Enhancing Breast Cancer Diagnosis: a CNN-based Approach for Medical Image Segmentation and Classification

Shoffan Saifullah[1,2][0000−0001−6799−3834] and Rafał Dreżewski[1][0000−0001−8607−3478]

[1] Faculty of Computer Science, AGH University of Krakow, Krakow, Poland
{saifulla,drezew}@agh.edu.pl
[2] Department of Informatics, Universitas Pembangunan Nasional Veteran Yogyakarta, Yogyakarta, Indonesia
shoffans@upnyk.ac.id

Abstract. This study introduces a novel Convolutional Neural Network (CNN) approach for breast cancer diagnosis, which seamlessly integrates segmentation and classification. The segmentation process achieves high precision, with Jaccard Index (JI) values of 0.89, 0.92, and 0.87 for Normal, Benign, and Malignant regions, respectively, resulting in an overall JI of 0.896. Similarly, the Dice Similarity Coefficient (DSC) values are notably high, with 0.94, 0.96, and 0.92 for the corresponding regions, yielding an overall DSC of 0.943. The CNN model exhibits high accuracy, specificity, precision, recall, and F1 score across all classes, establishing its reliability for clinical applications. This research comprehensively evaluates the model's performance metrics, addressing challenges in breast cancer diagnostics and proposing an innovative CNN-based solution. Beyond immediate applications, it lays a robust foundation for future medical imaging advancements, enhancing diagnostic accuracy and patient outcomes.

Keywords: Breast Cancer Detection · Medical Image Analysis · CNN · Image Segmentation · Image Classification.

1 Introduction

Breast cancer poses a health challenge globally [5], particularly impacting women, with advancements in its understanding and treatment. However, precise and timely diagnosis remains a formidable hurdle. These challenges underscore the urgent need for advancements in breast cancer diagnosis to navigate its intricate landscape with heightened accuracy and efficiency [14].

The complexities of breast cancer diagnosis arise from its diverse manifestations and progression patterns, influenced by genetic, molecular, and environmental factors [20]. Early symptoms often present subtly, complicating timely detection. Clinical examination and medical imaging interpretation further complicate matters, leading to delayed diagnoses and missed opportunities for early intervention [12, 22, 7]. Beyond clinical complexities, breast cancer

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_15

diagnosis presents societal and healthcare challenges [25], including emotional distress from false positives and negatives [24, 4]. Healthcare systems struggle to meet diagnostic demands [6], exacerbating disparities in access to advanced resources and early detection.

Researchers worldwide explore innovative solutions, such as advanced imaging, molecular diagnostics, and AI applications like Convolutional Neural Networks (CNNs)[23]. This study introduces a CNN-based approach to breast cancer diagnosis, aiming to improve outcomes by addressing existing challenges. The article's structure includes a review of related works (Section 2), detailed methodology (Section 3), comprehensive results analysis (Section 4), and conclusions with future research directions (Section 5).

2 Related Works

Breast cancer diagnosis has witnessed transformative advancements [3, 15] with integration of Convolutional Neural Networks (CNNs) into imaging technologies [16]. CNNs have revolutionized the field by significantly enhancing accuracy and efficiency across various diagnostic processes [10]. Traditionally, interpretations of medical images [11], such as mammograms and ultrasounds, were susceptible to human subjectivity, leading to inconsistent results [13]. However, recent studies have demonstrated the power of CNNs augmenting classification accuracy by leveraging extensive datasets containing diverse breast tissue images [8]. For instance, Li et al. [9] achieved a remarkable 94.55% accuracy in distinguishing malignant from benign lesions using mammography images, significantly reducing false positives and false negatives.

Furthermore, CNNs have played a crucial role in significantly improving medical image segmentation [16], a critical aspect of breast cancer diagnosis [14, 2]. Automated segmentation ensures consistency and expedites treatment accessibility by accurately delineating tumor boundaries within medical images. Despite these advancements, challenges persist, including the subtle nature of early-stage symptoms and the complexities involved in image interpretation [16]. Innovative concepts like artificial intelligence (AI) and deep learning techniques have been pivotal in addressing these challenges [20, 12], promising more accurate and efficient breast cancer diagnosis.

In summary, CNN integration in breast cancer diagnosis marks a paradigm shift, reshaping diagnostics and improving outcomes [22]. These advances highlight CNN's transformative potential in breast cancer management.

3 Methods

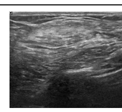
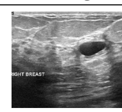
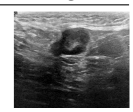
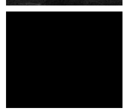
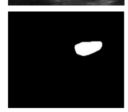
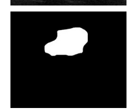
This section details our methodology, using CNNs for breast cancer diagnosis. Our approach improves accuracy and efficiency by combining image classification and segmentation. We cover data collection, preprocessing, CNN architecture, and workflow.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_15

3.1 Datasets of Breast Ultrasound Images and Preprocessing

Our dataset comprises 1,578 breast ultrasound images from Kaggle.com [21], categorized into Normal (266 images), Benign (891 images), and Malignant (421 images), each with ground truth labels (Table 1). Preprocessing steps ensured dataset consistency and readiness for CNN model training and evaluation. Standard normalization techniques and resizing were applied for uniformity, while data augmentation techniques, including rotation and flipping, enhanced model generalization [17].

Table 1. Samples of the dataset.

Category	Normal	Benign	Malignant
Breast Ultrasound Image			
Ground Truth (Mask)			

3.2 Convolutional Neural Networks (CNNs) Architecture

Our approach integrates two key components: U-Net for segmentation and CNN for classification, both pivotal for accurate breast cancer diagnosis. The U-Net architecture (Fig. 1), renowned for its effectiveness in segmentation tasks, operates with an $256 \times 256 \times 3$ input layer. It consists of a 5×5 Convolutional Block with ReLU activation, followed by Maxpooling and dropout to prevent overfitting. The bottleneck section employs a $16 \times 16 \times 1024$ feature map, which is crucial for capturing intricate details. The expansive path restores the feature map's original size, maintaining spatial details with upsampling layers and incorporating skip connections to preserve fine-grained details [1].

In parallel, the CNN classification network sequentially processes the segmented regions. Following the U-Net segmentation, the CNN operates on features extracted from the segmented images. Utilizing 5×5 convolutional blocks, the CNN discerns intricate patterns indicative of different breast cancer classes. Subsequent layers flatten the feature representation, followed by dense layers with dropout to learn complex relationships within the data. A softmax activation assigns probability scores to each class (Normal, Benign, Malignant), guided by categorical cross-entropy loss [2].

This integrated architecture optimizes segmentation and classification by leveraging the strengths of U-Net [19, 18] for precise delineation and CNN for accurate classification. Our commitment to reliable breast cancer diagnosis is evident in this comprehensive approach, promising better patient outcomes.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_15

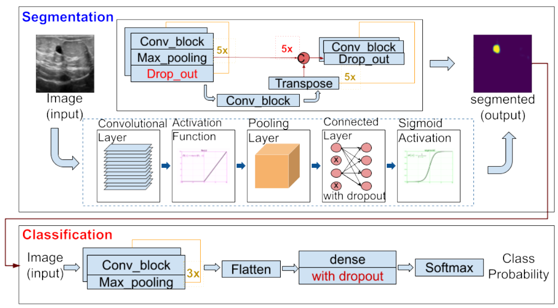


Fig. 1. CNN architecture used in the proposed approach.

3.3 Proposed Method

Our breast cancer detection method integrates segmentation and classification, inspired by [17, 18]. We preprocess breast cancer images to ensure dataset uniformity and quality, laying the groundwork for model training. Our approach utilizes a modified U-Net for lesion segmentation, coupled with a CNN classifier trained to identify malignancy patterns within segmented regions. Quantitative evaluation measures segmentation accuracy using Dice Similarity Coefficient (DSC) and Jaccard Index (JI). Comprehensive validation assesses accuracy, precision, recall, and F1-score metrics, distinguishing between benign and malignant cases. Model outputs, including segmentation and classification results, are visually interpreted to understand discriminative features aiding accurate diagnosis. The proposed method undergoes validation with real-world clinical data to ensure clinical relevance, aligning predictions with actual outcomes.

4 Results and Discussion

This section presents the performance of our CNN model in segmenting and classifying breast cancer images. We analyze the results, discuss the implications, and compare them with existing approaches in the field.

4.1 Breast Cancer Segmentation Results

Our proposed CNN-based model was meticulously evaluated to ensure its efficacy in segmenting breast cancer images. The training process, depicted in Fig. 2(a), illustrates convergence at epoch 1000, with high accuracy (0.9875) and minimal loss (0.0281) indicating optimal performance. Successful assessment using DSC (0.9063) and JI (0.8307) metrics confirm the model's segmentation efficacy, as illustrated in Fig. 2(b).

Performance metrics, detailed in Table 2, underscored the precision of our model in segmenting breast cancer regions. The DSC and JI scores, including 0.94 (DSC) and 0.89 (JI) for Normal, 0.96 (DSC) and 0.92 (JI) for Benign,

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_15

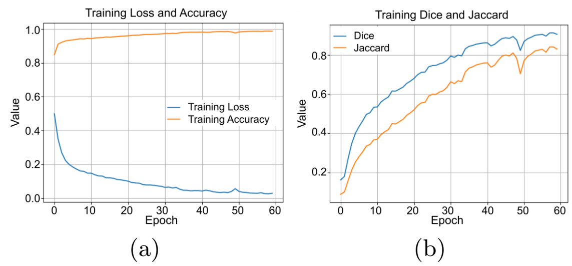


Fig. 2. Training performance of CNN: (a) model approach evaluated on accuracy and loss function, and (b) segmentation results assessed by DSC and JI.

Table 2. Evaluation metrics for segmented classes and overall performance.

No	Segmentation Metrics	Predicted			Overal
		Normal	Benign	Malignant	
1	Jaccard Index (JI)	0.89	0.92	0.87	0.896
2	Dice Similarity Coefficients (DSC)	0.94	0.96	0.92	0.943

and 0.92 (DSC) and 0.87 (JI) for Malignant regions, indicated a high degree of accuracy in delineating tissue classes. Furthermore, the consolidated DSC of 0.943 and JI of 0.896 affirmed the model's balanced accurate segmentation.

Fig. 3 shows the segmentation outcomes, demonstrating the model's accuracy in isolating distinct regions within breast images. This visualization aids in assessing segmentation accuracy and identifying areas for further analysis. The model's ability to delineate cancerous and non-cancerous regions enhances diagnostic reliability and treatment planning.

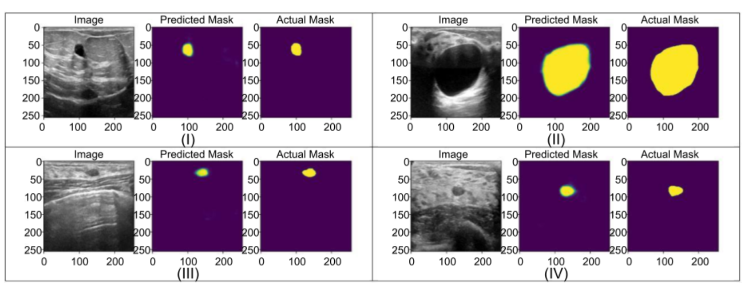


Fig. 3. Segmentation testing results in breast images, comparing model delineations with Ground Truth Masks for accuracy assessment.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_15

4.2 Breast Cancer Classification Results

The breast cancer classification results demonstrate the exceptional performance of our CNN model in accurately categorizing breast tissue into Normal, Benign, and Malignant classes. Fig. 4 depicts the model's training progress, showcasing the convergence of accuracy to near perfect levels and minimal loss, indicating successful adaptation to the dataset's complexities. The high accuracy of 0.9972 and low loss of 0.0022 highlight the model's proficiency in learning intricate tissue features during training.

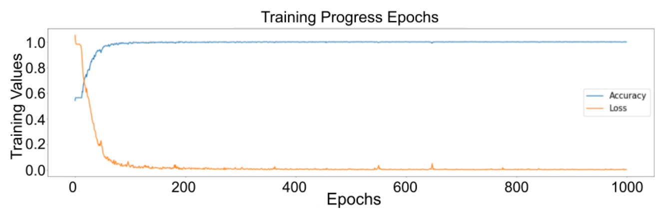


Fig. 4. Training progress depicting breast cancer classification performance, accuracy and loss metrics, showcasing model learning dynamics and convergence.

The model underwent testing with a separate dataset to evaluate generalization capabilities, following an 80%:20% split. Table 3 presents the testing phase's summary. With an accuracy of 0.896 and specificity of 0.943, the model effectively classifies breast cancer cases while minimizing false positives, essential for clinical reliability.

Table 3. Breast cancer classification efficacy metrics.

No	Segmentation Metrics	Predicted			Overall
		Normal	Benign	Malignant	
1	Accuracy	0.89	0.92	0.87	0.896
2	Specificity	0.94	0.96	0.92	0.943
3	Precision	1.0	0.9333	1.0	0.9778
4	Recall	0.9286	0.9333	0.85	0.904
5	F1-Score	0.963	0.9333	0.9189	0.9387

Precision values indicate the reliability of the model's predictions, with a near-perfect precision of 1.0 for the Benign class, ensuring confidence in classifying benign cases. However, there is a minor scope for improvement in the recall for the Malignant class, suggesting opportunities for enhancing the model's ability to capture all true positive instances of malignant cases. The F1-score of 0.9387 signifies a balanced assessment, reflecting the model's effectiveness in making accurate predictions across all classes and ensuring robust performance in breast cancer classification.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_15

5 Conclusions

This study demonstrates the efficacy of a CNN for breast cancer diagnosis, integrating medical image segmentation and classification. Precise segmentation identifies Normal, Benign, and Malignant tissues, while classification validates robustness across all classes. It contributes to medical image analysis by offering a reliable framework for accurate diagnosis, enhancing clinical decision-making. Future research may explore techniques (dataset split included) to improve generalization capabilities, advancing breast cancer diagnosis using deep learning.

Acknowledgement. This research was supported by the Polish Ministry of Science and Higher Education funds assigned to AGH University of Krakow and by PLGrid under grant no. PLG/2023/016757.

References

1. Alshehri, M.: Breast Cancer Detection and Classification Using Hybrid Feature Selection and DenseXtNet Approach. Mathematics **11**(23), 4725 (2023). https://doi.org/10.3390/math11234725
2. Balasubramaniam, S., Velmurugan, Y., Jaganathan, D., Dhanasekaran, S.: A Modified LeNet CNN for Breast Cancer Diagnosis in Ultrasound Images. Diagnostics **13**(17), 2746 (2023). https://doi.org/10.3390/diagnostics13172746
3. Chopra, S., Khosla, M., Vidya, R.: Innovations and Challenges in Breast Cancer Care: A Review. Medicina **59**(5), 957 (2023). https://doi.org/10.3390/medicina59050957
4. Gadaleta, E., Thorn, G.J., Ross-Adams, H., Jones, L.J., Chelala, C.: Field cancerization in breast cancer. The Journal of Pathology **257**(4), 561–574 (2022). https://doi.org/10.1002/path.5902
5. Hamdy, S., Nye, C.: Comics and revolution as global public health intervention: The Case of Lissa. Global Public Health **17**(12), 4056–4076 (2022). https://doi.org/10.1080/17441692.2019.1682632
6. Hunleth, J., Steinmetz, E.: Navigating Breast Cancer Screening in Rural Missouri: From Patient Navigation to Social Navigation. Medical Anthropology **41**(2), 228–242 (2022). https://doi.org/10.1080/01459740.2021.2015347
7. Jaiswal, V., Suman, P., Bisen, D.: An improved ensembling techniques for prediction of breast cancer tissues. Multimedia Tools and Applications (2023). https://doi.org/10.1007/s11042-023-16949-8
8. Labrada, A., Barkana, B.D.: A Comprehensive Review of Computer-Aided Models for Breast Cancer Diagnosis Using Histopathology Images. Bioengineering **10**(11), 1289 (2023). https://doi.org/10.3390/bioengineering10111289
9. Li, H., Zhuang, S., Li, D.a., Zhao, J., Ma, Y.: Benign and malignant classification of mammogram images based on deep learning. Biomedical Signal Processing and Control **51**, 347–354 (may 2019). https://doi.org/10.1016/j.bspc.2019.02.017
10. Liew, X.Y., Hameed, N., Clos, J.: A Review of Computer-Aided Expert Systems for Breast Cancer Diagnosis. Cancers **13**(11), 2764 (2021). https://doi.org/10.3390/cancers13112764

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_15

11. Lozano, A., Hassanipour, F.: Infrared imaging for breast cancer detection: An objective review of foundational studies and its proper role in breast cancer screening. Infrared Physics & Technology **97**, 244–257 (2019). https://doi.org/10.1016/j.infrared.2018.12.017
12. Najjar, R.: Redefining Radiology: A Review of Artificial Intelligence Integration in Medical Imaging. Diagnostics **13**(17), 2760 (2023). https://doi.org/10.3390/diagnostics13172760
13. Ortiz, M.M., Andrechek, E.R.: Molecular Characterization and Landscape of Breast cancer Models from a multi-omics Perspective. Journal of Mammary Gland Biology and Neoplasia **28**(1), 12 (2023). https://doi.org/10.1007/s10911-023-09540-2
14. Panico, A., Gatta, G., Salvia, A., Grezia, G.D., Fico, N., Cuccurullo, V.: Radiomics in Breast Imaging: Future Development. Journal of Personalized Medicine **13**(5), 862 (2023). https://doi.org/10.3390/jpm13050862
15. Pulumati, A., Pulumati, A., Dwarakanath, B.S., Verma, A., Papineni, R.V.L.: Technological advancements in cancer diagnostics: Improvements and limitations. Cancer Reports **6**(2) (2023). https://doi.org/10.1002/cnr2.1764
16. Rahman, H., Naik Bukht, T.F., Ahmad, R., Almadhor, A., Javed, A.R.: Efficient Breast Cancer Diagnosis from Complex Mammographic Images Using Deep Convolutional Neural Network. Computational Intelligence and Neuroscience **2023**, 1–11 (2023). https://doi.org/10.1155/2023/7717712
17. Saifullah, S., Drezewski, R.: Modified Histogram Equalization for Improved CNN Medical Image Segmentation. Procedia Computer Science **225**(C), 3020–3029 (2023). https://doi.org/10.1016/j.procs.2023.10.295
18. Saifullah, S., Suryotomo, A.P., Dreżewski, R., Tanone, R., Tundo, T.: Optimizing brain tumor segmentation through CNN U-Net with CLAHE-HE image enhancement. In: Proceedings of the 2023 1st International Conference on Advanced Informatics and Intelligent Information Systems (ICAI3S 2023). pp. 90–101. Atlantis Press (2024). https://doi.org/10.2991/978-94-6463-366-5_9
19. Saifullah, S., Yuwono, B., Rustamaji, H.C., Saputra, B., Dwiyanto, F.A., Drezewski, R.: Detection of Chest X-ray Abnormalities Using CNN Based on Hyperparameters Optimization. Engineering Proceedings **52**, 1–7 (2023). https://doi.org/10.3390/ASEC2023-16260
20. Sebastian, A.M., Peter, D.: Artificial Intelligence in Cancer Research: Trends, Challenges and Future Directions. Life **12**(12), 1991 (2022). https://doi.org/10.3390/life12121991
21. Shah, A.: Breast Ultrasound Images Dataset (2020), https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset
22. Singh, N., Srivastava, M., Srivastava, G.: Enhancing the Deep Learning-Based Breast Tumor Classification Using Multiple Imaging Modalities: A Conceptual Model. Communications in Computer and Information Science **1546**, 329–353 (2022). https://doi.org/10.1007/978-3-030-95711-7_29
23. Ting, F.F., Tan, Y.J., Sim, K.S.: Convolutional neural network improvement for breast cancer classification. Expert Systems with Applications **120**, 103–115 (2019). https://doi.org/10.1016/j.eswa.2018.11.008
24. Topol, E.: Deep medicine: how artificial intelligence can make healthcare human again. Hachette UK (2019)
25. Travado, L., Rowland, J.H.: Supportive Care and Psycho-oncology Issues During and Beyond Diagnosis and Treatment. In: Breast Cancer in Young Women, pp. 197–214. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-24762-1_17

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_15