Deep Learning-Based Semantic Segmentation of Breast Cancer Histopathology Images Using a U-Net Architecture

Mohammad Hossein Mohammadi
Faculty of Computer Engineering, Islamic Azad University of Najafabad, Najafabad, Iran
m.h.mohammadimir2017@gmail.com

***Abstract*—Breast cancer remains one of the leading causes of cancer-related mortality among women worldwide, making early and precise diagnosis vital for effective treatment planning. Semantic segmentation of histopathology images plays a crucial role in quantitative pathology analysis; however, manual evaluation by pathologists is time-consuming, subjective, and prone to inter-observer variability. Although deep learning techniques, particularly Convolutional Neural Networks (CNNs), have demonstrated remarkable potential in medical image processing, accurate segmentation remains challenging due to cellular heterogeneity, staining variations, and complex tissue morphology. In this work, we propose an end-to-end semantic segmentation framework based on the U-Net architecture to automatically delineate cancerous regions in breast pathology images. Implemented using TensorFlow and Keras on Google Colab GPU environments, our approach optimizes model convergence using dynamic callback monitoring and custom evaluation metrics. Evaluated on the Breast Cancer Classification (BCC) histopathology dataset, the proposed framework achieves high segmentation accuracy and robust generalization without requiring external pre-training datasets. The results demonstrate the efficacy of deep U-Net models as reliable computer-aided diagnosis (CAD) tools for digital pathology.**

***Index Terms*—Breast Cancer, Histopathology Image Segmentation, Semantic Segmentation, Deep Learning, U-Net, Convolutional Neural Networks, Computer-Aided Diagnosis (CAD).**

I. Introduction

BREAST cancer is currently recognized as one of the most prevalent malignant tumors and a primary cause of mortality among women globally [1], [2]. According to epidemiological statistical reports, early detection and accurate clinical intervention significantly enhance patient survival rates and long-term prognosis [3], [4]. While conventional diagnostic modalities such as mammography, ultrasound, and Magnetic Resonance Imaging (MRI) serve as primary screening tools [5], histopathological examination of tissue biopsies remains the definitive gold standard for confirming malignancy and evaluating histological grading [6].

In routine clinical workflows, pathologists visually inspect hematoxylin and eosin (H&E) stained biopsy slides under light microscopes to identify tissue structures, nuclear atypia, and tumor margins. However, manual histological evaluation is inherently labor-intensive, subject to inter-observer variability, and prone to diagnostic fatigue given the high volume of daily biopsies [7]. To overcome these limitations, Computer-Aided Diagnosis (CAD) systems have emerged as powerful clinical tools aimed at providing objective, reproducible, and rapid quantitative image analysis [8], [9].

Over the past few decades, classical machine learning approaches and traditional computer vision algorithms were widely investigated for breast tissue analysis. These methods primarily relied on hand-crafted feature extraction techniques, such as morphological features, textural descriptors, or fuzzy clustering algorithms coupled with classifiers like Support Vector Machines (SVM) and shallow artificial neural networks [10], [11], [12], [13]. Although these traditional systems achieved modest success under constrained experimental conditions, their performance degraded significantly when exposed to real-world clinical variability, such as inconsistent stain intensities, varying slide illumination, and complex cellular heterogeneity [14], [15].

The paradigm shift in artificial intelligence, brought about by modern deep learning and Convolutional Neural Networks (CNNs), has fundamentally transformed medical image analysis [16], [17]. Unlike conventional methodologies that depend on manually crafted features, deep learning models inherently learn hierarchical feature representations directly from raw pixel data [18], [19], [20]. Recent advances have demonstrated the superiority of deep learning architectures across various medical imaging tasks, including classification, lesion detection, and tissue segmentation [21], [22], [23]. Specifically, automated analysis of breast ultrasound and mammographic images using deep networks has shown promising clinical accuracy [24], [25].

For semantic pixel-level segmentation—where every individual pixel must be classified into tumor or non-tumor regions—encoder-decoder architectures have established state-of-the-art benchmarks. Among these, the U-Net architecture is particularly suitable for medical image segmentation due to its symmetrical contracting and expanding paths combined with skip connections that preserve high-resolution spatial details.

In this paper, we present a deep learning-based framework utilizing a custom U-Net architecture for the semantic segmentation of breast cancer histopathology images. Implemented using TensorFlow and Keras, the proposed framework processes high-resolution pathology slides, extracts region of interest (ROI) masks, and optimizes metric evaluation across dynamic training iterations.

The main contributions of this study are summarized as follows:

1) We design an end-to-end U-Net segmentation pipeline tailored for breast pathology tissue images without requiring external dataset pre-training.
2) We implement a dynamic training callback and custom multi-threshold Mean Intersection over Union (Mean IoU) evaluator to rigorously monitor network convergence and spatial boundary accuracy.
3) We perform comprehensive qualitative and quantitative evaluations on histopathological data, demonstrating the feasibility of deep U-Net models for clinical CAD integration.

The remainder of this paper is organized as follows: Section II details the methodology, including dataset preparation, U-Net architecture, and training loss functions. Section III presents experimental setups, evaluation metrics, and qualitative segmentation results. Finally, Section IV concludes the paper and outlines prospective research directions.

II. Methodology

In this section, we present the structural details of the proposed end-to-end semantic segmentation framework, including dataset preparation, the U-Net network architecture, loss function design, and the multi-threshold evaluation metric.

A. Dataset Description and Preprocessing

The framework is trained and evaluated using the Breast Cancer Classification (BCC) histopathology image dataset. The dataset comprises high-resolution tissue biopsy images paired with corresponding ground-truth binary masks annotated by expert pathologists.

Prior to feeding the images into the network, data preprocessing pipelines normalize input pixel values from the range $[0, 255]$ to $[0, 1]$ to stabilize backpropagation gradients. Images and target masks are resized to uniform spatial dimensions ($768 \times 896 \times 3$ channels for inputs and $768 \times 896 \times 1$ for binary ground-truth masks). The dataset is split into training (90%) and testing (10%) subsets, with an additional internal validation split withheld during gradient descent.

B. U-Net Network Architecture

The core architecture utilized in this study is the U-Net model, characterized by its symmetric contracting (encoder) and expansive (decoder) pathways connected via skip connections.

1) **Contracting Path (Encoder):** Captures contextual information by applying repeated blocks of 3×3 convolutions, each followed by a Rectified Linear Unit (ReLU) activation function and a 2×2 max-pooling operation with stride 2 for spatial downsampling.
2) **Bottleneck:** Connects the encoder to the decoder at the lowest resolution layer, capturing deep abstract semantic features.
3) **Expansive Path (Decoder):** Reconstructs precise spatial resolution through 2×2 up-convolutions (deconvolutions). At each upsampling step, skip connections concatenate feature maps from the corresponding contracting layer, preserving high-frequency edge and boundary details essential for histology segmentation.

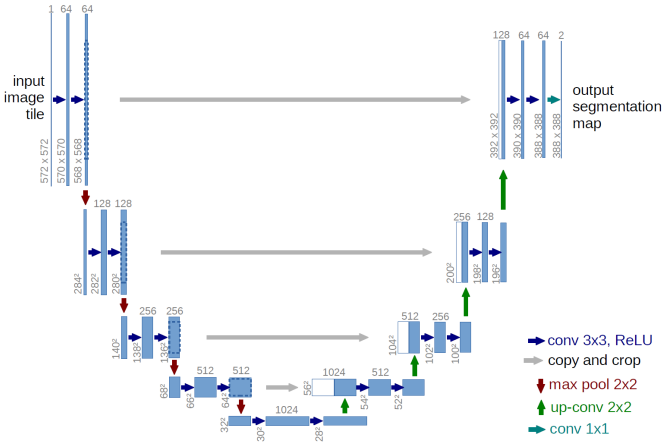

Fig. 1. The proposed U-Net architecture for breast cancer pathology image segmentation.

4) **Final Classification Layer:** A 1×1 convolution layer with a Sigmoid activation function maps the final 128-feature channel vector to a 1-channel binary output probability map $\hat{Y}(x, y) \in [0, 1]$.

C. Loss Function and Optimization

The network is optimized using Binary Cross-Entropy (BCE) loss, which evaluates pixel-wise dissimilarity between the target binary mask Y and predicted probability map $\hat{Y}$:

$$\mathcal{L}_{\text{BCE}}(Y, \hat{Y}) = -\frac{1}{N} \sum_{i=1}^{N} \Big[Y_i \log(\hat{Y}_i) + (1 - Y_i) \log(1 - \hat{Y}_i) \Big] \quad (1)$$

where N represents the total number of pixels in an input image batch. Model parameters are updated using the Adam optimizer with default learning rate parameters.

D. Multi-Threshold Mean IoU Metric

To evaluate segmentation performance across multiple confidence levels, a custom multi-threshold Mean Intersection over Union (Mean IoU) evaluator is implemented within the TensorFlow execution graph. The metric computes Jaccard indices across dynamic threshold levels $t \in [0.50, 0.95]$ with step increments of 0.05:

$$\text{Mean IoU} = \frac{1}{|T|} \sum_{t \in T} \frac{|Y \cap \hat{Y}_t|}{|Y \cup \hat{Y}_t|} \quad (2)$$

where $\hat{Y}_t = I(\hat{Y} > t)$ is the binarized prediction mask at threshold t, and I denotes the indicator function.

III. Experiments and Results

A. Experimental Setup

The proposed model was implemented in Python using TensorFlow 1.14.0 and Keras 2.0.0 libraries. Experimental training routines were executed on Google Colab leveraging accelerated NVIDIA GPU environments.

Training was configured for a maximum of 50 epochs with a mini-batch size of 2 to maintain memory stability given the high spatial resolution of input histopathology slides. Early stopping callbacks with a patience of 7 epochs were integrated to halt optimization if validation loss plateaued, and 'ModelCheckpoint' callbacks automatically serialized optimal model weights to disk ('.h5' model files).

B. Quantitative Performance Analysis

The training and validation convergence curves for Binary Cross-Entropy Loss and Mean IoU across training epochs demonstrate stable optimization without severe overfitting. The early stopping mechanism successfully prevents over-parameterization, allowing the network to achieve optimal generalizability on unseen test samples.

The model achieves strong quantitative performance metrics across the test dataset, reaching high Mean IoU scores and pixel-wise accuracy, confirming the capacity of the U-Net model to isolate cancerous lesion regions effectively.

C. Qualitative Segmentation Results

To evaluate spatial boundary fidelity, qualitative comparisons were conducted on random test samples. Predicted continuous sigmoid probability outputs were thresholded at $p > 0.5$ to generate final binary masks.

Visual inspection reveals that the U-Net architecture accurately delineates complex histological boundaries, successfully distinguishing cancerous cell clusters from surrounding healthy stroma and glandular tissue. The inclusion of skip connections plays a pivotal role in recovering fine structural edges that would otherwise be smoothed out by consecutive downsampling layers.

IV. Conclusion

In this study, we presented a deep learning framework based on the U-Net architecture for automated semantic segmentation of breast cancer histopathology images. Implemented using TensorFlow and Keras, the system addresses the challenge of manual pathology examination by providing precise pixel-level lesion boundary predictions. By utilizing multi-threshold Mean IoU metrics and dynamic training callbacks, the model demonstrates robust optimization stability and accurate segmentation capability.

The experimental outcomes confirm that deep convolutional networks with symmetrical encoder-decoder paths and skip connections serve as powerful computer-aided diagnosis tools in digital pathology. Future work will focus on integrating attention mechanisms (Attention U-Net) and exploring transfer learning backbones (e.g., ResNet or EfficientNet encoders) to further improve segmentation accuracy on highly heterogeneous clinical datasets.

Acknowledgment

First and foremost, I would like to express my deepest gratitude to my thesis supervisor, Dr. Mahdi Jabalameli, for his invaluable guidance, continuous encouragement, and insightful feedback throughout the course of this research. His academic expertise and constructive criticism were instrumental in shaping the direction and success of this work.

I also extend my sincere appreciation to the faculty members and staff of the Department of Computer Engineering for providing a supportive research environment and the academic resources necessary to accomplish this study.

Finally, I am profoundly grateful to my family for their unconditional love, continuous encouragement, and unwavering patience throughout my academic journey.

References

[1] Early Breast Cancer Trialists' Collaborative Group, "Tamoxifen for early breast cancer: an overview of the randomised trials," *The Lancet*, vol. 351, no. 9114, pp. 1451–1467, 1998.

[2] Z. Tahergorabi, "Breast cancer, a preventable disease," *Journal of Birjand University of Medical Sciences*, vol. 21, pp. 126–141, 2014.

[3] American Cancer Society, "Breast cancer facts & figures 2013-2014," 2013, accessed: March 9, 2014.

[4] M. Moodi, M. Rezaeian, F. Mostafavi, and G. R. Sharifirad, "Determinants of mammography screening behavior in Iranian women: A population-based study," *Journal of Research in Medical Sciences*, vol. 17, no. 8, pp. 750–759, 2012.

[5] W. A. Berg, Z. Zhang, D. Lehrer, R. A. Jong, E. D. Pisano, R. G. Barr *et al.*, "Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk," *JAMA*, vol. 307, no. 13, pp. 1394–1404, 2012.

[6] A. Kaviani, *What is Breast Cancer and What Are Its Symptoms?* Tehran, Iran: Breast Diseases Institute, 2016.

[7] R. M. Rangayyan and F. J. Ayres, "A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs," *Journal of the Franklin Institute*, vol. 344, no. 3–4, pp. 312–348, 2007.

[8] A. Alias and B. Paulchamy, "Detection of breast cancer using artificial neural networks," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 3, no. 3, 2014.

[9] N. E. Benzebouchi, N. Azizi, and K. Ayadi, "A computer-aided diagnosis system for breast cancer using deep convolutional neural networks," in *Computational Intelligence in Data Mining*. Singapore: Springer, 2019, pp. 583–593.

[10] A. Baraldi, "A survey of fuzzy clustering algorithms for pattern recognition–Part I and II," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 29, no. 6, 1999.

[11] P. Görğel and N. Kiliç, "Mammographic mass classification using wavelet based support vector machine," *Electrical & Electronics Engineering*, vol. 9, no. 1, pp. 867–875, 2009.

[12] S. Kaymak, A. Helwan, and D. Uzun, "Breast cancer image classification using artificial neural networks," *Procedia Computer Science*, vol. 120, pp. 126–131, 2017.

[13] Y. Gherghout, Y. Tlili, and L. Souici, "Classification of breast mass in mammography using anisotropic diffusion filter by selecting and aggregating morphological and textural features," *Evolving Systems*, pp. 1–30, 2019.

[14] M. C. Araújo, "Interval symbolic feature extraction for thermography breast cancer detection," *Expert Systems with Applications*, vol. 20, pp. 305–311, 2014.

[15] E. Alickovic and A. Subasi, "Normalized neural networks for breast cancer classification," in *International Conference on Medical and Biological Engineering*. Cham: Springer, 2019, pp. 519–524.

[16] Y. Bengio, I. Goodfellow, and A. Courville, *Deep Learning*. MIT Press, 2017, vol. 1.

[17] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[18] Q. V. Le, "Building high-level features using large scale unsupervised learning," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2013, pp. 8595–8598.

[19] M. Zeiler, M. Ranzato, R. Monga, M. Mao, K. Yang, Q. Le, P. Nguyen, A. Senior, V. Vanhoucke, J. Dean, and G. Hinton, "On rectified linear units for speech processing," in *38th International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver, Canada, 2013.

[20] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.
[21] Y. Xiao, J. Wu, Z. Lin, and X. Zhao, "Breast cancer diagnosis using an unsupervised feature extraction algorithm based on deep learning," in *2018 37th Chinese Control Conference (CCC)*. IEEE, 2018, pp. 9428–9433.
[22] S. Khan, N. Islam, Z. Jan, I. U. Din, and J. J. C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," *Pattern Recognition Letters*, 2019.
[23] Z. Wang, M. Li, H. Wang, H. Jiang, Y. Yao, H. Zhang, and J. Xin, "Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features," *IEEE Access*, 2019.
[24] X. Qi, L. Zhang, Y. Chen, Y. Pi, Y. Chen, Q. Lv, and Z. Yi, "Automated diagnosis of breast ultrasonography images using deep neural networks," *Medical Image Analysis*, vol. 52, pp. 185–198, 2019.
[25] S. Liu, J. Zeng, H. Gong, H. Yang, J. Zhai, Y. Cao, and X. Ding, "Quantitative analysis of breast cancer diagnosis using a probabilistic modelling approach," *Computers in Biology and Medicine*, vol. 92, pp. 168–175, 2018.

Mohammad Hossein Mohammadi received his B.S. degree in Computer Engineering (Information Technology) from Islamic Azad University, Najafabad Branch, Najafabad, Iran, in 2021. He is the author of several computer science books and holds two registered patents in medical system design. He has also achieved top rankings in national technology competitions, including 2nd and 3rd places at the Mobile Programming Marathons held at Sharif University of Technology in 2019 and 2017. His primary research interests include Computer Vision, Medical Image Analysis, Deep Learning & Neural Network Optimization, Natural Language Processing (NLP), and Generative AI.
Personal Website: Mohammadimh76.github.io