Employing Grad-CAM in DL Models for Tumour Segmentation and Visual Explanation: An Empirical Study

Dr. Rosy Sarmah, Ramakrishnananda, and Pranjal Singh Katiyar

Tezpur University, Napaam, Sonitpur, Tezpur, Assam 784028, India,
rosy8@tezu.ernet.in, rknanda.0111@gmail.com, pranjalsinghkatiyar@gmail.com

Abstract. This study employs Grad-CAM to enhance the interpretability of Deep Learning models for tumour segmentation on the Breast Ultrasound and LiTS17 datasets. We evaluated three segmentation models: U-Net, MultiResU-Net, and DCUNet, using metrics such as accuracy, precision, recall, Intersection over Union (IoU), Dice Coefficient, and various loss functions. The results indicate that MultiResU-Net consistently outperforms U-Net and DCUNet across both datasets. Grad-CAM heatmaps revealed that MultiResU-Net exhibits a high degree of focus on tumour regions, leading to superior segmentation accuracy. In contrast, DCUNet and U-Net showed moderate to low focus and targeting accuracy. By providing visual explanations of model decisions, Grad-CAM enhances the transparency and trustworthiness of segmentation models, thereby increasing trust among medical experts in the decision-making process.

Keywords: U-Net, DCU-Net, MultiRes-Unet, Grad-CAM, LiTS17

1 Introduction

Deep Learning (DL) according to Goodfellow et al.[1], a potent subset of Machine Learning (ML), has emerged as a transformative force in the domain of Artificial Intelligence (AI). Rooted in the concept of neural networks inspired by the intricate workings of the human brain, DL excels at processing vast amounts of unstructured data by automatically learning hierarchical representations through multiple layers of neurons. This paradigm shift has revolutionized various fields, including medical imaging. The application of segmentation techniques in medical imaging stands as a testament to the transformative potential of DL, particularly through the utilization of convolutional neural networks (CNNs), can be seen in Simonyan and Zisserman [2]. CNNs are a specialized type of neural network designed to effectively process visual data, making them particularly well-suited for tasks such as medical image segmentation. From identifying tumours in MRI (Magnetic Resonance Imaging) scans to delineating organs in CT (Computed Tomography) images, from Ultrasounds to X-Rays, CNN-based segmentation empowers healthcare professionals with invaluable insights, facilitating timely interventions and personalized treatments, as shown by Litijens et al. [3]. By automating tedious tasks and augmenting human expertise, medical image segmentation not only enhances diagnostic accuracy but also expedites workflows, ultimately improving patient outcomes. Furthermore, the benefits of leveraging CNNs for medical image segmentation extend beyond mere efficiency gains. It enables the development of predictive models, paving the way for early disease detection and prognostic assessments, see in Litijens et al. [4] and Sarker et al. [5]. Despite the remarkable success of CNN-based models in medical image segmentation, their inherent complexity often poses challenges in interpretation and explainability by Holzinger et al. [6]. In clinical settings, where decisions directly impact patient care, the opacity of these models can hinder their acceptance and trust among healthcare professionals. Consequently, there is a growing need for techniques that provide insights into the decision-making process of CNN-based models, particularly in the context of medical imaging.

One such technique, Gradient-weighted Class Activation Mapping (Grad-CAM) presented by Selvaraju et al. [7], has emerged as a promising approach for enhancing the interpretability of CNN-based models. By highlighting the regions of an image that are most influential in driving the model's decision, Grad-CAM offers insights into the areas of interest within medical images, thereby improving the transparency and trustworthiness of segmentation models. In the context of tumour segmentation, where precise delineation of tumour boundaries is critical for diagnosis and treatment planning, the integration of Grad-CAM holds potential for enhancing diagnostic accuracy and clinical decision-making, seen in Zhou et al. [8] and by Boykov and Jolly in [9].

This paper presents an empirical study on employing Grad-CAM in CNN-based models for tumour segmentation and visual explanation. By elucidating the decision-making process of segmentation models, we aim to bridge the gap between their impressive performance and the need for interpretability in medical imaging. Through the integration of Grad-CAM, we seek to empower clinicians with valuable insights into the underlying mechanisms driving model predictions, ultimately improving patient outcomes in clinical practice.

2 Literature Review

In recent years, DL techniques, such convolutional neural networks (CNNs) used by LeCun et al. in [10], have revolutionized medical image segmentation by providing powerful tools capable of extracting intricate patterns and features from complex medical images. Among the various segmentation architectures, U-Net presented by Ronneberger et al. [11], has emerged as a widely adopted framework for biomedical image segmentation. U-Net's architecture, characterized by a contracting path for feature extraction and an expansive path for precise localization, enables accurate segmentation of structures with fine details while mitigating the challenges posed by limited annotated data.

Building upon the success of U-Net, researchers have proposed extensions and variants tailored to specific segmentation tasks. One such extension is the DCU-Net (Dual-Context U-Net) presented by Li et al. in [12] which incorporates dual-context attention mechanisms to capture both local and global contextual information for improved segmentation performance. DCU-Net's ability to effectively integrate multi-scale contextual information makes it well-suited for challenging segmentation tasks, including tumour segmentation in medical imaging. Similarly, the MultiresU-Net (Multiresolution U-Net) by Ibtehaz et al. [20] which leverages a multi-resolution architecture to enhance segmentation accuracy and robustness. By integrating features extracted at multiple resolutions, MultiresU-Net effectively captures both coarse and fine details, facilitating accurate segmentation of complex structures in medical images. In the domain of tumour segmentation, previous studies have demonstrated the effectiveness of U-Net, DCU-Net, and MultiresU-Net architectures. For instance, Hossain et al. in [13] applied U-Net for brain tumour segmentation in MRI images, achieving competitive results compared to traditional methods. Similarly, Li et al. [14] employed DCU-Net for liver tumour segmentation in CT scans, demonstrating superior performance in delineating tumour boundaries and reducing false positives. Additionally, Gao et al. in [15] utilized MultiresU-Net for lung tumour segmentation in chest CT images, showcasing the architecture's ability to handle diverse tumour shapes and sizes across different imaging modalities.

While CNN-based approaches have shown promise in tumour segmentation, the interpretability of these models remains a critical concern. Explainable AI (XAI) aims to make the decision-making processes of AI systems transparent and understandable to humans. XAI methods provide insights into how models arrive at their predictions, fostering trust and facilitating the validation of model outputs. Techniques such as Grad-CAM given by Selvaraju et al. [7] are popular for visualizing regions of an image that significantly impact model predictions. By highlighting these areas, Grad-CAM helps clinicians understand and trust the decisions made by CNNs, thereby enhancing the interpretability of segmentation models. In the context of medical image segmentation, several studies have integrated Grad-CAM into CNN-based models to improve interpretability. For instance, Zhou et al. in [16] employed Grad-CAM in a DL model for lung nodule detection and localization in chest CT scans. Their results demonstrated that the visual explanations provided by Grad-CAM aided radiologists in confirming the presence of nodules and evaluating the model's performance. Similarly, Wang et al. in [17] utilized Grad-CAM in a DL model for breast cancer segmentation in mammography images, enabling clinicians to identify and verify the segmented regions with confidence. Latest work of field can be seen in [26]. Another paper by Zhang and Xiao [27] investigates medical image segmentation using Deep Neural Networks. It compares the performance of MobileNetV2 and ResNet backbones within the DeepLabV3+ structure. Grad-CAM is used to visualize how data augmentation and hyperparameter optimization affect the segmentation accuracy of these models.

Despite the progress made in incorporating visual explanation techniques into CNN-based models for medical image segmentation, challenges remain, particularly in achieving robust and generalizable models. Factors such as data heterogeneity, class imbalance, and variability in tumour morphology pose significant challenges to segmentation performance. Furthermore, the integration of domain-specific knowledge and expert feedback into the interpretation process could enhance the clinical utility of visual explanation techniques.

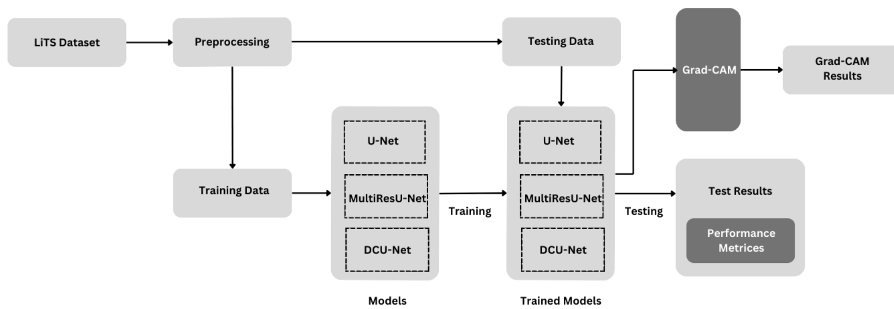

Fig. 1. Flow chart for the research

3 Methodology

In this section, we delineate the methodological framework employed to conduct our empirical study on employing Grad-CAM in DL (DL) models for tumour segmentation and visual explanation in medical imaging. Our methodology encompasses the selection and preparation of the dataset, preprocessing steps applied to enhance data quality, the implementation of segmentation models, and the integration of Grad-CAM for visual interpretation of model predictions, the flow of our proceure can be understood using the Flow chart in figure 1.

3.1 Dataset

The Datasets used in this study are (i) Breast Ultrasound dataset for binary segmentation, and (ii) LiTS17 Dataset for multiclass segmentation.
Breast Ultrasound Images Dataset [18] 2020 has a total of 780 images each of 512 X 512 pixels with three classes - normal, benign and malignant.
LiTS – Liver Tumour Segmentation Challenge Dataset (LiTS17) [19] comprises 130 CT scan images sourced from various clinical sites globally, formatted in NII (Neuroimaging Informatics Technology Initiative) standard. Each image has a resolution of 512 x 512 pixels with variable depths, with each image classified to either category of Background, Liver, and tumour, and the total dataset size is 49.9 GB.

3.2 Dataset Preprocessing

Proper preprocessing, including normalization, resizing, and format conversion, ensures images are suitable for ML tasks. For both datasets, images were downscaled to 128x128 using the resize() function from the scikit-image library, converted to NumPy arrays, and stored in .npy format for efficiency. The Breast Dataset was split into 10% for testing, with the remaining 90% for training, of which 20% was used for validation. For LiTS, 130 NII images were converted into 58,638 PNGs using the nilabel library by Abraham et al. [25]. A subset of 39,475 images had no liver representation, while 19,163 included the liver or liver tumor. From this cohort, 10% of the images were reserved for testing.

3.3 Segmentation Models

Segmentation models are crucial in medical imaging for accurately identifying and delineating tumours from surrounding tissues. In this study, we utilized three prominent models: U-Net, MultiResU-Net, and DCU-Net. U-Net is known for its encoder-decoder architecture with skip connections, allowing for precise localization and efficient training even with limited data. MultiResU-Net enhances the traditional U-Net by incorporating multiresolution analysis and residual connections, leading to better handling of complex and varied tumour structures. DCU-Net builds on the strengths of U-Net by introducing dense connectivity patterns, which improve feature propagation and capture finer details in the segmentation process. These models were selected for their proven effectiveness in medical image segmentation, enabling the accurate and reliable identification of tumour regions.

The U-Net architecture, introduced by Ronneberger [11], is a well-known CNN for semantic segmentation in biomedical images. It has an encoder with convolution and max-pooling layers, a bottleneck, and a decoder with transposed convolutions. Skip connections link the encoder and decoder to preserve spatial information. The final layer uses a 1x1 convolution with sigmoid or softmax activation for segmentation.

MultiResUNet by Ibtehaz et al. [20] is a DL model for medical image segmentation, effectively handling complex structures and size variations. It includes an encoder with convolution, batch normalization, ReLU activation, and pooling layers. MultiRes blocks capture multi-scale features, and a bottleneck processes high-level features. The decoder reconstructs details with up-sampling and convolution, using skip connections to retain spatial information. The final layer applies a convolution with softmax or sigmoid activation for segmentation.

DCU-Net (Dual Channel U-Net) presented by Jian et al. [21] is an advanced medical image segmentation model that enhances U-Net with DenseNet's dense connectivity for better feature propagation. It has an encoder with DCU blocks (convolution, batch normalization, ReLU, pooling), a bottleneck for high-level features, and a decoder that reconstructs details with up-sampling and dense blocks. Skip connections recover spatial information, and the final layer uses convolution with softmax or sigmoid activation for segmentation.

In this study, the ML models processes medical images of size 128x128, producing segmented images where each pixel is classified into different categories of the classes present.

3.4 Grad-CAM

Grad-CAM (Gradient-weighted Class Activation Mapping) presented by Selvaraju et al. [7], introduced in 2017, is a technique for visualizing and interpreting the decisions made by convolutional neural networks (CNNs). Unlike Class Activation Mapping (CAM), which requires network architecture modification, Grad-CAM offers a flexible approach that can be applied to any CNN-based model without changes. By leveraging the gradients of any target concept flowing into the final convolutional layer, Grad-CAM produces a heatmap that highlights important regions in the input image contributing to the model's prediction. The algorithm starts by creating a gradient model that maps the input image to the activations of the last convolutional layer and the model's output predictions. Using TensorFlow's GradientTape, it records operations for automatic differentiation, passing the input image through the gradient model to obtain the output of the last convolutional layer and the model's predictions. For multi-label segmentation, the class of interest can be specified, or the top predicted class can be selected using TensorFlow argmax, while for binary segmentation, the mean activation across the entire image is used. The algorithm computes the gradient of the class channel with respect to the output feature map of the last convolutional layer, performing global average pooling to capture the importance of different feature map channels. These pooled gradients weight the feature maps, which are summed to produce the class activation heatmap. The heatmap is normalized to the range [0, 1] by setting negative values to zero and dividing by the maximum heatmap value, then converted to a numpy array for further processing or visualization. Grad-CAM provides a visual explanation of the model's decision-making process by highlighting image regions that contribute most to predictions for multi-label and binary segmentation tasks. This helps in identifying whether the model is focusing on the right parts of an image or getting distracted by irrelevant features. Further it makes the "black-box" nature of CNNs more transparent, showing whether the model is focusing on relevant features or getting distracted by irrelevant ones. This transparency can increase user trust in the model's predictions and aids in diagnosing performance issues and guiding improvements by making model outcomes easier to interpret.

4 Results and Discussions

The study and analysis were conducted on the specified datasets, utilizing various segmentation models. The experiment involved thorough examination and preprocessing of the datasets to meet model-specific input requirements and optimize training efficiency. Different segmentation models were applied and trained with early stopping to prevent overfitting. Performance metrics were calculated on both training and testing data. Mask images were visualized for specific testing dataset images, and informative graphs depicting performance metrics during training epochs were generated. Heatmaps for each class were produced using Grad-CAM and superimposed on the original input images. The model used a learning rate of 0.001, the Adam optimizer, softmax for multiclass segmentation, ReLU for intermediate layers, and a batch size of 16. Training and testing were conducted on Google Colaboratory (Colab), utilizing its free GPU access to accelerate processing, support rapid prototyping, experimentation, and ensure reproducibility without relying on local computing resources.

Performance metrics (Table 1) in DL/ ML measure model effectiveness, guiding optimization for robust algorithms offering insights into a model's behavior and aiding in assessing strengths and weaknesses for a given task. For assessing the quality of image segmentation we have used 7 performance metrices, viz., Accuracy, Precision, Recall defined by Han et al. in book ([22]), IoU defined in Shanmugamani's book [23], Dice coefficient [24], Categorical Cross Entropy and Binary Cross Entropy defined by Goodfellow et al. in book ([1]) as given in Table 1. Confusion Matrix explained in Han et al. book [22], is crucial for evaluating model performance by comparing predictions with actual outcomes in both binary and multiclass scenarios. It breaks down predictions into True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN), representing correct and incorrect predictions. In image segmentation, these elements correspond to accurately identified targets, erroneously highlighted non-targets, correctly identified backgrounds, and missed targets, respectively. For multi-class classification, categorical cross entropy mentioned by Goodfellow et al. in book [1] is essential for handling class imbalances by optimizing the model. The implementation of Grad-CAM on Breast and LiTS17 dataset provides valuable insights into the models' focus during the segmentation task as discussed next.

Results on Breast Ultrasound Images Dataset with Grad-CAM : The heatmaps generated (Figures: 2, 3, 4) reveal that MultiResUNet exhibits a high degree of focus on the tumour region, as indicated by the dark red coloration, suggesting precise targeting and effective concentration on the tumour. This results in superior segmentation accuracy. In contrast, DCUNet's heatmap shows an orange hue over the tumour area, reflecting a moderate level of focus, indicating that while the model does concentrate on the tumour, its intensity is less compared to MultiResUNet. UNet's heatmap, with its lighter colors and smaller focus area, suggests a less accurate and less concentrated focus on the tumour region. These observations explain why MultiResUNet achieves better metric results (see Table 2) and outperforms both DCUNet and UNet. The dark red heatmap underscores MultiResUNet's enhanced performance, demonstrating its ability to accurately target and segment the tumour area with high precision.

Results on LiTS Dataset with Grad-CAM: The heatmaps (Figures: 5, 6, 7) reveal that MultiResUNet effectively targets the liver area without diverting to nearby organs, with colors well-contained within the liver. For tumour segmentation within the liver, MultiResUNet demonstrates precise targeting, indicated by red and orange patches on the heatmap for class 3, leading to superior segmentation results. Conversely, DCUNet's heatmap shows occasional deviation from the liver area during segmentation, sometimes focusing on other regions, which results in a decreased IoU.

Table 1. Performance Metrics

Metric	Formulae
Accuracy	$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$
Recall	$Recall = \frac{TP}{TP+FN}$
Precision	$Precision = \frac{TP}{TP+FP}$
IoU	$IoU = \frac{TP}{TP+FP+FN}$
Dice Coefficient	$Dice = \frac{2\times TP}{2\times TP+FP+FN}$
Categorical Cross Entr.	$H(y,p) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{i,j}\log(p_{i,j})$
Binary Cross Entr.	$BCE = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1-y_i)\log(1-p_i)]$

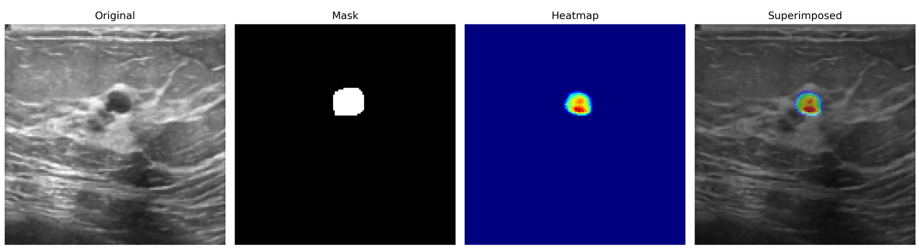

Fig. 2. Breast Ultrasound Dataset U-Net Grad-CAM Result

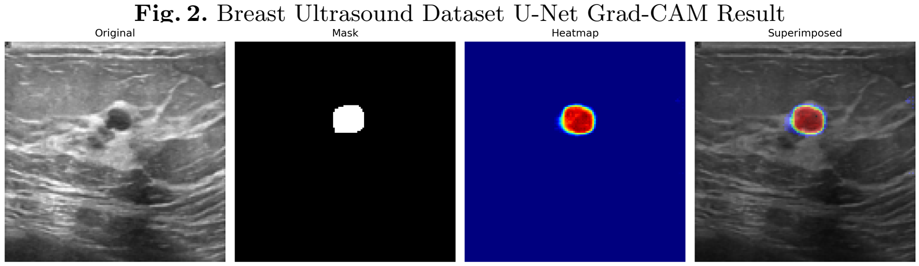

Fig. 3. Breast Ultrasound Dataset MultiresU-Net Grad-CAM Result

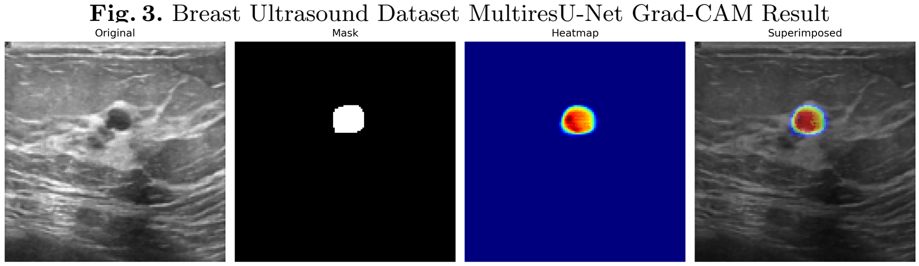

Fig. 4. Breast Ultrasound Dataset DCU-Net Grad-CAM Result

Table 2. Breast Ultrasound Dataset - Test Results

Metric	UNet	MultiResUNet	DCUNet
Accuracy	0.9536	0.9561	0.9444
Recall	0.3778	0.5513	0.5330
Precision	0.6116	0.5423	0.51207
IoU	0.7292	0.7836	0.7462
Dice coeff.	0.4657	0.5457	0.5263
Binary Cross Entr.	0.1224	0.2413	0.3834

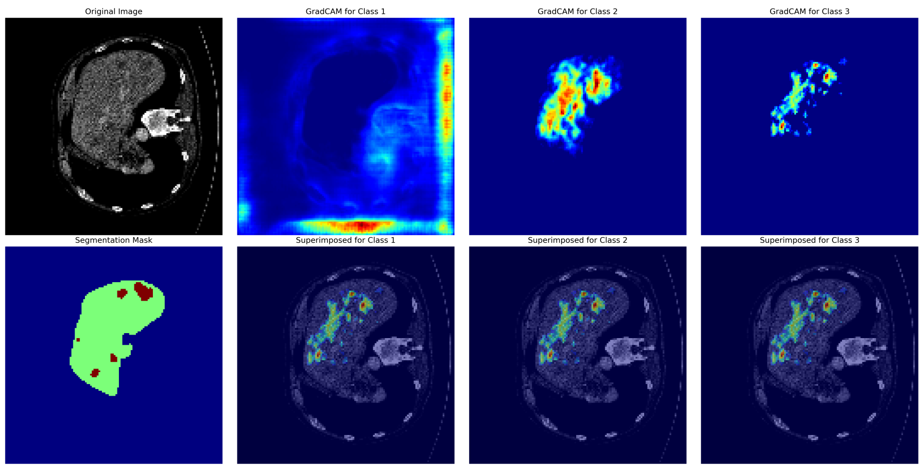

Fig. 5. LiTS Dataset U-Net Grad-CAM Result

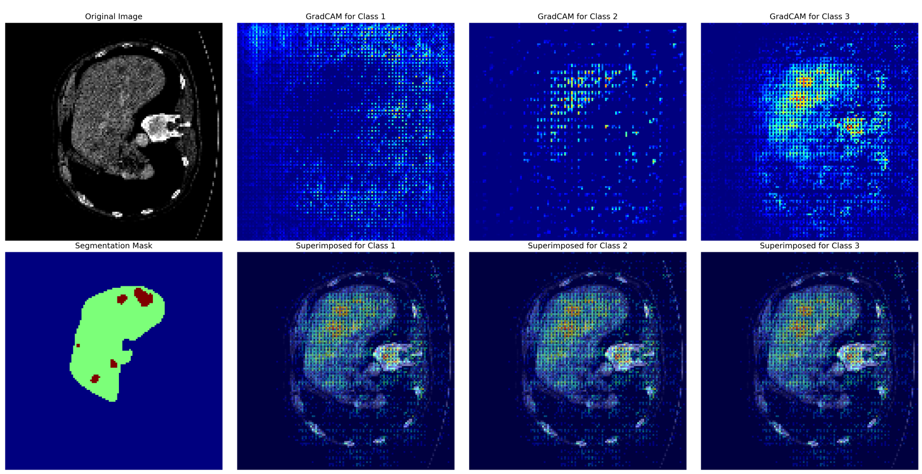

Fig. 6. LiTS Dataset MultiResU-Net Grad-CAM Result

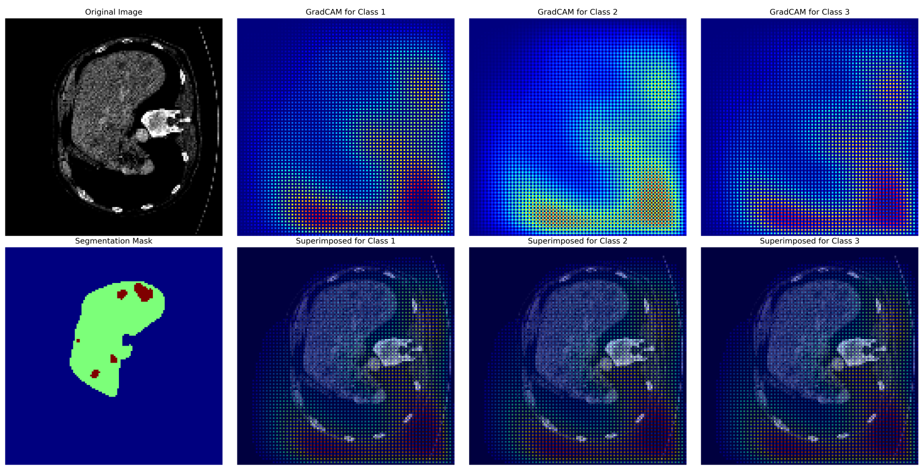

Fig. 7. LiTS Dataset DCU-Net Grad-CAM Result

While DCUNet generally targets the tumour region well, its occasional deviation affects its metrics compared to MultiResUNet. UNet's heatmap indicates frequent deviation from the liver area during liver segmentation and similar issues during tumour segmentation, leading to a lower IoU than MultiResUNet. Thus, the heatmaps underscore MultiResUNet's more effective focus on the liver and tumour regions, justifying its superior performance and better metric (see Table 3) results compared to DCUNet and UNet.

5 Conclusion

In this study, we used Grad-CAM to interpret DL models for tumour segmentation on the Breast Ultrasound and LiTS17 datasets. We assessed U-Net, MultiResU-Net, and DCUNet models using metrics like accuracy, precision, recall, IoU, Dice Coefficient, and Binary/Categorical Cross Entropy losses. MultiResU-Net consistently outperformed U-Net and DCUNet on both datasets. Grad-CAM heatmaps showed MultiResU-Net focusing well on tumour regions, while DCUNet had moderate focus, and U-Net the least. These results highlight the importance of model architecture in medical image segmentation and the value of Grad-CAM for visual explanations. MultiResU-Net's precise targeting supports its use in clinical decision-making. Future work could involve combining MultiResU-Net with other interpretability methods, testing on varied datasets, exploring real-time clinical applications, and using unsupervised learning to enhance performance with limited data.

Table 3. LiTS Dataset - Test Results

Metric	UNet	MultiResUNet	DCUNet
Accuracy	0.9974	0.9982	0.9979
Recall	0.9960	0.9973	0.9969
Precision	0.9962	0.9974	0.9970
IoU	0.9062	0.9373	0.9273
Dice coeff.	0.9961	0.9973	0.9969
Categorical Cross Entr.	0.0096	0.0073	0.0161

References

1. Goodfellow, I., Bengio, Y., Courville, A.: DL. MIT Press (2016).
2. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014).
3. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A., van Ginneken, B., Sánchez, C.I.: A survey on DL in medical image analysis. Medical Image Analysis 42, 60-88 (2017). doi:10.1016/j.media.2017.07.005
4. Litjens, G., et al. (2017). A survey on DL in medical image analysis. Medical Image Analysis, 42, 60-88. https://doi.org/10.1016/j.media.2017.07.005
5. Sarker, I.H., et al. (2020). Machine learning and DL in medical imaging: applications, challenges, and future trends. Springer, 30, 123-136. https://doi.org/10.1007/s00138-020-01145-8
6. Holzinger, A., Langs, G., Denk, H., Zatloukal, K., Müller, H.: Causability and explainability of artificial intelligence in medicine. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 9(4), e1312 (2019). doi:10.1002/widm.1312
7. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-CAM: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 618-626 (2017). doi:10.1109/ICCV.2017.74
8. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A.: Learning deep features for discriminative localization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2921-2929 (2016). doi:10.1109/CVPR.2016.319
9. Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary region segmentation of objects in N-D images. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 105-112 (2001). doi:10.1109/ICCV.2001.937505
10. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324 (1998). https://doi.org/10.1109/5.726791
11. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 234-241 (2015). doi:10.1007/978-3-319-24574-4_28
12. Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.W., Heng, P.A.: H-DenseUNet: Hybrid densely connected UNet for liver and tumour segmentation from CT volumes. IEEE Transactions on Medical Imaging 37(12), 2663-2674 (2018). doi:10.1109/TMI.2018.2845918

13. Hossain, M., Saifullah, S., Arifuzzaman, M.: DL in biomedical image segmentation: U-Net, DCU-Net, and their application in brain tumour segmentation. In: Proceedings of the IEEE International Conference on Biomedical Engineering and Applications, pp. 1-6 (2021). doi:10.1109/ICBEA.2021.9458353
14. Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.W., Heng, P.A.: H-DenseUNet: Hybrid densely connected UNet for liver and tumour segmentation from CT volumes. IEEE Transactions on Medical Imaging 37(12), 2663-2674 (2018). doi:10.1109/TMI.2018.2845918
15. Gao, F., Li, M., Zhu, Y., Feng, Q., Shen, D., Xu, X.: Multi-scale convolutional neural network for lung nodule classification in CT images. Neurocomputing 175, 827-835 (2016). doi:10.1016/j.neucom.2015.10.020
16. Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J.: UNet++: A nested U-Net architecture for medical image segmentation. In: Proceedings of the DL in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 3-11 (2018). doi:10.1007/978-3-030-00889-5_1
17. Wang, X., You, S., Zhu, L., Li, M., Qian, Y., Ye, H.: Weakly supervised breast cancer segmentation using positron emission tomography-computed tomography images. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1355-1361 (2019). doi:10.1109/BIBM47256.2019.8983191
18. Breast ultrasound images dataset. (2020) https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset
19. Lits – liver tumour segmentation challenge (LiTS17) organised in conjunction with isbi 2017 and MICCAI 2017. (2017). https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset
20. Ibtehaz, N., Rahman, M.S.: MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. Neural Networks 121, 74-87 (2020). doi:10.1016/j.neunet.2019.08.025
21. Jiang, Z., Ding, C., Liu, M., Tao, D.: Dual-Path U-Net for Medical Image Segmentation. In: Proceedings of the 24th International Conference on Pattern Recognition (ICPR), pp. 3505-3511 (2018). doi:10.1109/ICPR.2018.8546197
22. Han, J., Kamber, M., Pei, J.: Data Mining Concepts and Techniques, Third Edition. Morgan Kaufmann Publishers (2012). ISBN 978-0-12-381479-1.
23. Shanmugamani, R.: DL for Computer Vision. Packt Publishing (2018).
24. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778 (2016). doi:10.1109/CVPR.2016.90
25. Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., ... & Thirion, B. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics, 8*, 14. doi:10.3389/fninf.2014.00014
26. MDPI (2023). Deep Learning for Medical Image-Based Cancer Diagnosis. https://www.mdpi.com/2072-6694/15/14/3608
27. Zhang, X., Xiao, Z. (2020). A visualization method based on the Grad-CAM for medical image segmentation model. *In 2020 International Conference on Artificial Intelligence in Medicine (AIME)* (pp. 147-152). IEEE.