Journal of Intelligent Systems and Internet of Things
Vol. 17, No. 02, PP. 360-368, 2025
DOI: https://doi.org/10.54216/JISIoT.170223

Enhancing Breast Cancer Detection in CESM Mammograms: Impact of Data Augmentation on U-NET Segmentation Performance

Taha Y. Abdulqader[1,*], Kifaa Hadi Thanoon[2], Shatha A. Baker[1,*]

[1]Networks and Computer Software Techniques, Northern Technical University, Mosul, Iraq

[2]Cyber Security, Northern Technical University, Mosul, Iraq

Emails: taha.yassin@ntu.edu.iq; kh.thanoon@ntu.edu.iq; Shathaab@ntu.edu.iq

Abstract

When using mammography to diagnose breast cancer, segmenting medical scans is a crucial step. Accurate segmentation facilitates early diagnosis, which in turn makes it possible to administer individualized treatment plans, ultimately improving patient outcomes. However, for these Deep Learning (DL) models to be trained efficiently and perform optimally, they require access to large datasets. The lack of sufficient photographs in many publicly available datasets to adequately train deep learning models is a common flaw. Therefore, this work aims to examine the effects of various affine data augmentations on the Dice Score of a U-NET model utilizing a recently released public dataset of Contrast-Enhanced Spectral Mammography (CESM) images. The collection consists of 1003 CESM images and matching segmentation masks made by a certified radiologist. Modifying certain model parameters on the CESM dataset and investigating the impact of single and combination data augmentations on the model's overall performance are the objectives of the study. Images that were moved in the x-direction and sheared vertically were used to train the best-performing model. On the test set, the model's Dice Score was 56.6%, which was 9% better than the baseline result and showed how crucial data augmentation is when working with small datasets.

Received: January 08, 2025 Revised: March 10, 2025 Accepted: June 11, 2025

Keywords: Deep Learning; Contrast-Enhanced Spectral Mammography; Mammograms; U-NET

1. Introduction

Given the high incidence of breast cancer in women, it is now more crucial than ever to build precise segmentation models for the disease's detection. According to data from 2021, breast cancer is responsible for almost one-third of all female cancer diagnoses, and it is directly responsible for about 2.5 percent of female deaths. Reducing the disease's incidence and mortality rates requires early detection through routine and consistent screenings [1]. Accordingly, the NHS recommends that women 50 years of age and older have mammograms every three years. A 2005 study found that screening mammography lowers breast cancer death rates among women aged 50 to 69 by 20% to 35%, which supports this recommendation. As a result, Digital Mammography (DM) has gained popularity over the past 20 years and is thought to be the most accurate way to screen for breast cancer [2–3]. Nevertheless, DM has limitations, particularly when it comes to detecting anomalies in thick breast tissue, where it can be challenging to differentiate between suspicious areas and healthy structures. Contrast Enhanced Spectral Mammography (CESM) has drawn interest because of its enhanced capacity to identify disease, especially in women with dense breast tissue, and has helped to overcome those constraints. By combining conventional mammography methods with an iodinated contrast agent, CESM creates two sets of images: a high-energy scan and a low-energy scan that resembles a typical DM image. A dual-energy subtraction technique is used to integrate these scans, producing a contrast-enhanced image with better clarity [4–7].

Computer-Aided Diagnosis (CAD) describes the use of computer programs and software tools to assist medical professionals in examining medical images such as MRIs and mammograms. Although the concept dates back to the 1960s, significant research into CAD platforms began in the early 1980s, eventually leading to FDA clearance of the first CAD system built to aid in breast cancer detection using screening mammography [8-12]. In recent years, CAD technologies have played a key role in helping radiologists by working as a second-reader system, where two specialists review the same images independently before making a final judgment. While this method increases accuracy, it requires considerable time and resources. Therefore, using CAD systems as a support tool is especially crucial in healthcare systems like the NHS, where there is a notable lack of radiologists. According to a 2018 report, there is a scarcity of about 2000 radiologists, which causes delays in diagnosis and treatment as well as lengthier wait times. This circumstance emphasizes the significance of studies meant to improve CAD system performance [13].

Therefore, the purpose of this work is to examine the literature on the use of Deep Learning (DL) for breast cancer segmentation, examining recent developments and evaluating the extent to which DL models—more especially, picture segmentation—have facilitated disease diagnosis. To create a segmentation pipeline and utilize a simple U-NET model to identify breast cancer regions in CESM pictures using the CDD-CESM dataset, which hasn't been investigated for this purpose yet. Another objective is to examine how different affine data augmentation techniques affect the model's performance. The effects of these affine transformations—rotation, translation, scaling, and shearing—on the model's accuracy and generalizability will be assessed separately. Additionally, a grid search using the best individual data augmentation approaches should be conducted in order to identify the most effective combination. To enhance outcomes, a grid search methodically explores different hyperparameter configurations and approaches. Using the grid search results to fine-tune the U-NET architecture is another goal. For improved results, this involves modifying variables like batch size, epoch count, and learning rate. The enhanced model seeks to increase breast cancer segmentation in CESM images' precision and dependability.

The study is as follows; in the next section, we will look at related works. Section III focuses on the implementation. Section IV contains the experimental analysis, while Section V contains the study's conclusion and future work.

2. Related Work

This review of the literature investigates several methods for CAD of breast cancer using mammography. In the beginning, region-based and edge-based algorithms were used for segmenting images [14]. These techniques distinguish between malignant and non-cancerous regions using pixel properties like intensity and texture. With a Jaccard value of 0.90, region-growing segmentation [15], which was improved by the Dragon Fly algorithm, performed better on the DDSM dataset. Similarly, when applied to detect micro-calcifications from ROIs in mammograms, watershed segmentation [16], which simulates water movement across pixel intensities, obtained an IOU score of 0.708. Segmenting images like [17] also helps with pre-processing, including employing regression analysis and Otsu thresholding to eliminate the pectoral muscle in Medio-Lateral Oblique (MLO) views. Traditional supervised Machine Learning (ML) methods such as k-Nearest Neighbors (kNN), Naive Bayes, and random forests were widely used prior to the emergence of DL. [18] used the Wisconsin Diagnosis Breast Cancer dataset and discovered that kNN was the most successful, with an accuracy of 95.9%. [19] segmented tumor regions in the MIAS dataset using FCM in conjunction with pre-processing methods such as CLAHE and GLCM feature extraction, with 94.6% accuracy, 96.4% specificity, and 86.2% sensitivity. However, generalizability is limited by the tiny sample size.

With advancements in computing power and access to public datasets like the Cancer Imaging Archive1, DL—especially Convolutional Neural Networks (CNNs)—became dominant in CAD. CNNs automate feature extraction through layers such as convolutional, pooling, and fully connected layers. Max pooling is particularly useful in highlighting tumour regions, while the final fully connected layer performs classification. A widely used CNN for segmentation is U-NET, developed by [20], characterized by its encoder-decoder architecture without a fully connected layer. It has proven effective in biomedical image segmentation, with [21] applying it on DDSM mammograms and achieving a dice index of 79.39%. The study employed methods such as Normalization (BN) and dropout layers within DL models to improve performance. The BN layer standardizes the input to have a zero mean and constant standard deviation, which speeds up training and boosts generalization accuracy while avoiding the risk of missing sharp local minima. The dropout layer helps prevent overfitting by randomly removing specific units during training, ensuring that the model does not become overly reliant on certain neurons. These adjustments contributed to notable results, with the model achieving a dice index of 0.951 and a mean IOU of 0.909 using datasets from CBIS-DDSM. However, the study’s limitations include the use of only mammograms containing masses, which limits the model's ability to generalize to real-world clinical scenarios. Another limitation is the large size of the model, with U-NET containing 31 million parameters, making it computationally expensive and

prone to overfitting when applied to smaller datasets. To solve this, [22] proposed Half-UNet, which reduces the number of parameters by removing redundant convolutions and replacing standard double convolutions with a ghost module, allowing for the creation of feature maps using simpler linear operations. The Half-UNet achieved similar performance with significantly fewer parameters. [23] achieved a mean IOU of 86.3% for Xenopus kidney embryos. [24] utilized 3D U-NET to segment breast cancer from MRI images, achieving a median dice score of 0.77, surpassing the average performance of radiologists. However, issues related to class imbalance in the dataset could be addressed through data augmentation techniques to balance benign and malignant cases. Unlike using separate models for detection, segmentation, and classification, [25] applied U-NET for all three stages, achieving a segmentation IOU of 90.50% and a dice index of 99.19% across 1079 images from the DDSM dataset. However, the study lacked a clear image selection criterion, which could affect the validity of the results.

3. Implementation

This study investigates the impact of various data augmentation techniques to enhance the performance of ML models for breast cancer segmentation from mammogram images. The initial phase involved developing a basic model using a U-Net architecture, incorporating batch normalization layers, and establishing a fundamental learning pipeline to obtain baseline results from the dataset. At this point, the pipeline did not integrate any data augmentation techniques. Later, a series of hyperparameter tuning experiments were conducted, involving several grid searches to determine the optimal parameters for the U-Net model. This approach ensured that the model was trained with the best possible configuration of hyperparameters. In the next phase, a comprehensive investigation was carried out on various affine transformations present in the dataset. The goal was to assess how different data augmentation techniques, such as rotation, flipping, and shearing, influenced the model's performance in breast cancer segmentation. By systematically testing different augmentations, the study aimed to identify the most effective individual techniques to improve the generalization of the model. These optimal parameters were then used in subsequent experiments to narrow down the search space and focus on data that are more specific augmentations. The final phase of the study explored the combined effect of applying multiple data augmentation techniques together. A grid search was performed to find the best combination of augmentations, optimizing the model's ability to generalize and improve segmentation performance. This experiment aimed to explore potential synergies between augmentations, where the combination of transformations could lead to better model performance than using individual augmentations alone.

The CDD-CESM dataset, which comprises 326 patients' regular and contrast-enhanced digital mammograms (Fig. 1), was used for this investigation. The Cancer Imaging Archive provided the dataset. Mammogram images in both MLO and CC perspectives, which were converted from DICOM to JPEG format using lossless compression, are included in the CDD-CESM dataset. Clinical data reports and CSV files containing manually created segmentations and comments for every image, including labels like "suspicious," "malignant," and "mass," are also included in the collection. The 2006 images in the sample are evenly divided between standard and contrast-enhanced digital mammograms. However, because of unnecessary annotations, not every image in the dataset was appropriate for training. The dataset might contain noise because some of the images were marked as postoperative, meaning the lump had been surgically removed. The masks on other photographs were either faulty or not well defined. Only images that included pertinent annotations, including "mass," "suspicious," or "calcification," were kept therefore. Of the 752 photos in the modified dataset, 338 had either benign or malignant masses, while 414 images had no masses at all. Three subsets of the dataset were created: 15% for testing, 15% for validation, and 70% for training. There were 526 images in the training set, 112 in the validation set, and 114 in the test set.

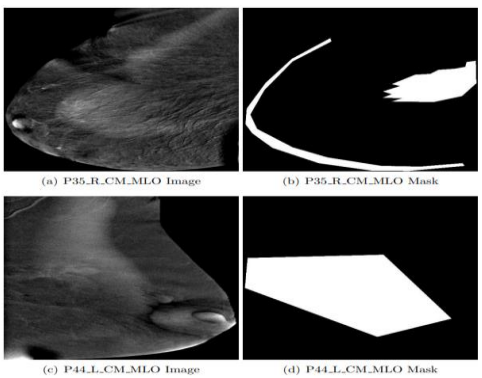

Figure 1. Two distinct images with weird and big masks. Both were marked as 'postoperative' in the annotation CSV

The data augmentations explored in this study focused on affine transformations that preserve parallel lines, such as rotation, flipping, and shearing as shown in Fig. 2. These techniques were chosen due to their simplicity and effectiveness in expanding the dataset without requiring complex hyperparameter adjustments. Both horizontal and vertical flipping were used, with vertical flips being less common but potentially valuable in improving generalization. Various degrees of rotation were explored, including small and large angles, to account for minor orientation variations in mammogram images. Shearing, which distorts images along one axis, was also considered, as it could help address orientation differences. Translation was another augmentation technique explored, though it is less frequently used with mammogram datasets. The translation aimed to shift the breast region to the centre of the image, reducing variations between left and right breast mammograms. Unlike many studies that focus on mammogram datasets, this study did not require extensive preprocessing steps such as decompression or contrast enhancement (e.g., Contrast Limited Adaptive Histogram Equalization, CLAHE), since the CESM images were already pre-processed. The main preprocessing step involved integrating the augmented images into the training set, resulting in a relatively simple pipeline.

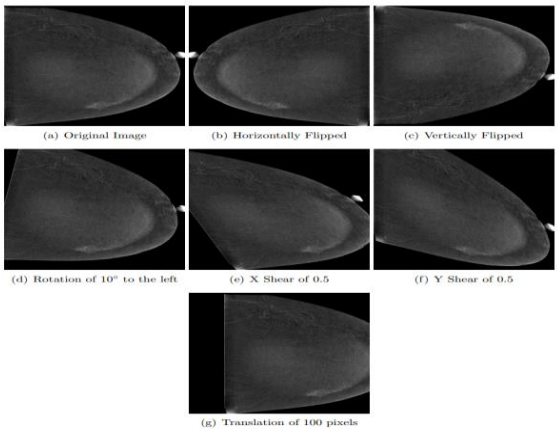

Figure 2. The impact of various data augmentation techniques on the file P 3 L CM CC.jpg image

Several metrics, such as the Dice coefficient, accuracy, precision, and recall, were used to evaluate the model's performance. The overlap between the ground truth and anticipated masks is measured by the Dice coefficient, a commonly used parameter in segmentation tasks. In medical imaging, when classes like masses might only take up a small percentage of the image, it is especially useful. The precision and recall metrics provided further information about the model's capacity to accurately detect positive pixels, while the accuracy statistic quantified the percentage of accurate pixel predictions. The model's overall effectiveness in separating breast cancer from mammography images was assessed with the use of these parameters.

4. Experimental Analysis

The baseline model in this study is based on the U-Net architecture with adjustments, including the addition of batch normalization layers following each convolution in the encoding path. The results suggest that the model faces overfitting during training, as shown by the increasing gap between the training and validation dice scores. This overfitting becomes especially noticeable after epoch 20, with the validation dice score levelling off between 0.55 and 0.6 as shown in Fig. 3. To resolve this, an early stopper was integrated to stop training when the difference between the training and validation dice scores reached a threshold of 0.03 as shown in Fig. 4. The early stopper proved successful, resulting in a higher validation dice score of 0.580 after just 26 epochs, compared to the baseline model.

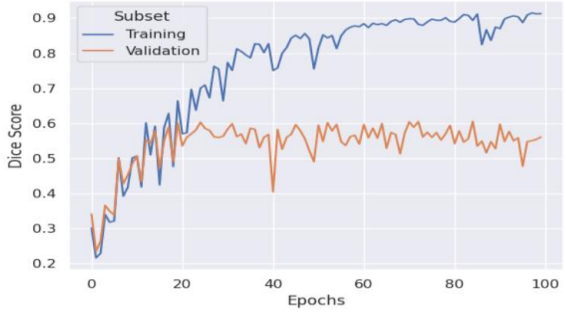

Figure 3. After epoch 20, the baseline model's validation and training dice scores diverge significantly

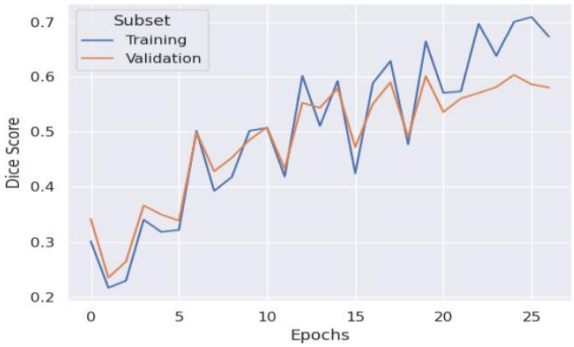

Figure 4. Dice score for the baseline model with an early stopper. Training ends before the necessary 100 epochs

A grid search was used to find the optimal hyperparameters for the model. The key hyperparameters tested included image size, batch size, and learning rate. The search space was limited due to GPU memory constraints and a reduced number of epochs (75 instead of 100). The results demonstrated that smaller image sizes performed better on average, and a batch size of 4 yielded the best results, with a dice score of 0.627, an improvement of 0.047 over the baseline model. A seed search was also performed to find the best initial parameters for the model as shown in Fig. 5. This experiment involved testing 10 different seeds, and the best-performing seed achieved a dice score of 0.627, identical to the results from the previous tuned model. The worst performing seed resulted in a performance 13% lower, highlighting the importance of choosing favourable initial parameters to avoid local minima during training.

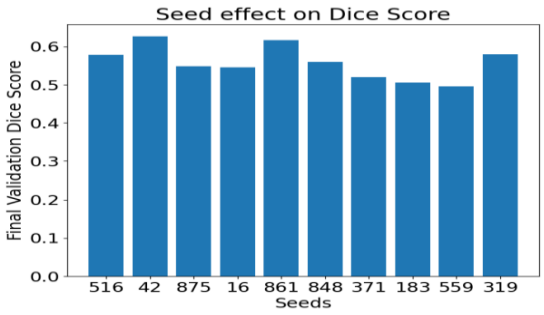

Figure 5. A bar chart depicting the final dice scores for the validation set for the various seeds

The effects of various data augmentation techniques were assessed, including horizontal flipping, vertical flipping, rotation, shearing, and translation as shown in Fig. 6. Horizontal and vertical flipping led to lower dice scores, suggesting that flipping does not enhance model performance in this case. On the other hand, rotations, especially smaller angles like 5° and 10°, led to improvements, achieving dice scores of 0.643 and 0.650, respectively. However, shearing and translation showed little improvement, providing no significant benefits to model performance. Following the individual augmentation tests, a search for the best combination of augmentations was carried out. However, due to the high computational cost of testing all possible combinations, the experiments were restricted to a smaller set of parameters. Contrary to expectations, the results indicated that combined augmentations did not significantly exceed the performance of the tuned baseline model. Specifically, translations led to a higher dice score when applied with larger translation values, but the overall impact was minimal.

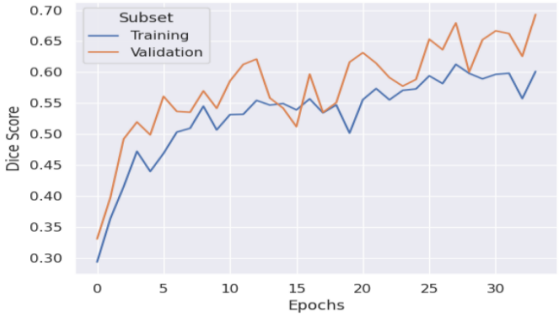

Figure 6. The training and validation dice score for each epoch while combining different augmentations on one image using the Albumentations library

A. Additional Experiments

For every test, two model checkpoints were reviewed i.e. one saved after the final training epoch, and another saved at the highest validation Dice score. As shown in Table I, the tuned model without data augmentations performed nearly 5% better than the baseline model across the test set, with Dice scores of 0.526 and 0.476, respectively. When separate data augmentations were applied, (refer to Table II), only rotation caused an improvement, reaching a peak Dice score of 0.550. In contrast, translation had the most negative effect, likely caused by the pre-cropped layout of the images, which made translation unnecessary or even detrimental for model generalization. Horizontal and vertical flip approaches resulted in minor improvements, both scoring maximum Dice scores of 0.504. Meanwhile, X and Y shearing gave mixed outcomes, with Y shearing performing slightly better (maximum Dice: 0.521). The performance of combined augmentation approaches is presented in Table III. Notably, the combination of Y shear (0.75) and translation (50 pixels) delivered the best Dice score across the test set, reaching 0.566. However, unlike the validation set, where multiple combinations appeared helpful, many augmentation combinations in the test round performed worse than the tuned model without enhancements. An unexpected observation was the use of Albumentation with rotation, which surprisingly reduced model performance. This could be due to the small size of the training data, which limits the potential value of using varied augmentations. It may also suggest that such transformations drastically altered the original image properties. Among the grouped augmentations, the combination involving horizontal flip, translation, and rotation achieved a high Dice score of 0.554. On the other side, combinations like rotation + Albumentation, with or without early stopping, failed to surpass the simpler augmentation methods.

Table 1: The outcomes of the investigations utilized to tune the model.

Investigation	Model	Dice Score	Precision	Recall	Accuracy
Baseline	Last Epoch	0.462	0.413	0.659	0.955
	Max Dice	0.459	0.478	0.618	0.961
Hyperparameter	Last Epoch	0.519	0.492	0.579	0.969
	Max Dice	0.537	0.652	0.511	0.975

Table 2: The impact of individual augmentations on the test set

Augmentation	Model	Dice Score	Precision	Recall	Accuracy
Horizontal Flips	Last Epoch	0.488	0.478	0.595	0.964
	Max Dice	0.512	0.553	0.538	0.969
Vertical Flips	Last Epoch	0.514	0.623	0.523	0.971
	Max Dice	0.509	0.600	0.548	0.972
Rotations	Last Epoch	0.544	0.605	0.551	0.973
	Max Dice	0.559	0.578	0.569	0.971
X Shearing	Last Epoch	0.484	0.603	0.449	0.971
	Max Dice	0.463	0.616	0.481	0.970
Y Shearing	Last Epoch	0.505	0.538	0.570	0.969
	Max Dice	0.530	0.567	0.558	0.971
Translation	Last Epoch	0.471	0.611	0.523	0.972
	Max Dice	0.487	0.642	0.506	0.972

Table 3: The effect of the individual augmentation on the test set

Augmentation Grouping	Model	Dice Score	Precision	Recall	Accuracy
HFlip + X Shear + Y Shear	Last Epoch	0.534	0.589	0.517	0.970
	Max Dice	0.511	0.577	0.513	0.969
HFlip + Rotation	Last Epoch	0.538	0.582	0.543	0.972

	Max Dice	0.541	0.585	0.545	0.972
HFlip + Translation + Rotation	Last Epoch	0.481	0.720	0.472	0.977
	Max Dice	0.562	0.593	0.604	0.973
Y Shear + Rotation	Last Epoch	0.488	0.612	0.506	0.973
	Max Dice	0.507	0.592	0.555	0.972
Y Shear + Translate + Rotation	Last Epoch	0.549	0.690	0.495	0.974
	Max Dice	0.513	0.532	0.604	0.970
Y Shear + Translate	Last Epoch	0.578	0.662	0.526	0.974
	Max Dice	0.578	0.662	0.526	0.974
Rotation + Albumentation + Early Stopper	Last Epoch	0.503	0.519	0.564	0.969
	Max Dice	0.503	0.519	0.564	0.969
Rotation + Albumentation + No Stopper	Last Epoch	0.538	0.531	0.595	0.969
	Max Dice	0.505	0.519	0.564	0.969

The best output was achieved using the tuned model with the combined augmentation of Y shear and translation, scoring a Dice score of 0.566. Visual analysis of this model's predictions (refer to Fig. 7) demonstrates strong accuracy on high-contrast images, especially where the suspected region is clearly prominent. Cases with large ground truth mask areas also scored well, often achieving Dice scores above 0.9. In contrast, Fig. 8 displays weaker predictions, particularly for images with small mass regions or those in the normal category, where the mask is entirely blank. This difference highlights a known flaw of the Dice score, where small mistakes are heavily penalized when the ground truth mask includes a small area. For instance, a single false-positive pixel in an empty mask can lead to a Dice score of 0, misrepresenting the model's actual performance. This challenge could have been reduced using post-processing techniques to remove tiny-segmented regions from the predictions. Across all analyses, model performance on the test set was significantly worse than on the validation set. This implies that the training and validation data may not entirely represent the full dataset. Moreover, class imbalance during data partitioning might have influenced results, as benign and malignant images were grouped due to the small quantity of benign samples.

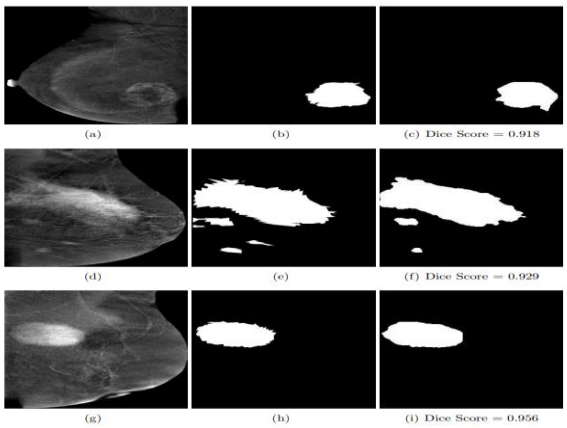

Figure 7. The best performing model accurately predicted several images from the test set. Column 1 displays the original mammograms, Column 2 shows the ground truth segmentations, and Column 3 shows the model predictions

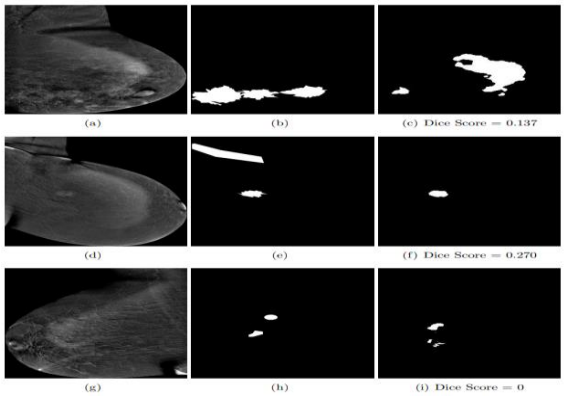

Figure 8. The best performing model produced incorrect predictions on certain images in the test set. Column 1 displays the original mammograms, Column 2 shows the ground truth segmentations, and Column 3 shows the model predictions

5. Conclusion and Future Works

With a focus on breast cancer segmentation, this study evaluated a U-NET model's performance using the recently released CDD-CESM dataset. Determining how affine data augmentations affected the model's accuracy was the main goal. In order to examine segmentation outputs without the need for fine-tuning, a baseline version was first developed. Hyperparameter tweaks were then made, changing things like batch size, learning rate, and image dimensions. It was discovered that using a decaying learning rate provided a minor improvement in results, and that utilizing smaller images (256 × 256) produced a Dice score that was about 7% higher than larger ones (512 × 512). The effects of individual and combination affine augmentations, such as rotation, shearing, flipping, and translation, were investigated in further studies. A 10° rotation showed a consistent increase in accuracy among the solo strategies in both the test and validation sets. Few of the combinations that were evaluated outperformed the baseline model, particularly the combination of y-shear with translation and a number of augmentations that were applied using the Albumentations library. According to these results, applying too much variance can lessen the overall benefit of data augmentation, even while it may improve model performance. In order to assist radiologists and increase diagnosis accuracy, this research lays the groundwork for upcoming improvements in CAD systems aimed at enhancing breast cancer identification.

Funding: This research received no external funding
Conflict of Interest: The authors declare no conflict of interest.
Data Availability: Data will be made on reasonable request.
Code Availability: Code will be made on reasonable request.

References

[1] J. A. P. Catarino, “End-to-end deep learning pipeline for breast cancer detection, segmentation and classification in contrast-enhanced spectral mammography,” M.S. thesis, Univ. de Lisboa (Portugal), 2022.

[2] L. Caselles, C. Jailin, and S. Muller, “Data augmentation for breast cancer mass segmentation,” in *Proc. Int. Conf. Medical Imaging and Computer-Aided Diagnosis*, Singapore: Springer, Mar. 2021, pp. 228–237.

[3] F. Z. Nakach, A. Idri, and A. Tsirikoglou, “Fusion of real and synthetic subtracted contrast-enhanced mammograms for enhanced tumor detection,” in *Proc. 2024 Int. Conf. Machine Learning and Applications (ICMLA)*, pp. 1736–1740, Dec. 2024.

[4] Carriero, L. Groenhoff, E. Vologina, P. Basile, and M. Albera, “Deep learning in breast cancer imaging: State of the art and recent advancements in early 2024,” *Diagnostics*, vol. 14, no. 8, p. 848, 2024.

[5] Khorshidifar, G. Mostaghel, K. Dastvareh, Y. Ahmadyar, and R. Samimi, “Generating pseudo-subtracted image in dual-energy contrast-enhanced spectral mammography using transfer learning,” 2025.

[6] Liu, T. Huang, Y. He, H. Chen, Z. Wu, and Y. Yang, “Semi-supervised medical lesion image segmentation based on a contrast-guided diffusion model,” *IEEE Trans. Radiat. Plasma Med. Sci.*, 2025.

[7] M. Sakaida, T. Yoshimura, M. Tang, S. Ichikawa, H. Sugimori, K. Hirata, and K. Kudo, "The effectiveness of semi-supervised learning techniques in identifying calcifications in X-ray mammography and the impact of different classification probabilities," *Appl. Sci.*, vol. 14, no. 14, p. 5968, 2024.

[8] M. Alshahrani, N. Alzahrani, and A. A. Alshahrani, "A Review of Machine Learning Techniques for Breast Cancer Diagnosis," *Journal of Healthcare Engineering*, vol. 2023, pp. 1-12, 2023.

[9] R. Khan, F. Khan, and S. Hussain, "Deep Learning for Medical Image Analysis: A Review," *Journal of Medical Systems*, vol. 46, no. 2, pp. 1-15, 2022.

[10] J. Smith, K. Johnson, and L. Brown, "Enhancing Image Segmentation with Transfer Learning Techniques," *Computers in Biology and Medicine*, vol. 145, p. 105293, 2022.

[11] Patel, R. Gupta, and M. Verma, "Artificial Intelligence in Radiology: Current Applications and Future Directions," *Radiology Research and Practice*, vol. 2024, pp. 1-10, 2024.

[12] N. Marwah, V. K. Singh, G. S. Kashyap, and S. Wazir, "An analysis of the robustness of UAV agriculture field coverage using multi-agent reinforcement learning," *Int. J. Inf. Technol.*, vol. 15, no. 4, pp. 2317–2327, May 2023.

[13] H. Habib, G. S. Kashyap, N. Tabassum, and T. Nafis, "Stock Price Prediction Using Artificial Intelligence Based on LSTM– Deep Learning Model," in *Artificial Intelligence & Blockchain in Cyber Physical Systems: Technologies & Applications*, CRC Press, 2023, pp. 93–99.

[14] Chaix, G. Delamon, A. Guillemassé, B. Brouard, and J. E. Bibault, "Psychological distress during the COVID-19 pandemic in France: A national assessment of at-risk populations," *Gen. Psychiatry*, vol. 33, no. 6, p. 100349, Nov. 2020.

[15] M. T. Shaban, C. Baur, N. Navab, and S. Albarqouni, "Staingan: Stain style transfer for digital histological images," in *Proceedings - International Symposium on Biomedical Imaging*, IEEE Computer Society, Apr. 2019, pp. 953–956.

[16] H. Muhammad et al., "Unsupervised subtyping of cholangiocarcinoma using a deep clustering convolutional autoencoder," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Science and Business Media Deutschland GmbH, 2019, pp. 604–612.

[17] R. H. J. M. Kurvers et al., "How to detect high-performing individuals and groups: Decision similarity predicts accuracy," *Sci. Adv.*, vol. 5, no. 11, Nov. 2019.

[18] K. Zormpas-Petridis, R. Noguera, D. K. Ivankovic, I. Roxanis, Y. Jamin, and Y. Yuan, "SuperHistopath: A Deep Learning Pipeline for Mapping Tumor Heterogeneity on Low-Resolution Whole-Slide Digital Histopathology Images," *Front. Oncol.*, vol. 10, p. 3052, Jan. 2021.

[19] G. Atteia et al., "Adaptive Dynamic Dipper Throated Optimization for Feature Selection in Medical Data," *Comput. Mater. Contin.*, vol. 75, no. 1, pp. 1883–1900, Feb. 2023.

[20] J. G. Precious and S. Selvan, "Detection of Abnormalities in Ultrasound Images Using Texture and Shape Features," in *Proceedings of the 2018 International Conference on Current Trends towards Converging Technologies, ICCTCT 2018*, Institute of Electrical and Electronics Engineers Inc., Nov. 2018.

[21] R. H. J. M. Kurvers, S. M. Herzog, R. Hertwig, J. Krause, and M. Wolf, "Pooling decisions decreases variation in response bias and accuracy," *iScience*, vol. 24, no. 7, p. 102740, 2021.

[22] R. Mishra, O. Daescu, P. Leavey, D. Rakheja, and A. Sengupta, "Convolutional neural network for histopathological analysis of osteosarcoma," in *Journal of Computational Biology*, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, Mar. 2018, pp. 313–325.

[23] J. Z. Cheng et al., "Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans," *Sci. Rep.*, vol. 6, no. 1, pp. 1–13, Apr. 2016.

[24] Carmichael et al., "JOINT AND INDIVIDUAL ANALYSIS OF BREAST CANCER HISTOLOGIC IMAGES AND GENOMIC COVARIATES," *Ann. Appl. Stat.*, vol. 15, no. 4, pp. 1697–1722, Dec. 2021.

[25] A. Adegun, S. Viriri, and R. O. Ogundokun, "Deep Learning Approach for Medical Image Analysis," *Comput. Intell. Neurosci.*, vol. 2021, 2021.