INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING
ISSN 2180-4370
Journal Homepage : https://ijic.utm.my/

A Sigmoid-Based Multi-Label U-Net with Weighted Tversky Loss for Cardiac Fat Segmentation in CT Imaging

Nur Atiqah Mohd Fuaad[1*] & Rohayanti Hassan[2]
Faculty of Computing
Universiti Teknologi Malaysia
81310 UTM Johor Bahru, Johor, Malaysia
Email: natiqah286@graduate.utm.my[1*], rohayanti@utm.my[2]

Submitted: 30/9/2025. Revised edition: 14/4/2026. Accepted: 4/5/2026. Published online: 10/6/2026
DOI: https://doi.org/10.11113/ijic.v16n1.674

***Abstract*—Cardiovascular diseases remain among the leading causes of mortality worldwide. Cardiac fat, particularly epicardial adipose tissue and mediastinal adipose tissue, has increasingly been recognized as a key indicator in the early detection of cardiovascular diseases. However, this detection remains challenging due to the thin and often indistinct pericardial membrane that separates epicardial and mediastinal in non-contrast computed tomography (CT) images. Accurate segmentation of these fat regions and the pericardial membrane in CT imaging is essential to support clinical risk assessment and timely medical interventions. Traditional manual segmentation techniques have been employed in the past, but they are time-consuming and prone to inter-observer variability. Meanwhile, conventional machine learning algorithms often struggle to capture the complex spatial information inherent in tomographic CT images. To address these challenges, this study proposes an improved U-Net model designed to enhance the accuracy and adaptability of coronary fat segmentation in non-contrast CT scans. The model specifically addresses issues such as class imbalance, ambiguous labeling, and limited generalizability by incorporating a sigmoid activation function, which enables binary predictions for each anatomical structure and allows the model to effectively handle overlapping regions in multi-class segmentation tasks. The proposed model is evaluated on an annotated CT dataset using standard segmentation performance metrics. The results demonstrate that the advanced U-Net architecture developed in this study achieves improved accuracy in cardiac fat segmentation, thereby offering potential to enhance clinical diagnosis and support better patient treatment in medical imaging.**

***Keywords*—Cardiac Fat Segmentation, Computed tomography (CT), Multi-Label U-Net**

I. INTRODUCTION

Cardiovascular diseases (CVDs) remain a leading cause of mortality worldwide [1], with epicardial and mediastinal adipose tissues recognized as critical key indicator due to their associations with inflammation, metabolism, and coronary artery disease [2]. Accurate segmentation of these cardiac fat regions in CT imaging is essential for reliable risk assessment and diagnosis. However, manual segmentation is time-intensive, inconsistent, and unsuitable for large-scale studies, creating a strong demand for robust automated methods [3].

Deep learning model such as U-Net has shown remarkable success in medical image segmentation due to its encoder–decoder structure and skip connections, which preserve fine-grained anatomical details [4, 5]. The conventional U-Net models encounter significant limitations in cardiac fat segmentation such as (i) class imbalance, where underrepresented fat tissues such as pericardium are poorly segmented or entirely missed; (ii) overlapping and ambiguous anatomical boundaries, which lead to blurry predictions and misclassifications; and (iii) insufficient sensitivity of standard loss functions which underrepresent the impact of false positive and false negative errors in evaluation resulting in weak generalization across diverse fat regions. To address these challenges, this study proposes a Sigmoid-Based Multi-Label U-Net architecture designed for non-contrast CT images which introduces sigmoid activation for multi-label segmentation that allow independent probability estimation for overlapping structures and integrated with the weighted Tversky loss

function to handle class imbalance by proportionally balancing both false negatives and false positives [6].

The objectives of this study are:

(a) To develop a sigmoid-based multi-label U-Net for simultaneous segmentation of epicardial, mediastinal, and pericardial fat.
(b) To optimize the weighted Tversky loss with class weighting for improving performance in minority fat regions.
(c) To evaluate and compare the proposed model's performance against the baseline U-Net using standard metrics such as Dice similarity coefficient, precision, recall, and F1-score.

This paper is organized as follows which start with Section 1 that serves as the introduction for this research study. Section 2 reviews related work on cardiac fat segmentation and deep learning-based medical image analysis. Section 3 describes the methodology, dataset, and the proposed U-Net model with sigmoid activation. Section 4 discuss the results and comparative analysis with baseline model and future works. Section 5 reports the conclusion of this study.

II. LITERATURE REVIEW

Cardiac adipose tissue is an important marker in cardiovascular health and is mainly classified into two regions: epicardial fat and mediastinal fat [7]. Epicardial fat lies between the myocardium and visceral pericardium, in direct contact with the heart and sharing its microcirculation, which makes it metabolically active and strongly linked to inflammation and coronary artery disease. Mediastinal fat, on the other hand, is located outside the parietal pericardium within the mediastinal space, separated from the heart by the fibrous pericardium, yet it remains clinically relevant due to its association with systemic inflammation and cardiovascular risk factors. The pericardium which consists of an outer parietal layer and an inner visceral layer also known as epicardium which separates these two fat compartments and clear terminology distinguishing them is essential for accurate cardiovascular imaging and research [8, 9]. Advancements in deep learning have greatly improved cardiac fat segmentation by replacing traditional techniques such as thresholding and region-growing with convolutional neural networks such as U-Net. The U-Net's encoder–decoder design with skip connections enables learning of complex spatial features while retaining contextual information making it the foundation of many state-of-the-art methods [5].

A. U-Net Architecture and its Variants in Cardiac Fat Segmentation

U-Net is a widely adopted architecture in biomedical image segmentation due to its encoder–decoder design with skip connections which preserves spatial details and enables accurate pixel-level classification even with limited training data [5]. In cardiac imaging, U-Net has proven effective for segmenting epicardial and mediastinal adipose tissues in non-contrast CT scans, where subtle boundaries such as the pericardium are usually challenging to outlined. Its strength lies in capturing both local and global image features allowing reliable segmentation in complex or low-contrast regions.

TABLE I. U-NET VARIANTS AND THEIR KEY FEATURES IN CARDIAC FAT SEGMENTATION

U-Net Variants	Key Features
2D U-Net [11]	Processes individual 2D CT slices with encoder-decoder with skip connections efficient for feature extraction.
3D U-Net with Attention Gates [10]	Extends U-Net to volumetric input enabling captures spatial continuity across slices with attention gates suppress irrelevant activations.
HU-Attention U-Net (DeepFat) [12]	Integrates Hounsfield Unit (HU)-based attention that highlights fat tissues within specific HU ranges which improves discrimination of adipose tissue.
PoinUNet and Hyperbolic U-Net [13]	Combines multi-scale feature extraction with attention layers; enhances recognition of fine details and overlapping regions at different resolutions.

Several U-Net variants have been explored in the field of cardiac fat segmentation to overcome the limitations of the standard architecture. The classical 2D U-Net [11] remains widely used due to its encoder–decoder structure with skip connections that enables efficient extraction of features from individual CT slices. The 3D U-Net with attention gates which capture better volumetric information extends the framework to process full volumetric inputs that allow spatial continuity across slices while using attention gates to suppress irrelevant activations and focus on critical structures [10]. Building upon domain-specific properties, the HU-Attention U-Net (DeepFat) [12] incorporates Hounsfield Unit (HU)-based attention, which highlights fat tissues within specific HU ranges, thereby improving the model's ability to discriminate adipose tissue from neighboring structures. More recent developments such as PoinU-Net and Hyperbolic U-Nets [13] introduce non-Euclidean (hyperbolic) embeddings which enable the network to better capture hierarchical and complex anatomical geometries while preserving fine tissue boundaries. Finally, multi-scale attention U-Nets [14] combine hierarchical feature extraction with attention mechanisms that enhance the recognition of fine details and overlapping fat regions across different spatial resolutions. Overall, these architectural innovations contributed the segmentation performance of U-Net models in cardiac imaging applications.

B. U-Net Architecture Limitations

Despite the advancements introduced by different U-Net variants, several limitations remain in their application to cardiac fat segmentation. The conventional 2D U-Net [11] is constrained by its inability to capture volumetric context, which often leads to discontinuities between slices and difficulty in modeling 3D anatomical relationships. Extending to volumetric data, the 3D U-Net with attention gates [10] improves spatial continuity but introduces high computational and memory demands, and its performance can be unstable on smaller datasets, making it prone to overfitting. The HU-Attention U-Net (DeepFat) [12] leverages Hounsfield Unit information to highlight adipose tissues, yet its reliance on HU

values makes it sensitive to scanner variability and imaging protocols, which may compromise robustness in noisy or low-quality scans. More advanced frameworks such as PoinU-Net and hyperbolic U-Nets [13] offer improved boundary preservation using non-Euclidean embeddings, but they remain relatively new, with limited validation in clinical practice, and involve significant mathematical and implementation complexity. Finally, multi-scale attention U-Nets [14] enhance recognition of small and overlapping regions through hierarchical features, but their increased architectural complexity leads to longer training times, redundancy in extracted features, and a greater dependency on large annotated datasets for stability.

III. METHODOLOGY

This section presents the workflow adopted for cardiac fat segmentation using the proposed Sigmoid Based Multi-Label Cardiac U-Net Architecture model. The process consists of four main phases as shown in Fig. 1 below: (i) data preparation and preprocessing, (ii) model design and development, (iii) model optimization and fine-tuning, and (iv) result analysis and evaluation. Each phase is described in detail, including dataset characteristics, preprocessing strategies, model architecture, training setup, and evaluation metrics.

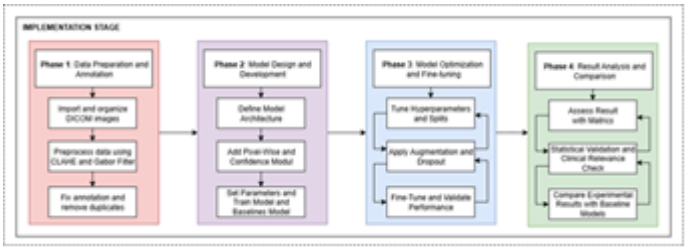

Fig. 1. Implementation framework.

A. Data Preparation and Annotation

The dataset used in this study was obtained from the Visual Computing Group, Universidade Federal Fluminense, originally introduced by Rodrigues *et al*. (2016) [9] and Kazemi *et al*. (2019) [15] for cardiac fat segmentation. It has since been cited in multiple works and was recently employed in Medical Engineering & Physics (2024) [16], reflecting continued academic relevance. The dataset comprises non-contrast cardiac CT scans from 20 patients totalling approximately 1000 image slices. For this study, only the fat images and ground-truth segmentations in the range of [−200, −30] HU were used. While the dataset size is relatively limited, this is a common constraint in medical imaging due to the complexity and cost of obtaining expert annotations and highlights the capability of the proposed model to learn effectively from limited data while still achieving better segmentation performance. For preprocess CLAHE is used to enhance image contrast by improving local details while limiting noise amplification, making important structures in CT images more visible. The Gabor filter is applied to extract texture and edge information by detecting specific frequency and orientation patterns. Together, these techniques improve feature quality and support more accurate segmentation. The final step involves correcting annotation errors to ensure accurate and consistent labelling and duplicate images are removed to prevent redundancy and improve the quality of model training.

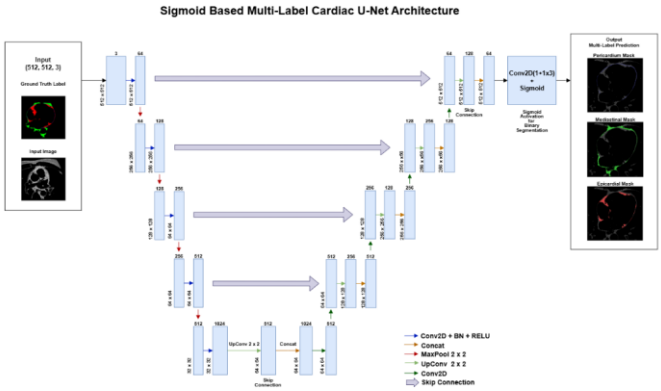

Fig. 2. Sigmoid based multi-label cardiac U-Net architecture

This study proposes a Sigmoid-Based Multi-Label Cardiac U-Net Architecture designed for simultaneous segmentation of multiple cardiac fat tissues from CT images as shown in Fig. 2 above. The model adopts a symmetric encoder–decoder structure consisting of a contracting path (encoder), a bottleneck, and an expanding path (decoder). The encoder is composed of four convolutional blocks, each containing Conv2D + Batch Normalization + ReLU units followed by 2×2 MaxPooling, progressively reducing spatial dimensions from 512×512 to 32×32 while increasing filters from 64 to 1024 to capture hierarchical features. the purple arrows denote skip connections that link corresponding layers in the encoder and decoder paths. Thes10e connections transfer feature maps directly from the contracting path to the expanding path at the same spatial resolution. Unlike the main forward flow, which progressively downsamples and abstracts features, skip connections preserve low-level spatial information such as edges, boundaries, and fine anatomical structures. At the bottleneck, the network learns high-level abstract representations with 1024 filters, serving as the semantic bridge between encoder and decoder. The decoder path mirrors the encoder, employing transpose convolutions for upsampling and concatenating feature maps from the corresponding encoder blocks via skip connections to recover fine-grained spatial details. This process restores the feature maps back to 512×512 resolution while gradually reducing the number of filters. The architecture enables pixel-level predictions across multiple classes by applying a sigmoid activation at the output layer, allowing overlapping regions to be captured effectively.

B. Model Optimization and Fine Tuning

This stage focuses on improving the robustness and generalization of the proposed network through a series of optimization strategies. Hyperparameter tuning was conducted to determine suitable values for learning rate, batch size and number of epochs to stable convergence without overfitting.

Dataset splitting strategies were carefully designed with images partitioned into training, validation, and testing sets to achieve a balanced evaluation across subjects and avoid data leakage. To further enhance generalization, data augmentation techniques such as rotation, flipping, scaling and elastic deformation were applied enabling increasing variability and reducing the risk of model bias. Dropout layers were incorporated within the network to prevent overfitting by randomly deactivating neurons during training.

To address the inherent class imbalance in cardiac fat segmentation, the Tversky loss function was employed [6]. This loss provides flexibility by weighting false positives and false negatives asymmetrically making it particularly effective in segmenting small or overlapping fat regions. For the output layer, a sigmoid activation function was applied enabling multi-label predictions and allowing pixels to belong simultaneously to more than one anatomical class. This is especially relevant in regions where epicardial, mediastinal, and pericardial fat overlap. Finally, the model was fine-tuned iteratively using the validation set, where training parameters were adjusted based on feedback from evaluation metrics to progressively improve segmentation performance.

C. *Result Analysis and Comparison*

The final stage evaluates segmentation performance using a set of benchmark metrics. A confusion matrix was constructed to derive key indicators, including Dice similarity score, precision, accuracy, and recall sensitivity which collectively measure overlap quality and classification reliability [17–21]. To better address class imbalance and overlapping anatomical regions, the Tversky coefficient was employed, while cross-entropy loss was monitored to quantify pixel-wise prediction error during training. The performance of the proposed sigmoid-based multi-label U-Net was then compared against the baseline U-Net, demonstrating improved segmentation accuracy, robustness, and consistency with clinical annotations.

IV. RESULT & DISCUSSION

This section presents the result and analysis through the evaluation phase of the proposed sigmoid-based multi-label U-Net, comparing its performance to the baseline SoftMax-based U-Net for cardiac fat segmentation.

The evaluation metrics used are precision, recall, F1-score, and support. Prediction accuracy is measured by precision, which is the percentage of accurately predicted positive instances among all predicted positives and the model's ability to identify a class is shown in recall, which is the percentage of properly identified positive cases out of all actual positives, the F1-score offers a fair assessment of both metrics since it is the harmonic mean of precision and recall and lastly support shows the distribution of classes in the dataset by providing the total number of ground truth instances (pixels) for each class.

A. *Performance Result & Analysis of Baseline U-Net Model*

The performance result of the baseline U-Net model based on SoftMax-activated output for multi-class semantic segmentation were tabulated in Table II below where the results show that the model achieved a high overall Dice Coefficient of 0.97 and a weighted average F1-score of 0.76. While the Dice Coefficient remains high due to the background class that makes up most of the data and has a big impact on the total score, the weighted average reflects the challenges the model faced with the smaller cardiac fat classes.

TABLE II. PERFORMANCE RESULT & ANALYSIS OF BASELINE U-NET

Metric	Precision	Recall	F1-Score	Support
Background	0.99	1.00	0.99	9445408
Epicardial	0.90	0.33	0.49	238654
Mediastinal	0.75	0.68	0.71	249900
Pericardial	0.37	0.00	0.01	27510
Dice Coefficient	N/A	N/A	0.97	N/A
Macro Average	0.75	0.50	0.55	9961472
Weighted Average	0.87	0.74	0.76	19922944

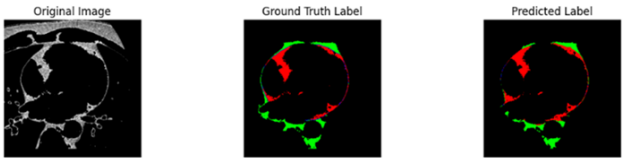

Fig. 3. Result summary of baseline U-Net model

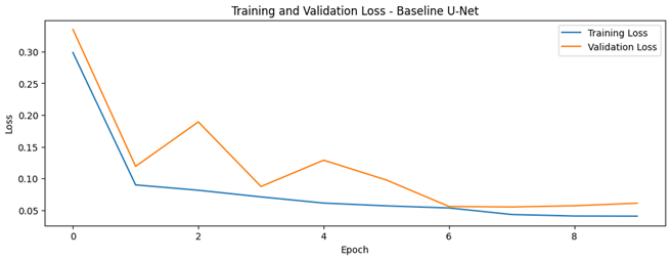

Fig. 4. Training and validation loss of baseline U-Net model

The result summary of the image and the training and validation loss were illustrated in Fig. 3 and Fig. 4 respectively. Among the cardiac fat classes, the model performed moderately on mediastinal fat with an F1-score of 0.71, followed by epicardial fat with an F1-score of 0.49. However, pericardial fat (Pericardium) was significantly under-segmented, with a recall of 0.00 and an F1-score of only 0.01, which shows the model's difficulty in detecting smaller, less distinct anatomical structures. The difference in performance between classes is also shown by the macro-average F1-score of 0.55, which underscores the imbalanced learning typical of SoftMax-based segmentation models. The main weakness of this method is that it forces each area to belong to only one class which makes it hard for the model to identify overlapping regions, hence the areas where different fat types are close together or mixed such as epicardial and mediastinal fat often overlap are labelled incorrectly or missed completely.

B. Performance Result & Analysis of Sigmoid-Based Multi-Label U-Net Model

The performance results of the Sigmoid-Based Multi-Label U-Net Model were tabulated in Table III below which show the model demonstrates very high accuracy in segmenting the Background class, achieving a precision of 0.95, recall of 1.00, and an F1-score of 0.97.

TABLE III. PERFORMANCE RESULT & ANALYSIS OF SIGMOID-BASED MULTI-LABEL U-NET

Metric	Precision	Recall	F1-Score	Support
Background	0.95	1.00	0.97	9445408
Epicardial	0.57	0.95	0.71	238591
Mediastinal	0.51	0.95	0.66	246289
Pericardial	0.19	0.82	0.31	27184
Dice Coefficient	N/A	N/A	0.97	N/A
Macro Average	0.55	0.93	0.66	9957472
Weighted Average	0.74	0.96	0.81	19914944

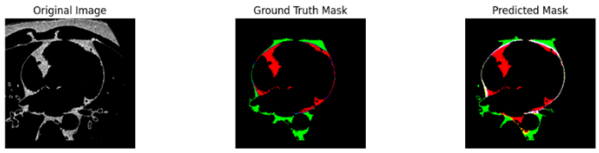

Fig. 5. Result summary of sigmoid-based multi-label U-Net.

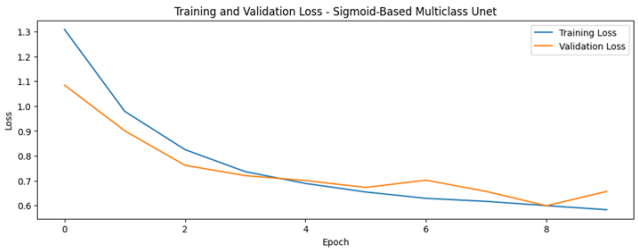

Fig. 6. Result summary of sigmoid-based multi-label U-Net.

TABLE IV. PERFORMANCE RESULT & ANALYSIS OF SIGMOID-BASED MULTI-LABEL U-NET

Metric	Precision	Recall	F1-Score
Dice Coefficient F1	0.97	0.97	Equal
Macro Average F1	0.55	0.66	Sigmoid U-Net
Weighted Average F1	0.76	0.81	Sigmoid U-Net
Background F1	0.99	0.97	Baseline U-Net
Epicardial F1	0.49	0.71	Sigmoid U-Net
Mediastinal F1	0.71	0.66	Baseline U-Net
Pericardial F1	0.01	0.31	Sigmoid U-Net

Based on performance analysis shown in Table IV the Sigmoid-Based Multi-Label Cardiac U-Net achieves higher overall segmentation performance compared to Baseline U-Net. Even though both models obtain a high Dice coefficient of 0.97 which demonstrates good overlap with the ground truth, the sigmoid approach delivers better class balance, improving the macro-average F1-score from 0.55 to 0.66 and the weighted F1-score from 0.76 to 0.81. Class-level analysis further highlights this advantage: epicardial fat performance improves from 0.49 to 0.71, and pericardial fat (Pericardium) rises from 0.01 to 0.31, demonstrating the model's enhanced ability to capture smaller regions. Mediastinal fat performance remains comparable, with the baseline retaining a slight lead (0.71 vs. 0.66), representing a minor trade-off to achieve better overall class balance.

The performance from all the tables highlights the F1-score as a principal evaluation metric because of its appropriateness for imbalanced segmentation tasks. In cardiac fat segmentation, class imbalance and overlapping regions hinder the efficacy of utilizing precision or recall independently. The F1-score, characterized as the harmonic mean of precision and recall, offers a fair assessment by concurrently considering false positives and false negatives. This is especially crucial for precisely identifying smaller and underrepresented adipose areas, where detection sensitivity and predictive accuracy are paramount. Consequently, prioritizing the F1-score provides a more thorough and dependable evaluation of the model's segmentation efficacy. Overall, the Sigmoid-Based Multi-Label Cardiac U-Net demonstrates superior performance in handling multiple cardiac fat classes simultaneously.

V. CONCLUSION

This study proposed a Sigmoid-Based Multi-Label U-Net to improve automated cardiac fat segmentation in non-contrast CT images, addressing the limitations of the conventional U-Net in handling class imbalance, overlapping structures, and poor detection of underrepresented fat tissues. By incorporating sigmoid activation for independent multi-label predictions, along with a weighted Tversky loss and class weighting, the model achieved significant improvements in precision, recall, and F1-scores, particularly for pericardial and epicardial fat where the baseline U-Net performed poorly. These results demonstrate the effectiveness and clinical potential of the proposed approach in providing accurate, consistent, and automated quantification of cardiac adipose tissue, thereby reducing observer variability and supporting more reliable cardiovascular risk assessment and diagnosis.

ACKNOWLEDGMENT

This work was supported by Universiti Teknologi Malaysia (UTM) under the UTM Hi-Tech (F4) Grant, Grant No. Q.J130000.4628.00Q48.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

[1] Gaidai, O., Cao, Y., Loginov, S. (2023). Global cardiovascular diseases death rate prediction. *Current Problems in Cardiology*, *48*(5), 101622. https://doi.org/10.1016/j.cpcardiol.2023.101622.

[2] Benčević, M., Galić, I., Habijan, M., & Pižurica, A. (2022). Recent progress in epicardial and pericardial adipose tissue segmentation and quantification based on deep learning: a systematic review. *Applied Sciences*, *12*(10), 5217. https://doi.org/10.3390/app12105217.

[3] Kaushal, A., Gupta, A.K., Sehgal, V.K. (2024). A semantic segmentation framework with UNet-pyramid for landslide prediction using remote sensing data. *Scientific Reports. 14*(1), 30071. https://doi.org/10.1038/s41598-024-79266-6.

[4] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*. Cham: Springer international publishing. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.

[5] Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A.W.M., van Ginneken, B., Sánchez, C.I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*. *42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005.

[6] Abraham, N., Khan, N. (2023). A novel Focal Tversky Loss Function with improved attention U-Net for lesion segmentation. In *2019 IEEE 16th international symposium on biomedical imaging* (ISBI 2019) *IEEE.* 683–687. https://doi.org/10.32920/22734398.v1.

[7] Rebelo, A.F., Ferreira, A.M., Fonseca, J.M. (2022). Automatic epicardial fat segmentation and volume quantification on non-contrast cardiac computed tomography. *Computer Methods and Programs in Biomedicine Update*. *2*, 100079. https://doi.org/10.1016/j.cmpbup.2022.100079.

[8] Rodrigues, É., Cordeiro de Morais, F., Conci, A. (2015). On the automated segmentation of epicardial and mediastinal cardiac adipose tissues using classification algorithms. *MEDINFO 2015: eHealth-enabled Health.* IOS Press. https://doi.org/10.3233/978-1-61499-564-7-726.

[9] Rodrigues, É.O., Morais, F.F.C., Morais, N.A.O.S., Conci, L.S., Neto, L.V., Conci, A. (2016). A novel approach for the automated segmentation and volume quantification of cardiac fats on computed tomography. *Computer Methods and Programs in Biomedicine*, *123*, 109–128. https://doi.org/10.1016/j.cmpb.2015.09.017.

[10] Liu, L., Ma, R., van Ooijen, P.M.A., Oudkerk, M., Vliegenthart, R., Veldhuis, R.N.J., Brune, C. (2023). The U-Net family for epicardial adipose tissue segmentation and quantification in low-dose CT. *Technologies*, *11*(4), 104. https://doi.org/10.3390/technologies11040104.

[11] Su, R., Zhang, D., Liu, J., Cheng, C. (2021). MSU-net: Multi-Scale U-net for 2d medical image segmentation. *Frontiers in Genetics*, *12*, 639930.

[12] Hoori, A., Hu, T., Lee, J., Al-Kindi, S., Rajagopalan, S., Wilson, D.L. (2022). Deep learning segmentation and quantification method for assessing epicardial adipose tissue in CT calcium score scans. *Scientific Reports*, *12*(1), 2276. https://doi.org/10.1038/s41598-022-06351-z.

[13] Firouznia, M., Ylipää, E., Henningsson, M., Carlhäll, C.-J. (2025). Poincare guided geometric UNet for left atrial epicardial adipose tissue segmentation in Dixon MRI images. *Scientific Reports*, *15*(1), 25549. https://doi.org/10.1038/s41598-025-10110-1.

[14] Liu, L., Wu, K., Wang, K., Han, Z., Qiu, J., Zhan, Q., Wu, T., Xu, J., Zeng, Z. (2024). SEU 2-Net: multi-scale U 2 -Net with SE attention mechanism for liver occupying lesion CT image segmentation. *PeerJ Computer Science*, *10*, e1751. https://doi.org/10.7717/peerj-cs.1751.

[15] Kazemi, A., Keshtkar, A., Rashidi, S., Aslanabadi, N., Khodadad, B., Esmaeili, M. (2019). Segmentation of cardiac epicardial and pericardial fats by using Gabor Filter Bank Based GLCM. In: *2019 26th National and 4th International Iranian Conference on Biomedical Engineering* (ICBME). *IEEE.* 177–182. https://doi.org/10.1109/ICBME49163.2019.9030380.

[16] Santos da Silva, G., Casanova, D., Oliva, J.T., Rodrigues, E.O. (2024). Cardiac fat segmentation using computed tomography and an image-to-image conditional generative adversarial neural network. *Medical Engineering & Physics*, *124*(1), 104104. https://doi.org/10.1016/j.medengphy.2024.104104.

[17] Zheng, Y., Tian, B., Yu, S., Yang, X., Yu, Q., Zhou, J., Jiang, G., Zheng, Q., Pu, J., Wang, L. (2025). Adaptive boundary-enhanced Dice loss for image segmentation. *Biomedical Signal Processing and Control*, *106*, 107741. https://doi.org/10.1016/j.bspc.2025.107741.

[18] Taha, A.A., Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging*, *15*(1), 29. https://doi.org/10.1186/s12880-015-0068-x.

[19] Supriyadi, M.R., Samah, A.B.A., Muliadi, J., Awang, R.A.R., Ismail, N.H., Majid, H.A., Othman, M.S. Bin, Hashim, S.Z.B.M. (2025). A systematic literature review: exploring the challenges of ensemble model for medical imaging. *BMC Med Imaging*, *25*, 128. https://doi.org/10.1186/s12880-025-01667-4.

[20] Salehi, A.W., Khan, S., Gupta, G., Alabduallah, B.I., Almjally, A., Alsolai, H., Siddiqui, T., Mellit, A. (2023). A study of CNN and transfer learning in medical imaging: advantages, challenges, future scope. *Sustainability, 15*, 5930. https://doi.org/10.3390/su15075930.

[21] Goodfellow, I.J., Shlens, J., Szegedy, C. (2014). Explaining and harnessing adversarial examples. *3rd International Conference on Learning Representations.*