzy membership function is defined as:

$$\mu(\mathbf{x}_i) = \left| \frac{1}{1 + \left| \frac{x_i - c}{\sigma} \right|} \right| \quad (1)$$

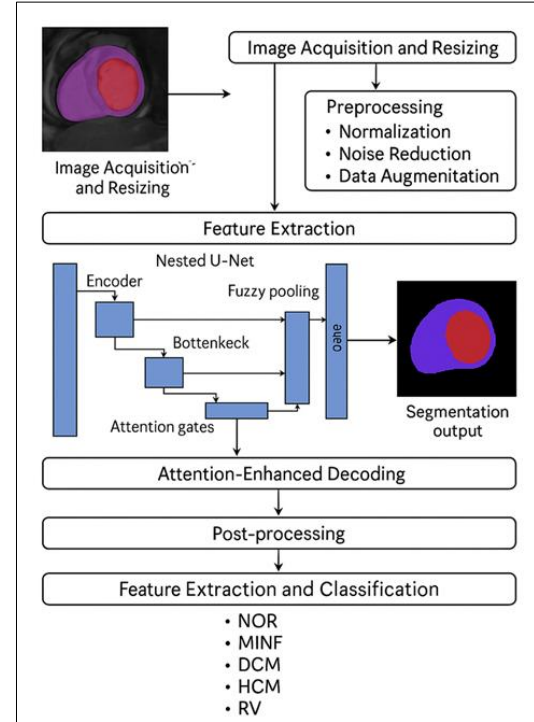


Fig. 2: Workflow of the proposed model.

where:

- x_i is the input value,
- c is the center of the fuzzy set (typically the mean of the window),
- σ controls the spread (standard deviation or other scaling),
- $p \geq 1$ determines the sharpness of the membership function.

The final pooled output y is computed as a fuzzy-weighted average:

$$y = \frac{\sum_{i=1}^n \mu(\mathbf{x}_i) \cdot x_i}{\sum_{i=1}^n \mu(\mathbf{x}_i)} \quad (2)$$

This ensures smoother transitions and more robust feature preservation compared to traditional max pooling or average pooling.

In our model, we replace standard maximum pooling layers with fuzzy pooling layers, as illustrated in Fig. 3. This ensures that important structural and contextual details are retained while the coverage is reduced. When integrated into Nested U-Net, fuzzy pooling enhances the continuity of features across layers and also helps capture boundaries more accurately, especially in low-noise or low-contrast images.

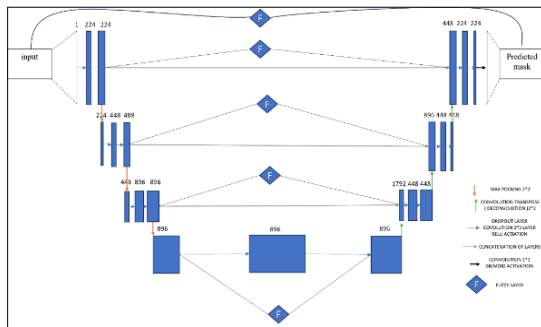


Fig. 3: Fuzzy U-Net Neural Network Design.

3.3. Attention mechanism

Attention mechanisms promote segmentation by allowing the model to focus on the areas most relevant to the input [26]. In the proposed model, attention gateways (AGs) are incorporated into the skip connections of a nested U-Net, resulting in feature filtering before decoding. These gates amplify important structures (e.g., myocardium, tumor boundaries) and suppress irrelevant noise, thereby improving localization and structural clarity [30].

Fig. 4, that this attention-enhanced U-Net architecture improves segmentation in noisy MRI scans or when the target area is small.

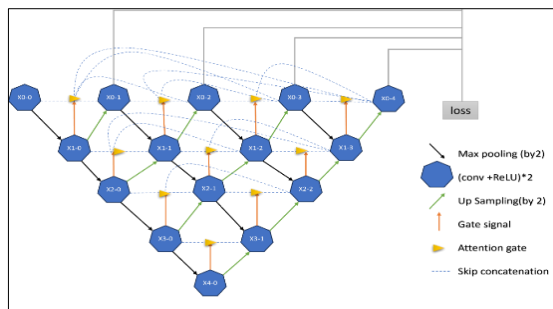


Fig. 4: Attention-based nested U-Net.

Fig. 5 shows how the attention mechanism operates, as the weighted inputs assign greater importance to critical areas in the image [31]. Context awareness enhances focus on diagnostic-related regions, such as organs or lesions [32]. Dynamic focus adjusts attention based on the Image's context, for example, focusing within tumor areas on healthy tissue [27].

3.4. Training strategy

The model has been implemented in PyTorch and trained with the Adam optimizer. A composite loss function that combines Dice loss and binary cross-entropy addresses class imbalance. The training process includes 150 epochs, with early stops based on dice score validation. The model was trained on 100 patients and validated on a separate group of 50 patients.

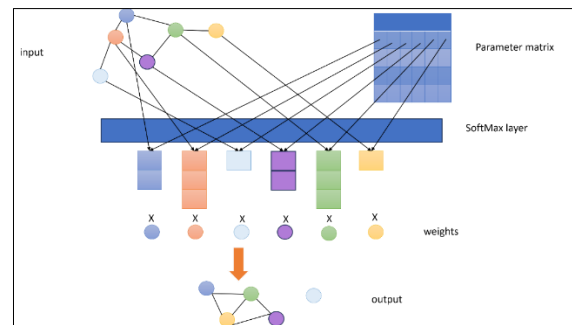


Fig. 5: The work of the attention mechanism.

4. EXPERIMENTAL SETUP

4.1. Software and libraries

The following libraries and toolboxes are used for building the model and simulating the work:

- PyTorch (v1.10.0 or above): a basic framework for building, training, and evaluating the zigzag U-Net model. PyTorch's dynamic calculation chart allows for easy experimentation and correction.
- TorchIO: A medical imaging library used to increase and normalize data and advance the processing of 3D medical images.
- NumPy: for numerical processes such as matrix processing, image conversions, and data processing.
- Matplotlib & Seaborn: to visualize results and create performance plots (e.g., dice scores, IoU, confusion matrix) and attention maps.
- SimpleITK: To handle and process NIfTI medical image files, regrouping procedure, and other standard imaging processes.
- SciPy: for statistical analysis, specifically in calculating Hausdorff's distance and making different image metrics.
- OpenCV: for image processing tasks, including post-processing steps such as thresholding and morphological processes.

4.2. Hardware specifications

All experiments were conducted on a system with the following specifications:

- Processor: 13th generation Intel (R) Core (TM) i7-13620H @ 2.40 GHz
- RAM: 16.0 GB (15.7 GB usable)
- System Type: 64-bit operating system, x64 processor

These specifications enable the efficient development of computational training and evaluation of models, removing it as it is mentioned in the previous comment, as well as the use of 16 GB RAM to easily handle large datasets and the intensive computational requirements of

deep learning models. A high-speed, 24-hour multi-core processor also ensures rapid training and processing of the model, particularly when handling large volumes of MRI data.

4.3. Dataset

The dataset used in this study is a publicly available ACDC dataset, part of the Automated Cardiac Diagnostic Challenge [ACDC Dataset, 2017]. It consists of short-axis cardiac MRI scans from 150 patients, as they were obtained during the end-systolic and end-diastolic phases. The data set was divided into 100 training cases and 50 testing cases. It includes patients diagnosed with NOR (normal), MINF (myocardial infarction), DCM (dilated cardiomyopathy), HCM (hypertrophic cardiomyopathy), and ARV (abnormal right ventricle) as explained in Table 1, covering a wide range of cardiac conditions. This dataset was then selected for high-quality ground truth annotations and widely used in benchmarking segmentation and classification models [19].

Table 1: Patients diagnosed.

Label	Condition	Number of Patients
NOR	Normal	7
MINF	Myocardial Infarction	12
DCM	Dilated Cardiomyopathy	3
HCM	Hypertrophic Cardiomyopathy	9
RV	Abnormal Right Ventricle	19
Total		50

Each volume of MRI volumes includes detailed anatomical structures of the left ventricle (LV), right ventricle (RV), and heart muscle (MYO) - the three core areas targeted for fragmentation. To ensure anatomical diversity and improve the robustness of the segmentation model, the dataset encompasses a wide range of cardiac morphologies and diseases.

Before training the model, all images underwent a comprehensive pipeline that involved normalizing density to standardize image variation and brightness across subjects. In addition, data augmentation techniques, such as geometric shifts and morphological processes, have been applied to artificially expand the dataset and increase its variability. This approach helps mitigate excessive processing and enhances the model's ability to generalize to unseen situations.

The use of a clearly defined train-test split not only supports replicability but also facilitates a fair and impartial assessment of the model's performance. This chapter ensures that the model's segmentation accuracy, generalization capability, and computational efficiency are assessed on

completely invisible data, reflecting its real-world applicability.

4.4. Evaluation metrics

To evaluate the performance of segmentation and classification models, several well-established metrics have been used, including the Dice Similarity Coefficient (DSC) for interference accuracy, the Hausdorff Distance (HD) for boundary accuracy, accuracy, precision, and recall, and the F1 score for classification quality and cross-union (IoU) for area-based interference. These metrics provide a comprehensive evaluation framework that addresses various aspects of model performance. These measures are shown in Table 2.

5. EXPERIMENTAL RESULTS

In this section, we present the results of experiments conducted to evaluate the proposed performance of Nested U-Net with an attention mechanism and Fuzzy Pooling model for segmenting medical MRI images. We compare the performance of our model to that of many current modern models to demonstrate its superior capabilities in segmentation accuracy, computational efficiency, and robustness.

5.1. Results of disease classification

Each test patient has been classified into one of the five subgroups for heart disease: normal (NOR), myocardial infarction (MINF), dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), and abnormal cardiomyopathy (Rether Ventricle). The classification results presented in Table A1 in the appendix show a high level of accuracy, with only minor misclassifications were observed.

The classification results are summarized in the confusion matrix of Figure 6, which demonstrates the model's robust classification capability, providing clear insight into true positives, false positives, and false negatives across the five categories. The model shows strong taxonomic performance, with remarkable accuracy in differentiating similar heart diseases.

Table 3 shows key performance metrics, including accuracy, recall, F1 score, and support for each category. These metrics demonstrate the model's ability to rank, with values as high as 1.00 indicating complete agreement between the expected and true categories, which further measures the effectiveness of the rating for the category. The high accuracy and recall values across all classes confirm the model's ability to identify and classify different types of diseases correctly. Overall assessment accuracy was 98%.

Table 2: Evaluation metrics.

Metric	Description	Formula
Dice Similarity Coefficient (DSC)	Measures overlap between predicted and true regions.	$DSC = 2 X \cap Y / X + Y $
Hausdorff Distance (HD)	Measures the maximum distance between predicted and true boundaries.	$d_H(X, Y) = \max\{d_{XY}, d_{YX}\} = \max\{\max_{x \in X} \min_{y \in Y} d(x, y), \max_{y \in Y} \min_{x \in X} d(x, y)\}$
Accuracy	Ratio of correct pixels (both positives and negatives).	$Accuracy = (TP + TN) / (TP + TN + FP + FN)$
Precision	The proportion of true positives among predicted positives.	$Precision = TP / TP + FP$
F1-Score	Harmonic mean of Precision and Recall.	$F1 = 2 * Precision * Recall / (Precision + Recall)$
Intersection over Union (IoU)	Measures overlap of predicted and true regions divided by their union.	$IOU = A \cap B / A \cup B $

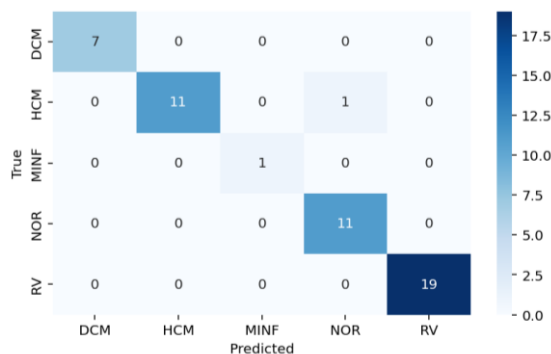


Fig. 6: Confusion matrix.

Table 3: Classification results.

Class	Precision	Recall	F1-score	Support
DCM	1.00	1.00	1.00	1
HCM	0.92	1.00	0.96	11
MINF	1.00	0.92	0.96	12
NOR	1.00	1.00	1.00	7
RV	1.00	1.00	1.00	19

5.2. Segmentation performance

The segmentation performance was assessed using the Dice-like coefficient (DSC) of the three main anatomical structures: the left ventricle (LV), the right ventricle (RV), and the

heart muscle (MYO). As shown in Fig. 7 and Table 4, the proposed model significantly outperformed current approaches, achieving Dice scores of 98.20%, 98.00%, and 98.42% for LV, RV, and MYO, respectively.

The precision of the boundary has been assessed using the Hausdorff (HD) distance, which is 80.16, 76.67, and 78.50 for LV, RV, and MYO, respectively. The model achieved minimum HD values, which indicates high boundary accuracy.

Table 4: Comparison of obtained segmentation results with recent studies in terms of DSC.

Method	Avg. DSC	RV	MYO	LV
R50 UNet [33]	87.55	87.10	80.63	94.92
R50 Att-UNet [34]	86.75	87.58	79.20	93.47
ViT [35]	81.45	81.46	70.71	92.18
R50 ViT [35]	87.57	86.07	81.88	94.75
TransUNet [36]	89.71	88.86	84.53	95.73
Swin UNet [37]	90.00	88.55	85.62	95.83
LeViT-UNet384 [38]	90.32	89.55	87.64	93.76
nnUNet [39]	91.61	90.24	89.24	95.36
nnFormer [40]	91.78	90.22	89.53	95.59
FCT224 w/o D.S. [39]	91.49	90.32	89.00	95.17
FCT224 full D.S. [39]	91.49	90.49	88.76	95.23
FCT224[39]	92.84	92.02	90.61	95.89
FCT384[39]	93.02	92.64	90.51	95.90
Proposed model	98.20	98.00	98.42	98.19

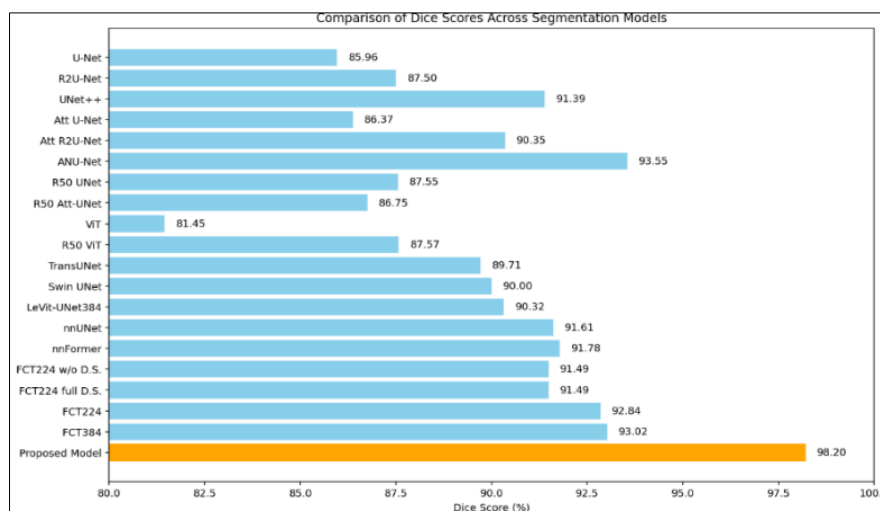


Fig. 7: Dice Similarity Coefficient (DSC) Comparison.

5.3. Comparison with recent models

The proposed model has been compared with the latest current models using dice gauges, accuracy, and recall. As indicated in Table 5, all existing methods of segmentation accuracy and classification reliability have exceeded, confirming the effectiveness of integrating interest and ambiguous aggregation into the Internet's overlapping framework.

Table 5: Comparison with Latest Models.

Method	Dice (%)	Precision	Recall
U-Net [33]	85.96	0.8731	0.8465
R2U-Net [41]	87.50	0.9211	0.8339
UNet++ [42]	91.39	0.9306	0.8979
Attention U-Net [42]	86.37	0.9111	0.8209
Attention R2U-Net [42]	90.35	0.9474	0.8640
ANU-Net [42]	93.55	0.9423	0.9288
Proposed model	98.20	0.9683	0.9683

5.4. Computational efficiency

The computational efficiency of the model has been assessed in terms of training time, evaluation time, memory use, and productivity. As shown in Table 6, the model exhibits high efficiency, with a training time of 1501.45 seconds, an evaluation time of 146.67 seconds, a memory usage of 1414.38 MB, and a throughput of 17.18 (inf/sec). This helps in implementing the model on mobile devices, thereby facilitating its use in real-time clinical applications at a lower cost.

Table 6: Computational Efficiency.

Metric	Value
Training Time	1501.45 seconds
Evaluation Time	146.67 seconds
Testing Time	146.67 seconds
Memory Usage	1414.38MB
Throughput	17.18(inf/sec)

5.5. Summary of results

The U-Net Network, combined with the proposed attention mechanism and Fuzzy Pooling, has achieved superior performance in all aspects evaluated. Exceptional dice scores across major anatomical structures, robust classification metrics, and minimal errors confirm the model's accuracy and reliability. In addition, an effective computational footprint makes it suitable for clinical applications. These results highlight the advantages of combining attention units and a mysterious gathering to segment a strong medical image.

6. DISCUSSION

This study presents a hybrid framework for cardiac MRI segmentation that integrates the overlapping U-Net architecture with spatial attention and fuzzy pooling mechanisms to

enhance segmentation accuracy and diagnostic reliability. The proposed model has shown exceptional performance, achieving average Dice scores of 98.20% for the left ventricle LV, 98.00% for the right ventricle RV and 98.42% for the myocardium MYO, besides low Hausdorff distances LV: 80.16, RV: 76.67, MYO: 78.50. This highlights its accuracy in determining limits. Compared to other advanced models such as nnU-Net, FCT384, and ANU-Net, the proposed method consistently achieved higher segmentation accuracy and classification performance, reaching an overall diagnostic accuracy of 98% across five cardiology categories.

The integration of units of attention allows the model to focus on clinically relevant areas, while fuzzy pooling preserves subtle spatial features that are often lost in traditional clustering layers. In addition, the model achieves efficient computational performance with minimal training and inference times, as well as a low memory footprint, making it suitable for clinical settings where speed and scalability are crucial. Key contributions include the development of an efficient hybrid architecture, a robust segmentation and classification pipeline, and validation of a known cardiac dataset. However, the study was limited to cardiac MRI data, and the classification is based solely on segmentation-derived features, which may overlook other clinical variables.

Future work will focus on integrating multimodal data (e.g., MRI with ECG or clinical records), CT scans (e.g., lung nodule or liver lesion segmentation), and Ultrasound (e.g., fetal or thyroid imaging), using domain adaptation techniques for broader generalization, and integrating interpretable AI methods to enhance clinical interpretability and confidence.

7. CONCLUSION

This paper presented an improved nested U-Net model with integrated spatial attention and fuzzy pooling for segmentation of cardiac MRI images. The model achieved superior performance in dice score (98.20%) and Hausdorff distance, surpassing modern deep learning architectures. The combination of attentional mechanisms and fuzzy pooling allowed enhanced focus on key anatomical regions while maintaining spatial integrity. High computational efficiency also supports real-time clinical applicability. Future work will expand the model to incorporate other methods, such as computed tomography and ultrasound, and investigate explainable AI strategies to enhance clinical confidence.

REFERENCES

- [1] Y. M. A. Mohammed, S. El Garouani, and I. Jellouli, "A survey of methods for brain tumor segmentation-based MRI images," *J. Comput. Des. Eng.*, vol. 10, no. 1, pp. 266–293, 2023, doi: 10.1093/jcde/qwac141.
- [2] A. Kumar, "Study and analysis of different segmentation methods for brain tumor MRI application," *Multimed. Tools Appl.*, vol. 82, no. 5, pp. 7117–7139, 2023, doi: 10.1007/s11042-022-13636-y.
- [3] N. Abdullah and S. Al-Kazzaz, "Evaluation of Physiotherapy Exercise by Motion Capturing Based on Artificial Intelligence: A Review," *Al-Rafidain Engineering Journal (AREJ)*, 28, no. 2, pp. 237–251, 2023, doi: 10.33899/rengj.2023.138562.1235.
- [4] R. Yousef, G. Gupta, N. Yousef, and M. Khari, A holistic overview of deep learning approach in medical imaging, vol. 28, no. 3. Springer Berlin Heidelberg, 2022. doi: 10.1007/s00530-021-00884-5.
- [5] H. Cui, H. Pan, and K. Zhang, "SCU-Net++: A Nested U-Net Based on Sharpening Filter and Channel Attention Mechanism," *Wirel. Commun. Mob. Comput.*, vol. 2022, 2022, doi: 10.1155/2022/2848365.
- [6] R. A. Moyassar and M. A. M. Abdullah, "Advancing EMG Finger Movement Classification with Feature Extraction and Machine Learning," *Al-Rafidain Eng. J.*, vol. 30, no. 1, pp. 150–163, 2025, doi: 10.33899/arej.2024.152539.1368.
- [7] A. Al-Saegh, S. A. Dawwd, and J. M. Abdul-Jabbar, "Towards Efficient Motor Imagery EEG-based BCI Systems using DCNN," 2024 Arab ICT Conf. AICTC 2024, pp. 59–66, 2024, doi: 10.1109/AICTC58357.2024.10735028.
- [8] H. Peng and S. Yu, "A Systematic IoU-Related Method: Beyond Simplified Regression for Better Localization," *IEEE Trans. Image Process.*, vol. 30, no. i, pp. 5032–5044, 2021, doi: 10.1109/TIP.2021.3077144.
- [9] Z. Fu, H. Chen, M. He, and L. Liu, "An Enhanced U-Net with Combining PPM and CBAM for Medical Image Segmentation," *IEEE Access*, vol. 12, no. June, pp. 107098–107112, 2024, doi: 10.1109/ACCESS.2024.3426518.
- [10] S. R. Gunasekara, H. N. T. K. Kaldera, and M. B. Dissanayake, "A systematic approach for MRI brain tumor localization and segmentation using deep learning and active contouring," *J. Healthc. Eng.*, vol. 2021, pp. 1–13, 2021, doi: 10.1155/2021/6695108.
- [11] M. A. Ottom, H. A. Rahman, and I. D. Dinov, "Znet: Deep learning approach for 2D MRI brain tumor segmentation," *IEEE J. Transl. Eng. Health Med.*, vol. 10, no. 1, pp. 1–8, 2022, doi: 10.1109/JTEHM.2022.3176737.
- [12] X. Guan et al., "3D AGSE-VNet: an automatic brain tumor MRI data segmentation framework," *BMC Med. Imaging*, vol. 22, no. 1, pp. 1–18, 2022, doi: 10.1186/s12880-021-00728-8.
- [13] A. Al-Saegh, A. Daood, and M. H. Ismail, "Dual Optimization of Deep CNN for Motor Imagery EEG Tasks Classification," *Diyala J. Eng. Sci.*, vol. 17, no. 4, pp. 75–91, 2024, doi: 10.24237/djes.2024.17405.
- [14] B. Chen, Y. Liu, Z. Zhang, G. Lu, and A. W. K. Kong, "TransAttUnet: Multi-Level Attention-Guided U-Net with Transformer for Medical Image Segmentation," *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 8, no. 1, pp. 55–68, 2024, doi: 10.1109/TETCI.2023.3309626.
- [15] S. Sinha, M. Bhatt, A. Anand, A. S. Areckal, and V. Aparna, "EnigmaNet: A Novel Attention-Enhanced Segmentation Framework for Ischemic Stroke Lesion Detection in Brain MRI," *IEEE Access*, vol. 12, no. June, pp. 91480–91498, 2024, doi: 10.1109/ACCESS.2024.3422025.
- [16] G. Han et al., "Improved U-Net based insulator image segmentation method based on attention mechanism," *Energy Reports*, vol. 7, pp. 210–217, 2021, doi: 10.1016/j.egy.2021.10.037.
- [17] L. Shen, Q. Wang, W. Wang, R. Wang, Y. Yu, and C. Zhang, "Deep attention enhanced networks for medical image segmentation," vol. 13167, no. Rsmip, p. 126, 2024, doi: 10.1117/12.3029812.
- [18] Y. Zhou, Q. Kong, Y. Zhu, and Z. Su, "MCFA-UNet: Multiscale Cascaded Feature Attention U-Net for Liver Segmentation," *Irbm*, vol. 44, no. 4, p. 100789, 2023, doi: 10.1016/j.irbm.2023.100789.
- [19] M. T. Rostami and A. Motamedi, "IRA-Unet: Inception Residual Attention U-Net in adversarial network for cardiac MRI segmentation," *TechRxiv*, preprint, pp. 1–10, Aug. 2023, doi: 10.36227/techrxiv.23896641.
- [20] S. Kuang et al., "MSCDA: Multi-level semantic-guided contrast improves unsupervised domain adaptation for breast MRI segmentation in small datasets," *Neural Networks*, vol. 165, pp. 119–134, 2023, doi: 10.1016/j.neunet.2023.05.014.
- [21] F. Zhang et al., "Deep learning based segmentation of brain tissue from diffusion MRI," *Neuroimage*, vol. 233, no. July 2020, p. 117934, 2021, doi: 10.1016/j.neuroimage.2021.117934.
- [22] Y. Gu, Z. Piao, and S. J. Yoo, "STHardNet: Swin Transformer with HardNet for MRI Segmentation," *Appl. Sci.*, vol. 12, no. 1, 2022, doi: 10.3390/app12010468.
- [23] C. Zhang, A. Achuthan, and G. M. S. Himel, "State-of-the-Art and Challenges in Pancreatic CT Segmentation: A Systematic Review of U-Net and Its Variants," *IEEE Access*, vol. 12, no. June, pp. 78726–78742, 2024, doi: 10.1109/ACCESS.2024.3392595.
- [24] H. Wang, S. Qiu, B. Zhang, and L. Xiao, "Multilevel Attention Unet Segmentation Algorithm for Lung Cancer Based on CT Images," *Comput. Mater. Contin.*, vol. 78, no. 2, pp. 1569–1589, 2024, doi: 10.32604/cmc.2023.046821.
- [25] K. Wisaeng, "U-Net++DSM: Improved U-Net++ for Brain Tumor Segmentation With Deep Supervision Mechanism," *IEEE Access*, vol. 11, no. October, pp. 132268–132285, 2023, doi: 10.1109/ACCESS.2023.3331025.
- [26] T. Zhou, X. Ye, H. Lu, X. Zheng, S. Qiu, and Y. Liu, "Dense convolutional network and its application in medical image analysis," *Biomed. Res. Int.*, vol. 2022, pp. 1–11, 2022, doi:

- 10.1155/2022/2384830.
- [27] Z. Mahmoud and M. A. M. Abdullah, "Comparison of Deep Learning and Machine Learning Techniques for Automated Diagnosis of Acute Lymphoblastic Leukemia," *Al-Rafidain Eng. J.*, vol. 30, no. 1, pp. 175–185, 2025, doi: 10.33899/arej.2024.152478.1366.
- [28] Z. Li, H. Zhang, Z. Li, and Z. Ren, "Residual-attention UNet++: A nested residual-attention U-Net for medical image segmentation," *Appl. Sci.*, vol. 12, no. 14, pp. 1–14, 2022, doi: 10.3390/app12147149.
- [29] Z. Al-Mokhtar, F. Ibraheem, and H. Al-Layla, "A Review of Digital Image Fusion and its Application," *Al-Rafidain Eng. J.*, vol. 26, no. 2, pp. 309–322, 2021, doi: 10.33899/rengj.2021.127928.1055.
- [30] M. Cui, K. Li, J. Chen, and W. Yu, "CM-Unet: A Novel Remote Sensing Image Segmentation Method Based on Improved U-Net," *IEEE Access*, vol. 11, no. June, pp. 56994–57005, 2023, doi: 10.1109/ACCESS.2023.3282778.
- [31] A. Al-Saegh, "Identifying a Suitable Signal Processing Technique for MI EEG Data," *Tikrit J. Eng. Sci.*, vol. 30, no. 3, pp. 140–147, 2023, doi: 10.25130/tjes.30.3.14.
- [32] N. S. Punni and S. Agarwal, "Modality specific U-Net variants for biomedical image segmentation: A survey," *Artif. Intell. Rev.*, vol. 55, no. 7, pp. 5845–5889, 2022, doi: 10.1007/s10462-022-10152-1.
- [33] M. Ali, S. O. Gilani, A. Waris, K. Zafar, and M. Jamil, "Brain Tumour Image Segmentation Using Deep Networks," *IEEE Access*, vol. 8, pp. 153589–153598, 2020, doi: 10.1109/ACCESS.2020.3018160.
- [34] O. Oktay et al., "Attention U-Net: Learning where to look for the pancreas," *arXiv preprint*, 2018. [Online]. Available: <http://arxiv.org/abs/1804.03999>
- [35] A. Dosovitskiy et al., "An image is worth 16×16 words: Transformers for image recognition at scale," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2021. [Online]. Available: <http://arxiv.org/abs/2010.11929>
- [36] J. Chen et al., "TransUNet: Transformers make strong encoders for medical image segmentation," *arXiv preprint*, pp. 1–13, 2021. [Online]. Available: <http://arxiv.org/abs/2102.04306>
- [37] H. Cao et al., "Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13803 LNCS, pp. 205–218, 2023, doi: 10.1007/978-3-031-25066-8_9.
- [38] G. Xu, X. Zhang, X. He, and X. Wu, "LeViT-UNet: Make Faster Encoders with Transformer for Medical Image Segmentation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 14432 LNCS, pp. 42–53, 2024, doi: 10.1007/978-981-99-8543-2_4.
- [39] A. Tragakis, C. Kaul, R. Murray-Smith, and D. Husmeier, "The Fully Convolutional Transformer for Medical Image Segmentation," *Proc. - 2023 IEEE Winter Conf. Appl. Comput. Vision, WACV 2023*, pp. 3649–3658, 2023, doi: 10.1109/WACV56688.2023.00365.
- [40] H.-Y. Zhou, J. Guo, Y. Zhang, L. Yu, L. Wang, and Y. Yu, "nnFormer: Interleaved Transformer for volumetric segmentation," *arXiv preprint*, pp. 1–10, 2021. [Online]. Available: <http://arxiv.org/abs/2109.03201>
- [41] M. Z. Alom, C. Yakopcic, M. Hasan, T. M. Taha, and V. K. Asari, "Recurrent residual U-Net for medical image segmentation," *Journal of Medical Imaging*, vol. 6, no. 1, pp. 014006-1–014006-13, 2019, doi: 10.1117/1.jmi.6.1.014006.
- [42] C. Li et al., "ANU-Net: Attention-based nested U-Net to exploit full resolution features for medical image segmentation," *Comput. Graph.*, vol. 90, pp. 11–20, 2020, doi: 10.1016/j.cag.2020.05.003.

APPENDIX A

Table A1: Patient Classification Results

Patient	True Class	Predicted Class	Patient	True Class	Predicted Class
patient101	RV	RV	patient126	RV	RV
patient102	NOR	NOR	patient127	HCM	HCM
patient103	NOR	NOR	patient128	MINF	MINF
patient104	HCM	HCM	patient129	HCM	HCM
patient105	MINF	HCM	patient130	NOR	NOR
patient106	RV	RV	patient131	RV	RV
patient107	DCM	DCM	patient132	RV	RV
patient108	RV	RV	patient133	RV	RV
patient109	RV	RV	patient134	NOR	NOR
patient110	MINF	MINF	patient135	HCM	HCM
patient111	HCM	HCM	patient136	RV	RV
patient112	MINF	MINF	patient137	NOR	NOR
patient113	DCM	DCM	patient138	RV	RV
patient114	MINF	MINF	patient139	RV	RV
patient115	RV	RV	patient140	DCM	DCM
patient116	HCM	HCM	patient141	HCM	HCM
patient117	MINF	MINF	patient142	HCM	HCM
patient118	DCM	DCM	patient143	HCM	HCM
patient119	NOR	NOR	patient144	RV	RV
patient120	MINF	MINF	patient145	DCM	DCM
patient121	RV	RV	patient146	RV	RV
patient122	HCM	HCM	patient147	NOR	NOR
patient123	RV	RV	patient148	MINF	MINF
patient124	MINF	MINF	patient149	HCM	HCM
patient125	RV	RV	patient150	HCM	HCM

تقسيم صور الرنين المغناطيسي الطبي باستخدام شبكة U المتداخلة مع آلية الانتباه والتجميع المضرب

علي الصائغ
ali.alsaegh@uomosul.edu.iq

نور محمد بشر
noor.23enp115@student.uomosul.edu.iq

قسم هندسة الحاسوب، كلية الهندسة، جامعة الموصل، الموصل، العراق

تاريخ القبول: 28 يوليو 2025

استلم بصيغته المنقحة: 27 يونيو 2025

تاريخ الاستلام: 22 ابريل 2025

الملخص

بعد تجزئة الصور الطبية خطوة حاسمة في تشخيص وعلاج الأمراض المختلفة، لا سيما عند تحليل فحوصات التصوير بالرنين المغناطيسي (MRI). على الرغم من التقدم الكبير، لا يزال التجزئة الدقيقة تمثل تحديًا بسبب تعقيد الهياكل التشريحية والحدود الغامضة والاختلافات في جودة الصورة. تقترح هذه الورقة هيكلًا محسنًا لـ Nested U-Net يتضمن آلية انتباه وتجميعًا غامضًا لتحسين أداء البيع بالتجزئة. تستفيد U-Net المتداخلة من اتصالات التخطي الواسعة لالتقاط المعلومات السياقية متعددة المستويات، بينما تعزز آلية الانتباه قدرة النموذج على التركيز على الميزات ذات الصلة وقمع الضوضاء. بالإضافة إلى ذلك، نستبدل طبقات التجميع المتطرفة التقليدية بتجميع غامض، مما يمكن الشبكة من التعامل مع الغموض المكاني بشكل أكثر فعالية وإنتاج خرائط بيع بالتجزئة أكثر دقة وصقلًا. تظهر النتائج التجريبية على مجموعات بيانات التصوير بالرنين المغناطيسي المتاحة للجمهور أن النموذج المقترح يحقق تحسينات ملحوظة، مع زيادة بنسبة 4.97٪ في معامل النرد، وتحسن بنسبة 2.76٪ في الدقة، وزيادة بنسبة 4.25٪ في الاستدعاء مقارنة بهيكل U-Net القياسية. علاوة على ذلك، لوحظت تحسينات كبيرة في مقاييس المسافة في Housdorf.

الكلمات الدالة:

آلية الانتباه، التعلم العميق في التصوير الطبي، التجمع الغامض، شبكة U-Net المتداخلة، تجزئة صور التصوير بالرنين المغناطيسي، التجزئة الدلالي.