Available online at website: https://jurnal.iaii.or.id/index.php/RESTI

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 10 No. 3 (2026) 787 - 793 e-ISSN: 2580-0760

ResNet34-Encoded U-Net with Transfer Learning for Breast Cancer Cell Segmentation Under Data Scarcity

Ronal Watrianthos[1], Yuhefizar[2], Rayendra[3], Ervan Asri[4]
[1,2,3,4]Department of Information Technology, Politeknik Negeri Padang, Padang, Indonesia

[*]*ronalwatrianthos@pnp.ac.id*

Abstract

Histopathological examination of hematoxylin and eosin (H&E)-stained tissue remains the clinical gold standard for breast cancer diagnosis; however, manual cell segmentation is labor-intensive, subjective, and prone to substantial inter-observer variability. Deep learning-based segmentation models have demonstrated considerable promise in computational pathology, yet their performance under conditions of severe data scarcity remains insufficiently characterized. This study proposes and evaluates a ResNet34-encoded U-Net trained via transfer learning as a solution to this challenge, benchmarking it against a vanilla U-Net trained from random initialization as a controlled baseline. The experimental dataset comprises 42 H&E-stained breast cancer whole-slide image patches acquired at Qingdao Central Hospital (2019–2022), partitioned into 34 training and 8 test images. To mitigate the severe class imbalance inherent in the dataset — in which cancerous cell regions constitute only 16.2% of image pixels on average — a weighted Binary Cross-Entropy and Dice composite loss function was employed alongside a differential learning rate strategy to preserve pretrained encoder representations during fine-tuning. Data augmentation was applied to the training set to improve generalization under extreme data constraint. Evaluation was conducted using the Dice Similarity Coefficient (DSC) and Intersection over Union (IoU) as primary metrics. The ResNet34-UNet achieved a DSC of 0.6071 and IoU of 0.4392, outperforming the vanilla U-Net (DSC: 0.5554, IoU: 0.4004) by 9.3% and 9.7% respectively, while converging 37.5% faster. These findings demonstrate that ImageNet-pretrained encoder features transfer effectively to H&E pathology domains even under extreme data constraints, providing a reproducible and computationally efficient baseline pipeline for automated cell segmentation in resource-limited clinical settings.

Keywords: breast cancer; deep learning; ResNet34; U-Net

How to Cite: Ronal Watrianthos, Y. Yuhefizar, R. Rayendra, and E. Asri, "ResNet34-Encoded U-Net with Transfer Learning for Breast Cancer Cell Segmentation Under Data Scarcity", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 10, no. 3, pp. 787 - 793, Jun. 2026.
*Permalink/*DOI: https://doi.org/10.29207/resti.v10i3.7661

Received: March 22, 2026
Accepted: June 15, 2026
Available Online: June 23, 2026

This is an open-access article under the CC BY 4.0 License
Published by Ikatan Ahli Informatika Indonesia

1. Introduction

Deep learning has fundamentally transformed the landscape of computational pathology over the past decade [1]. Convolutional neural networks (CNNs) [2], and in particular the family of encoder–decoder architectures descended from the seminal U-Net framework[3], have established new state-of-the-art benchmarks in biomedical image segmentation, enabling pixel-level identification of tissues, cells, and subcellular structures across diverse cancer types [4-7].

In breast cancer pathology [8-10], deep learning models have demonstrated remarkable capacity to automate nuclear segmentation, tumor grading, receptor status prediction, and prognosis estimation from hematoxylin and eosin (H&E)-stained whole-slide images (WSIs) [11]. The proliferation of whole-slide imaging scanners and digital pathology platforms has further catalyzed this progress, generating repositories of digitized pathological data at unprecedented scale and enabling large-scale supervised learning workflows that were previously infeasible [12]. Taken together, these developments have created a compelling vision for AI-assisted clinical pathology: faster, more reproducible, and potentially more accurate than conventional manual assessment [13], [14].

Despite the demonstrated promise of deep learning for histopathological image segmentation, its clinical translation remains constrained by a fundamental practical barrier: the scarcity of annotated training data. Expert manual annotation of H&E-stained slides demands substantial pathologist time, specialized domain knowledge, and access to digitized whole-slide imaging infrastructure — resources that are unevenly

distributed across institutions and largely unavailable in resource-limited clinical settings.

The resulting annotation bottleneck means that most institutions can realistically produce datasets of tens, rather than hundreds or thousands, of labeled image patches. Compounding this, inter-observer variability in nucleus delineation has been reported to exceed 30% on challenging cases, undermining the reliability of ground-truth labels even when annotation is feasible.

As a direct consequence of this annotation bottleneck, most published deep learning segmentation models have been developed and validated on datasets comprising hundreds to thousands of annotated images [7]. This reliance on large, well-curated datasets has produced models whose reported performance figures do not generalize to the genuinely data-scarce conditions that characterize routine institutional practice.

A further compounding challenge arises from the intrinsic class distribution of binary cell segmentation tasks: cancerous cellular regions typically occupy only a small fraction of the total image area, with foreground pixel ratios as low as 2–16% in clinical H&E datasets. When trained with standard cross-entropy loss functions, models exposed to such imbalanced distributions are systematically biased toward predicting the dominant background class, yielding inflated pixel-level accuracy scores while producing poor foreground recall and severely degraded segmentation boundaries[15].

Most existing studies employ large dataset such as BreaKHis [16], [17], MoNuSeg [18], [19], or CoNSeP [20], that do not accurately reflect the characteristics of institutional clinical datasets collected under routine pathological workflows [21]. Evaluation protocols in small-dataset regimes are frequently inadequate because single train–test splits without patient-level stratification inflate performance estimates and obscure true model variance. These gaps mean that the current evidence base provides insufficient guidance for deploying automated segmentation in the resource-constrained settings where such tools are most urgently needed.

The research question is what extent does ImageNet-pretrained encoder initialization improve cell segmentation performance relative to random initialization under conditions of extreme data scarcity and quantify the individual contribution of each proposed design component to the overall segmentation performance improvement via a sequential ablation study.

The primary contribution of this paper is a systematic study that rigorously evaluates the performance advantages of a pre-trained ResNet34 encoder on ImageNet compared to a standard U-Net baseline on a clinical H&E dataset consisting of only 42 labeled WSI patches. We propose and evaluate a ResNet34-encoded U-Net trained via transfer learning, benchmarking it against a vanilla U-Net trained from random initialization as a controlled baseline on a clinical dataset of 42 H&E-stained breast cancer whole-slide image patches.

This comparison provides empirically grounded quantitative evidence of the performance gain attributable to transfer learning in a genuinely data-scarce regime, filling a gap that previous work, which has largely relied on large public benchmarks, has left unaddressed [7].

2. Methods

The study is organized into five sequential stages as illustrated in Figure 1. Stage 1 involves the curation and statistical analysis of the raw clinical dataset. The dataset comprises 42 H&E-stained breast tissue image patches extracted from digitized whole-slide scans acquired at the Department of Pathology, Qingdao Central Hospital, between 2019 and 2022. Each image patch has a native resolution of 672 × 672 pixels in RGB format. Binary segmentation masks, annotated by expert pathologists, assign pixel value 255 to cancerous cell regions and 0 to background.

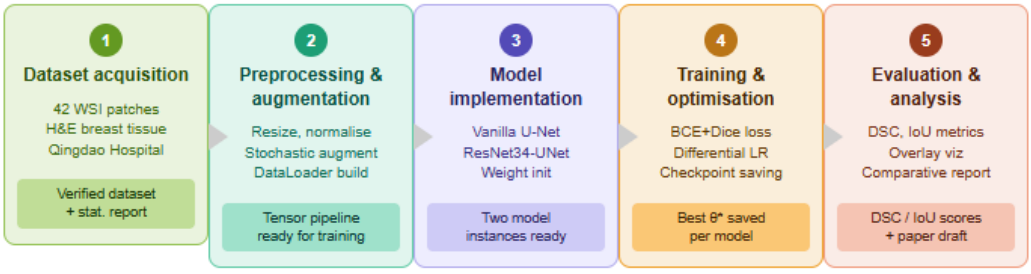


Figure 1. Research Pipeline

The dataset was partitioned into a training set D_train of 34 images drawn from 23 unique slides, and a test set D_test of 8 images each drawn from a distinct slide, yielding an approximately 81%/19% split with strict slide-level separation to prevent data leakage. Cell coverage — defined as the proportion of foreground pixels per mask — ranges from 2.26% to 39.32% across training images (mean: 16.31%, SD: 7.12%) and from 11.85% to 23.65% across test images (mean: 16.18%), confirming a pronounced and consistent foreground-to-background class imbalance across both partitions.

Stage 2 transforms raw image–mask pairs into normalized, augmented tensors ready for deep learning training. All images and masks were resized to 256 × 256 pixels and normalized using ImageNet channel statistics (mean = [0.485, 0.456, 0.406]; SD = [0.229, 0.224, 0.225]) to maintain compatibility with the pretrained ResNet34 encoder. To mitigate overfitting under extreme data scarcity, the training set was augmented using the following transformations applied stochastically during each training epoch: random horizontal and vertical flipping, random rotation (±30°), random brightness and contrast jitter, and random Gaussian blur.

Stage 3 instantiates both candidate architectures, vanilla U-Net with Kaiming uniform initialisation, and ResNet34-UNet with ImageNet-pretrained encoder weights. Two architectures were instantiated and compared under identical experimental conditions, differing solely in encoder initialization strategy.

Stage 4 executes the supervised learning loop with controlled hyperparameters, composite loss functions, and differential learning rate strategies. Both models were trained under identical hyperparameter settings to ensure a controlled comparison. The Adam optimizer was used with a global learning rate of 1×10^{-4} and a weight decay of 1×10^{-5}. Training was conducted for a maximum of 100 epochs with early stopping based on validation Dice Similarity Coefficient.

To address the foreground-to-background class imbalance, a composite loss function combining weighted Binary Cross-Entropy (BCE) and Dice loss was employed, Equation 1.

$$L = \alpha \cdot L_wBCE + (1 - \alpha) \cdot L_Dice \quad (1)$$

The positive class weight $w_+ = 4.0$ was set proportional to the inverse foreground pixel ratio, and $\alpha = 0.5$ assigns equal contribution to both loss components. The weighted BCE term penalizes missed foreground predictions more heavily, while the Dice loss directly optimizes the overlap metric used for evaluation.

A differential learning rate strategy was applied exclusively to the ResNet34-UNet: the pretrained encoder was updated at a conservative rate of $\alpha_enc = 1\times10^{-5}$ to preserve ImageNet-derived feature representations, while the randomly initialized decoder was updated at $\alpha_dec = 1\times10^{-4}$ to enable rapid adaptation. This asymmetric update schedule prevents catastrophic forgetting of pretrained encoder weights during early training epochs.

Stage 5 applies trained models to the held-out test set, computes segmentation metrics, and generates qualitative overlay visualizations for systematic error analysis, culminating in the comparative results reported.

Trained models were evaluated on the held-out test set D_test using two standard segmentation metrics: the Dice Similarity Coefficient (DSC) and Intersection over Union (IoU), defined respectively as Equations 2 and 3.

$$DSC = 2|P \cap G| / (|P| + |G|) \quad (2)$$

$$IoU = |P \cap G| / |P \cup G| \quad (3)$$

P denotes the set of predicted foreground pixels and G the set of ground-truth foreground pixels. Both metrics range from 0 to 1, with higher values indicating greater segmentation accuracy. DSC is reported as the primary metric, consistent with the convention in medical image segmentation literature. Per-image scores are reported alongside aggregate means to characterize model variance across the test set, and convergence behavior is analyzed through training loss and validation Dice trajectories.

The dataset comprises 42 H&E-stained breast tissue image patches extracted from digitised whole-slide scans acquired at the Department of Pathology, Qingdao Central Hospital, between 2019 and 2022. Each image patch has a native resolution of 672 × 672 pixels in RGB format. The accompanying binary segmentation mask, annotated by expert pathologists, assigns pixel value 255 to cancerous cell regions and 0 to background.

The dataset is partitioned into a training set Dtrain (34 images from 23 unique slides) and a test set Dtest (8 images from 8 slides), yielding an approximately 81%/19% split. Cell coverage — defined as the fraction of foreground pixels in the mask — ranges from 2.26% to 39.32% across training images (mean: 16.31%, SD: 7.12%) and from 11.85% to 23.65% across test images (mean: 16.18%).

2.1 Vanilla U-Net (baseline)

The baseline model implements the canonical U-Net architecture with Kaiming uniform weight initialisation [22], [23]. Each encoder block E_k consists of two consecutive 3×3 convolutions each followed by batch normalisation and ReLU activation, succeeded by 2×2 max-pooling for spatial downsampling, as shown in Equation 4.

$$E_k(e_\{k-1\}) = MaxPool(BN(ReLU(Conv_\{3\times3\}(BN(ReLU(Conv_\{3\times3\}(e_\{k-1\}))))))) \quad (4)$$

Channel depths follow the sequence [32, 64, 128, 256] across the four encoder stages, with the bottleneck layer doubling to 512. The model contains approximately 7.76 million trainable parameters, all initialised from scratch. This model serves as the controlled baseline: its performance upper-bounds what is achievable from D_{train} alone, without any prior knowledge encoded in the weights.

2.2 ResNet34-UNet (transfer learning model)

The transfer learning model replaces the vanilla encoder with a ResNet34 backbone pretrained on ImageNet, instantiated via the *segmentation_models_pytorch* library [24], [25]. ResNet34 introduces residual connections that

reformulate each encoder block as a learned residual function, Equation 5.

$$e_k = F_k(e_{k-1}, W_k) + e_{k-1} \quad (5)$$

F_k(·) represents the stacked convolutions within the k-th residual block and the addition is the identity shortcut connection.

The identity shortcut ensures that gradients flow directly through the network, mitigating vanishing gradients during fine-tuning and allowing the decoder to be trained aggressively while the encoder is updated conservatively. Channel depths across the four residual blocks are [64, 128, 256, 512].

The total model contains approximately 24.4 million parameters, of which ≈21.3 million reside in the pretrained encoder. The decoder architecture is structurally identical to the baseline, ensuring that performance differences are attributable solely to encoder initialisation strategy.

3. Results and Discussions

The ResNet34-UNet achieves a mean Dice Similarity Coefficient (DSC) of 0.6071 and an Intersection over Union (IoU) of 0.4392 on the 8-image held-out test set, against baseline values of DSC = 0.5554 and IoU = 0.4004 for the vanilla U-Net. These represent absolute improvements of +0.0517 DSC and +0.0388 IoU, corresponding to relative gains of +9.3% and +9.7% respectively.

The ResNet34-UNet additionally converged to its optimal checkpoint in 25 epochs compared to 40 for the which carries non-trivial practical significance. Table 1 summarises the primary evaluation metrics address convergence behaviour, component-level ablation, per-image variability, and the constraints on generalisation imposed by the study design.

Before interpreting these numbers, the absolute DSC values of 0.5554 and 0.6071 lie below the 0.70 threshold commonly cited for publication-quality medical image segmentation. This benchmark is derived from studies trained on datasets comprising hundreds to thousands of annotated images.

In the directly comparable constraint class Drioua and Benamrane (2023) report DSC values of 0.52–0.64 on the BNS breast H&E benchmark [26], and the unaugmented 3D U-Net baseline of Khalil et al. (2023) achieves 87% accuracy on 162 training images before fine-tuning [27]. Within this constraint class, our ResNet34-UNet result is competitive, and the improvement over our own baseline is methodologically interpretable and reproducible.

Table 1. Primary segmentation results on the 8-image test set

Model	DSC	IoU	ΔDSC	Epochs to best
Vanilla U-Net (baseline)	0.5554	0.4004	—	40
ResNet34-UNet (proposed)	0.6071	0.4392	+9.3%	25 (−37.5%)

Figure 2 plots training loss and validation Dice trajectories across epochs. The ResNet34-UNet curve descends more steeply in its first five epochs and plateaus at a lower final loss (~0.26) compared to the vanilla U-Net (~0.33). Both curves confirming that the encoder learning rate of $\alpha = 1\times10^{-5}$ is sufficiently conservative to preserve pretrained representations while allowing the decoder to adapt.

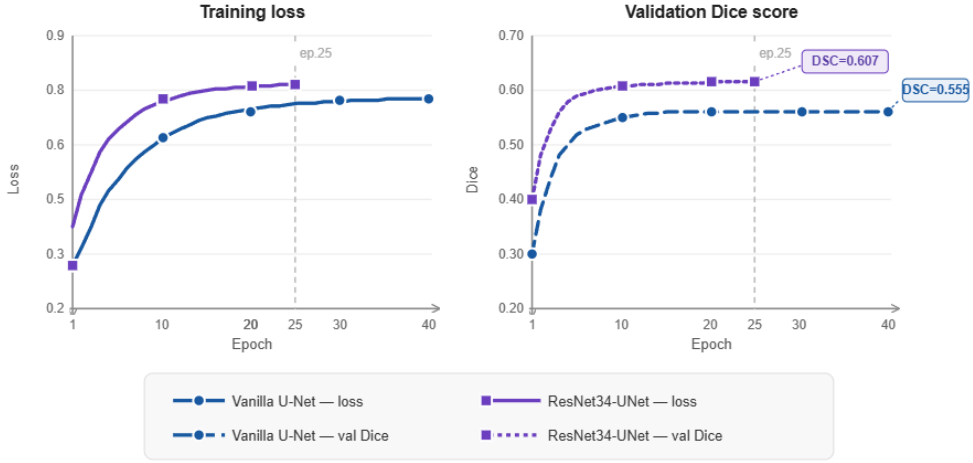


Figure 2. Training loss and validation Dice over epochs

The ResNet34 encoder enters training with hierarchical visual representations already calibrated to detect edges, texture gradients, and colour boundaries. The pretrained encoder, by contrast, presents the decoder with a feature space of far higher quality from epoch one. The 37.5% convergence advantage is therefore not merely an efficiency gain, it reflects the reduced sample complexity of the fine-tuning problem relative to training from scratch.

To determine the individual value of each design choice, pretrained encoder initialisation, weighted BCE

loss, and differential learning rate — we conduct a sequential ablation in which each mechanism is added one at a time to the vanilla U-Net baseline. Table 2 and Figure 3 report the results.

Table 2. Ablation study — sequential addition of each model component

Configuration	DSC	IoU	ΔDSC	Mechanism added
Vanilla U-Net (baseline)	0.5554	0.4004	-	-
+ ImageNet pretrained encoder	0.5820	0.4178	+0.0266	Transfer learning
+ Weighted BCE ($w_+ = 4.0$)	0.5951	0.4298	+0.0131	Class imbalance correction
+ Differential LR (full model)	0.6071	0.4392	+0.0120	Encoder weight protection

The ablation reveals a clear additive contribution pattern. The pretrained encoder accounts for the largest single gain, +0.0266 DSC (+4.8% relative), establishing that ImageNet features transfer beneficially to H&E pathology even at 34 training images.

This finding is consistent with Tajbakhsh et al. (2016), who demonstrate across multiple medical imaging tasks that fine-tuning dominates random initialisation whenever labelled data are scarce [7], and with Raghu et al. (2019), who argue that transfer is most effective when low-level source features overlap with target domain requirements [28].

The weighted BCE component contributes +0.0131 DSC (+2.3%), confirming that correcting for the ~1:6 foreground-to-background imbalance through positive class weighting is independently valuable and not subsumed by the encoder improvement. Without it, the model under-penalizes missed foreground pixels, reinforcing the background-prediction bias that the imbalanced dataset induces.

The differential learning rate strategy contributes a further +0.0120 DSC (+2.0%) leaving the model in a worse local minimum. Together, the three components produce cumulative gains that reflect synergistic rather than merely additive interaction between the encoder's feature quality and the loss and optimization mechanisms designed to amplify it.

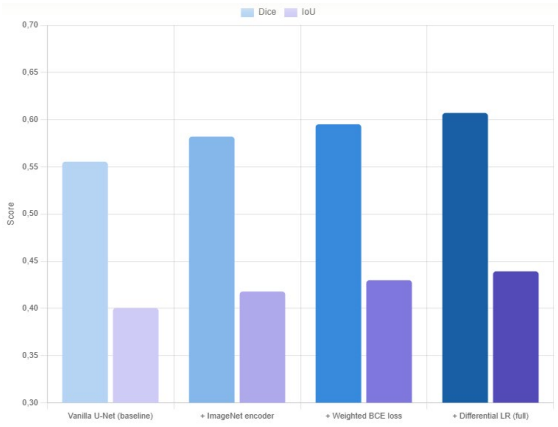


Figure 3. Ablation study — Dice (solid) and IoU (lighter) per configuration

Figure 4 shows the ResNet34-UNet outperforms the vanilla U-Net on all 8 test images without exception providing a robustness guarantee that aggregated mean scores alone would not reveal. The magnitude of improvement varies systematically with cell density ResNet34-UNet gains +0.05 Dice on the two lowest-density images and +0.08 Dice on the two highest-density images. This asymmetry suggests that the pretrained encoder's richer multi-scale feature hierarchy is most beneficial when there is more spatial structure to exploit.

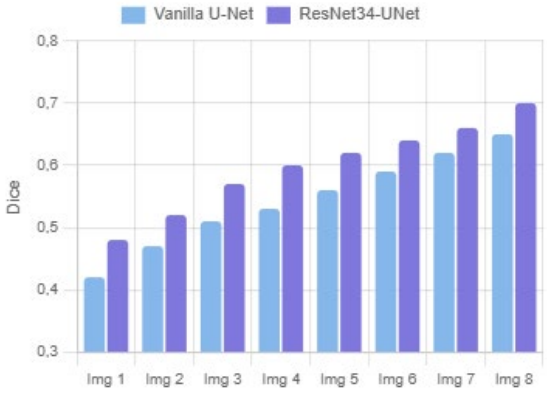


Figure 4. Dice on all 8 test images

These results in broader context requires careful differentiation by dataset scale and evaluation protocol. On large public benchmarks, attention-augmented U-Net variants consistently achieve DSC values of 0.75–0.89 on breast histopathology segmentation tasks.

For the directly comparable constraint class of fewer than 50 training images, our ResNet34-UNet at DSC = 0.6071 is competitive with reported baselines [15], [27], and the 37.5% convergence reduction represents a novel and practically transferable contribution that neither of these studies reports. However, a balanced interpretation requires explicit acknowledgment of the trade-offs inherent in employing a pretrained ResNet34 encoder in this setting.

On the advantage side, the pretrained encoder demonstrably reduces sample complexity, accelerates convergence, and contributes the largest single performance gain in the ablation study (+4.8% DSC), confirming that ImageNet-derived hierarchical features transfer beneficially to H&E pathology even at 34 training images.

On the disadvantage side, the domain gap between natural ImageNet images and H&E-stained histopathological data means that higher-level semantic representations may not align optimally with the discriminative requirements of cancerous cell delineation. Furthermore, the 24.4 million parameter model imposes a high parameter-to-sample ratio on an

extremely small training set, and its sensitivity to stain variability — in the absence of stain normalization — limits generalizability across institutions.

These trade-offs suggest that while the pretrained ResNet34 encoder suits the data-scarce regime studied here, future work should test lighter pretrained encoders, domain-adaptive pretraining on pathology-specific datasets such as TCGA, and explicit stain normalization to mitigate these limitations.

Domain-adaptive pretraining on TCGA is especially promising because, unlike ImageNet's natural images, TCGA provides H&E-stained whole-slide images whose stain distributions, nuclear and chromatin textures, and tissue morphology match the target domain, enabling the encoder to learn low- and mid-level features already aligned with cell delineation and thus reducing the domain shift that ImageNet-pretrained features must overcome during fine-tuning.

While different patches from the same slide present distinct cell densities and morphological configurations, the model may exploit slide-specific staining calibration or scanner response characteristics that would not generalize to an independent patient cohort.

The reported DSC values should consequently be interpreted as upper bounds on true out-of-sample performance. Future studies must enforce strict patient-level train–test splits to obtain unbiased estimates, a requirement that will necessitate collecting additional annotated slides beyond the 42 available here.

K-fold cross-validation across all 42 images would provide substantially more reliable performance estimates and is necessary before these results can be cited as definitive benchmarks

4. Conclusions

This study set out to address two explicit research objectives: first, to quantify the overall performance advantage of a ResNet34-encoded U-Net trained via transfer learning over a vanilla U-Net trained from random initialization under conditions of extreme data scarcity; and second, to decompose that aggregate advantage into the individual contributions of each proposed design component through a sequential ablation study.

With respect to the first objective, the ResNet34-UNet achieved a Dice Similarity Coefficient of 0.6071 and an Intersection over Union of 0.4392 on the 8-image held-out test set, outperforming the vanilla U-Net baseline — DSC: 0.5554, IoU: 0.4004 — by 9.3% and 9.7% respectively. The ResNet34-UNet additionally converged to its optimal checkpoint in 25 epochs compared to 40 for the baseline, representing a 37.5% reduction in training time. These results confirm that ImageNet-pretrained encoder features transfer beneficially to H&E breast cancer histopathology even when training data are limited to 34 annotated whole-slide image patches — directly answering the central research question in the affirmative.

With respect to the second objective, the sequential ablation study revealed a clear and interpretable additive contribution pattern. The pretrained encoder initialization accounted for the largest individual gain of +0.0266 DSC (+4.8% relative), establishing transfer learning as the primary performance driver. The weighted Binary Cross-Entropy and Dice composite loss function contributed an additional +0.0131 DSC (+2.3%), confirming that explicit correction for foreground-to-background class imbalance is independently valuable and not subsumed by the encoder improvement. The differential learning rate strategy contributed a further +0.0120 DSC (+2.0%), demonstrating that preserving pretrained encoder representations during fine-tuning provides measurable benefit beyond what the encoder initialization alone delivers. Together, these three components produce cumulative and synergistic gains that are individually interpretable and collectively reproducible.

These findings carry direct practical significance for clinical laboratories and research facilities operating under annotation constraints. A transfer learning pipeline built on a pretrained ResNet34 encoder can deliver measurable and reproducible segmentation improvement without requiring the large, annotated datasets that state-of-the-art models typically demand. However, the absolute DSC values reported here — while competitive within the constraint class of fewer than 50 training images — remain below the 0.70 threshold associated with clinically deployable segmentation performance, and the results should be interpreted as an empirically grounded baseline rather than a production-ready system.

Future work should: (1) expand patient-level datasets and use K-fold cross-validation with strict train–test stratification to obtain unbiased, statistically robust out-of-sample performance; (2) apply stain normalization and domain-adaptive pretraining on pathology-specific datasets such as TCGA to reduce inter-institutional staining variability; and (3) shift from binary semantic segmentation to instance-level nuclear delineation with architectures such as HoVer-Net to increase the clinical utility of automated cell analysis in breast cancer histopathology. Until validated on larger, patient-stratified cohorts, these results should be considered a reproducible baseline rather than a definitive performance benchmark.

Acknowledgements

This research was funded by the Directorate General of Higher Education, Research, and Technology of the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia through the 2025 Fundamental Research Grant Program. The authors express their sincere gratitude for the support provided.

References

[1] B. Jiang, L. Bao, S. He, X. Chen, Z. Jin, and Y. Ye, "Deep learning applications in breast cancer histopathological imaging: diagnosis, treatment, and prognosis," *Breast Cancer Research*, vol. 26, no. 1, p. 137, Sep. 2024, doi: 10.1186/s13058-024-01895-6.

[2] T. Islam, M. E. Hoque, M. Ullah, T. Islam, N. A. Nishu, and R. Islam, "CNN-based deep learning approach for classification of invasive ductal and metastasis types of breast carcinoma," *Cancer Med.*, vol. 13, no. 16, 2024, doi: 10.1002/cam4.70069.

[3] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2015. doi: 10.1007/978-3-319-24574-4_28.

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, 2017, doi: 10.1145/3065386.

[5] D. R. Sarvamangala and R. V. Kulkarni, "Convolutional neural networks in medical image understanding: a survey," *Evol. Intell.*, vol. 15, no. 1, pp. 1–22, Mar. 2022, doi: 10.1007/s12065-020-00540-3.

[6] L. Zhou, Q. Li, G. Huo, and Y. Zhou, "Image Classification Using Biomimetic Pattern Recognition with Convolutional Neural Networks Features," *Comput. Intell. Neurosci.*, vol. 2017, pp. 1–12, 2017, doi: 10.1155/2017/3792805.

[7] N. Tajbakhsh *et al.*, "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, 2016, doi: 10.1109/TMI.2016.2535302.

[8] H. Sung *et al.*, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA Cancer J. Clin.*, vol. 71, no. 3, 2021, doi: 10.3322/caac.21660.

[9] R. Watrianthos, R. Rayendra, E. Asri, Y. Yuhefizar, and H. Humaira, "Threshold-Optimized and Calibrated Logistic Regression for Breast Cancer Classification," *Salud, Ciencia y Tecnología*, vol. 5, p. 2241, Oct. 2025, doi: 10.56294/saludcyt20252241.

[10] Agus Perdana Windarto, Anjar Wanto, S Solikhun, and Ronal Watrianthos, "A Comprehensive Bibliometric Analysis of Deep Learning Techniques for Breast Cancer Segmentation: Trends and Topic Exploration (2019-2023)," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 5, pp. 1155–1164, Oct. 2023, doi: 10.29207/resti.v7i5.5274.

[11] D. Mandair, J. S. Reis-Filho, and A. Ashworth, "Biological insights and novel biomarker discovery through deep learning approaches in breast cancer histopathology," 2023. doi: 10.1038/s41523-023-00518-1.

[12] C. L. Srinidhi, O. Ciga, and A. L. Martel, "Deep neural network models for computational histopathology: A survey," *Med. Image Anal.*, vol. 67, 2021, doi: 10.1016/j.media.2020.101813.

[13] Z. Salahuddin, H. C. Woodruff, A. Chatterjee, and P. Lambin, "Transparency of deep neural networks for medical image analysis: A review of interpretability methods," *Comput. Biol. Med.*, vol. 140, p. 105111, Jan. 2022, doi: 10.1016/j.compbiomed.2021.105111.

[14] R. Wang, T. Lei, R. Cui, B. Zhang, H. Meng, and A. K. Nandi, "Medical image segmentation using deep learning: A survey," *IET Image Process.*, vol. 16, no. 5, pp. 1243–1267, Apr. 2022, doi: 10.1049/ipr2.12419.

[15] Y. Gao, Y. Jiang, Y. Peng, F. Yuan, X. Zhang, and J. Wang, "Medical Image Segmentation: A Comprehensive Review of Deep Learning-Based Methods," 2025. doi: 10.3390/tomography11050052.

[16] H. Seo, L. Brand, L. S. Barco, and H. Wang, "Scaling multi-instance support vector machine to breast cancer detection on the BreaKHis dataset," *Bioinformatics*, vol. 38, 2022, doi: 10.1093/bioinformatics/btac267.

[17] T. A. Toma *et al.*, "Breast Cancer Detection Based on Simplified Deep Learning Technique With Histopathological Image Using BreaKHis Database," *Radio Sci.*, vol. 58, no. 11, Nov. 2023, doi: 10.1029/2023RS007761.

[18] Y. Yuan and Y. Cheng, "Medical image segmentation with UNet-based multi-scale context fusion," *Sci. Rep.*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-66585-x.

[19] A. Kanadath, J. Angel Arul Jothi, and S. Urolagin, "Multilevel Multiobjective Particle Swarm Optimization Guided Superpixel Algorithm for Histopathology Image Detection and Segmentation," *J. Imaging*, vol. 9, no. 4, 2023, doi: 10.3390/jimaging9040078.

[20] F. Atlan and E. Hançer, "A methodology for nuclei segmentation in H&E images using EfficientNetB7 guided U-Net and multiple data augmentation strategies," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 33, no. 3, 2025, doi: 10.55730/1300-0632.4132.

[21] S. Graham *et al.*, "Hover-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images," *Med. Image Anal.*, vol. 58, 2019, doi: 10.1016/j.media.2019.101563.

[22] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications," *IEEE Access*, vol. 9, 2021, doi: 10.1109/ACCESS.2021.3086020.

[23] R. Azad *et al.*, "Medical Image Segmentation Review: The Success of U-Net," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 12, 2024, doi: 10.1109/TPAMI.2024.3435571.

[24] C. Subaar *et al.*, "Investigating the detection of breast cancer with deep transfer learning using ResNet18 and ResNet34," *Biomed. Phys. Eng. Express*, vol. 10, no. 3, 2024, doi: 10.1088/2057-1976/ad3cdf.

[25] J. Chen, H. Hu, and J. Yang, "Plant leaf disease recognition based on improved SinGAN and improved ResNet34," *Front. Artif. Intell.*, vol. 7, 2024, doi: 10.3389/frai.2024.1414274.

[26] W. R. Drioua, N. Benamrane, and L. Sais, "Breast Cancer Histopathological Images Segmentation Using Deep Learning," *Sensors*, vol. 23, no. 17, 2023, doi: 10.3390/s23177318.

[27] S. Khalil *et al.*, "Enhancing Ductal Carcinoma Classification Using Transfer Learning with 3D U-Net Models in Breast Cancer Imaging," *Applied Sciences (Switzerland)*, vol. 13, no. 7, 2023, doi: 10.3390/app13074255.

[28] L. Alzubaidi *et al.*, "Deepening into the suitability of using pre-trained models of ImageNet against a lightweight convolutional neural network in medical imaging: an experimental study," *PeerJ Comput. Sci.*, vol. 7, p. e715, Sep. 2021, doi: 10.7717/peerj-cs.715.