Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control
Journal homepage: http://kinetik.umm.ac.id
ISSN: 2503-2267
Vol. 11, No. 3, August, Pp. 441-450

Automated breast cancer cell counting: comparing multi-class segmentation and two-stage classification strategies

Dzaky Hanif Arjuna*[1], Edy Kurniawan[1], Imam Susilo[2], Reza Fuad Rachmadi[1,3], I Ketut Eddy Purnama[1,3]
Electrical Engineering, Institut Teknologi Sepuluh Nopember, Indonesia[1]
Anatomical Pathology, Universitas Airlangga, Indonesia[2]
Computer Engineering, Institut Teknologi Sepuluh Nopember, Indonesia[3]

Article Info
Keywords:
Breast Cancer, Cell Classification, Deep Learning, Hematoxylin and Eosin, MobileNetV2, Multi-class Segmentation, U-Net

Article history:
Received: December 14, 2025
Accepted: June 01, 2026
Published: August 01, 2026

Cite:
D. H. Arjuna, E. Kurniawan, R. F. Rachmadi, and I. K. E. Purnama, "Automated Breast Cancer Cell Counting: Comparing Multi-class Segmentation and Two-stage Classification Strategies", *KINETIK*, vol. 11, no. 3, Aug. 2026.
https://doi.org/10.22219/kinetik.v11i3.2639

*Corresponding author.
Dzaky Hanif Arjuna
E-mail address:
6022241128@student.its.ac.id

Abstract
The manual interpretation of Hematoxylin and Eosin (H&E) histopathology images for breast cancer diagnosis is hindered by time constraints and observer bias. This research aims to develop an automated system using deep learning for cell detection and classification by evaluating two key approaches: Multi-class Segmentation (single-stage) and Segmentation followed by Classification (two-stage). The U-Net architecture was employed for segmentation, while MobileNetV2 and VGG16 were used for classification. The models were evaluated on the public IHC4BC dataset and primary data from Airlangga University Hospital (RSUA). The study also evaluated the impact of Resizing and Tiling data processing strategies. Experimental results showed that although the MobileNetV2 and VGG16 classification models achieved a high testing accuracy of 98.80%, the integrated two-stage system exhibited a high counting error, with a Mean Absolute Error (MAE) of 119.87 for positive cells, primarily due to under-segmentation of overlapping cells. In contrast, the Multi-class Segmentation approach utilizing the Tiling strategy demonstrated superior performance. This model effectively preserved spatial resolution while distinguishing cell types simultaneously, achieving the lowest MAE of 18.46 for positive cells and 1.66 for negative cells. This study concludes that Multi-class Segmentation with the Tiling strategy is the most effective and accurate approach for automated cell counting in histopathology images.

1. Introduction

Cancer remains a critical public health challenge and a leading cause of mortality worldwide. Among all malignancies, breast cancer has the highest incidence and is the primary driver of cancer-related deaths among women, both globally and in Indonesia [1], [2]. Given its continuously increasing prevalence, robust early detection and accurate prognostic strategies are essential for improving patient survival rates and treatment outcomes [3].

The standard clinical approach for diagnosing and grading breast cancer involves the histological examination of tissue biopsy samples stained with Hematoxylin and Eosin (H&E). In this process, hematoxylin imparts a blue-purple hue to cell nuclei, while eosin stains the cytoplasm pink [4]. Pathologists manually analyze these slides to quantify specific biomarkers such as Ki-67 by counting positive (proliferating) and negative (non-proliferating) tumor cells—a metric that directly reflects tumor aggressiveness and informs therapeutic decisions [5]. Recent advances in Convolutional Neural Networks (CNNs) and deep learning have enabled Computer-Aided Diagnosis (CAD) systems to achieve exceptional performance in classification, detection, and segmentation tasks, positioning them as a viable alternative to manual analysis [8], [9].

Manual quantification of tumor cells is labor-intensive, time-consuming, and highly prone to inter-observer and intra-observer variability, particularly when analyzing images containing thousands of cells [6], [7]. To address this challenge, the prevailing deep learning strategy employs a *two-stage* cascade framework: a segmentation model (e.g., U-Net) first isolates cellular regions, after which a classifier (e.g., VGG16 or MobileNetV2) categorizes each detected cell [10], [11]. However, despite achieving high classification accuracy, these approaches suffer from a critical and underreported bottleneck: **under-segmentation**. In dense tissue regions where cells heavily overlap or cluster together, binary segmentation models tend to merge multiple nuclei into a single object. Consequently, the classifier treats this merged cluster as a single cell, leading to severe undercounting errors regardless of the classifier's individual precision [12]. Additionally, the common practice of resizing input images to reduce computational load degrades spatial resolution, further obscuring the fine morphological details necessary for accurate cell separation [13].

To overcome these limitations, this study proposes and evaluates a *single-stage multi-class segmentation* approach as an alternative to the conventional two-stage pipeline. The proposed method employs a modified U-Net architecture that simultaneously classifies each pixel into one of three classes—background, positive cell, or negative cell—thereby eliminating the need for a separate classification stage. The key novelty of this work is

twofold: (1) by learning semantic class boundaries and spatial context concurrently, the model is better equipped to delineate touching and overlapping cells than binary segmentation; and (2) the adoption of an *image tiling strategy* preserves the native high resolution of histopathology images, preventing the information loss associated with conventional resizing methods.

This study conducts a systematic comparison between the proposed single-stage approach and the conventional two-stage segmentation-classification pipeline using the public IHC4BC dataset and primary clinical data from Airlangga University Hospital (RSUA). The primary contribution of this work is to demonstrate that, although two-stage models achieve high standalone classification accuracy, the single-stage multi-class approach with Tiling strategy significantly outperforms them in end-to-end cell counting accuracy, providing a more reliable and clinically relevant tool for automated tumor proliferation analysis.

2. Research Method

This study employs a quantitative experimental approach to compare two deep learning strategies for automated cell counting: a single-stage Multi-Class Segmentation and a Two-Stage Segmentation-Classification framework. The methodology encompasses data collection, preprocessing, architectural design, post-processing for instance extraction, and performance evaluation.

2.1 Dataset Preparation

This study utilizes a dataset obtained from Airlangga University Hospital (RSUA), consisting of histopathology images stained with Hematoxylin and Eosin (H&E). The dataset contains a total of 480 images acquired at 40× magnification with a resolution of 1000 × 1000 pixels. Ground truth annotations were generated using QuPath and validated by pathologists to label cell nuclei into two biological categories: Positive (proliferating) and Negative (normal/non-proliferating). A 70:20:10 data split was employed, allocating 70% of the images for training, 20% for validation, and the remaining 10% for testing.

2.2 Data Preprocessing Strategies

To assess how input dimensions influence the model performance, two distinct preprocessing strategies were implemented.

2.2.1 Image Resizing

The original images (1000 × 1000 pixels) were resized to 512 × 512 pixels. While computationally efficient, this standard approach may result in the loss of fine morphological details that are essential for separating dense cell clusters. The resizing workflow is illustrated in Figure 1.

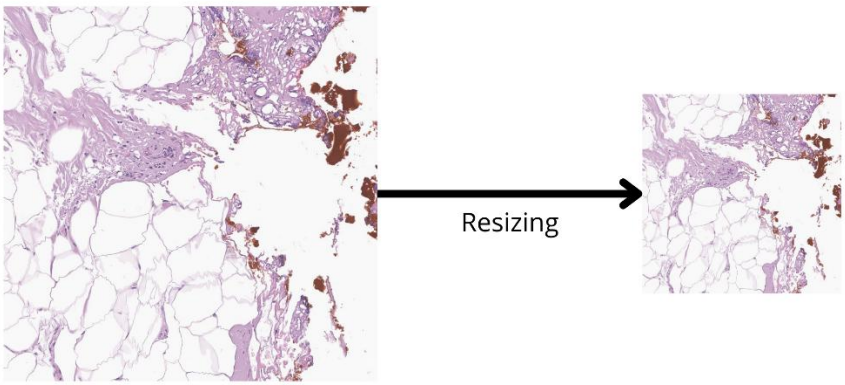

Figure 1. Illustration of the Image Resize Strategy

2.2.2 Image Tiling

The original images were divided into four non-overlapping 500 × 500 pixel patches without scaling. This tiling strategy preserves the native resolution and detailed cellular features, thereby supporting improved separation of overlapping cells. Figure 2 illustrates how each original image was partitioned into four non-overlapping patches while preserving its native spatial resolution.

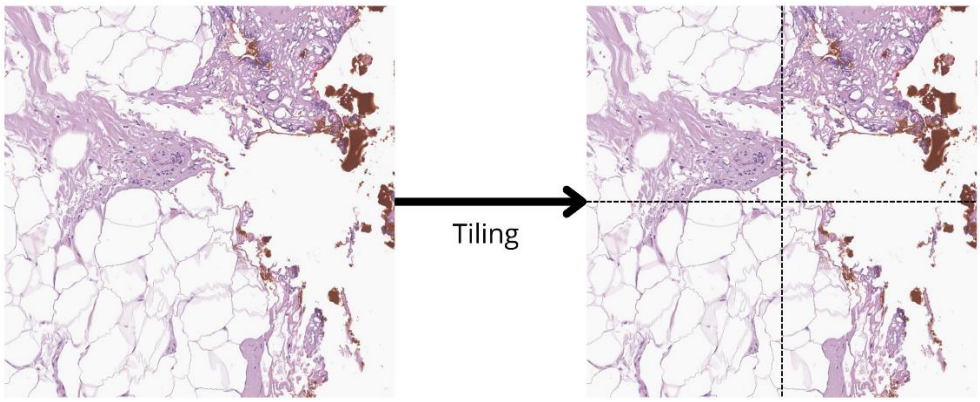

Figure 2. Illustration of the Image Tiling Strategy

Dynamic augmentation was utilized during the training phase to prevent the network from overfitting. By applying random flips (horizontal and vertical) and rotations of 90°, 180°, and 270°, the model was exposed to the diverse spatial orientations typically found in histological slides.

2.3 Single-Stage Multi-Class Segmentation Architecture

In the first approach, a modified U-Net architecture was employed to perform end-to-end multi-class segmentation [14]. The architecture is designed to perform pixel-wise classification by assigning each pixel to one of three distinct classes: Background, Positive Cell, or Negative Cell. The architecture consists of a contracting path (encoder) with convolutional blocks that progressively increase from 32 to 512 filters, and an expansive path (decoder) to ensure accurate localization. To mitigate the loss of spatial resolution typically caused by downsampling operations, the model employs skip connections that transfer feature maps directly from the encoder to the decoder. Finally, the output layer generates the multi-class prediction map using a 1 × 1 convolution followed by a Softmax activation function. The overall encoder-decoder structure and skip connections are shown in Figure 3.

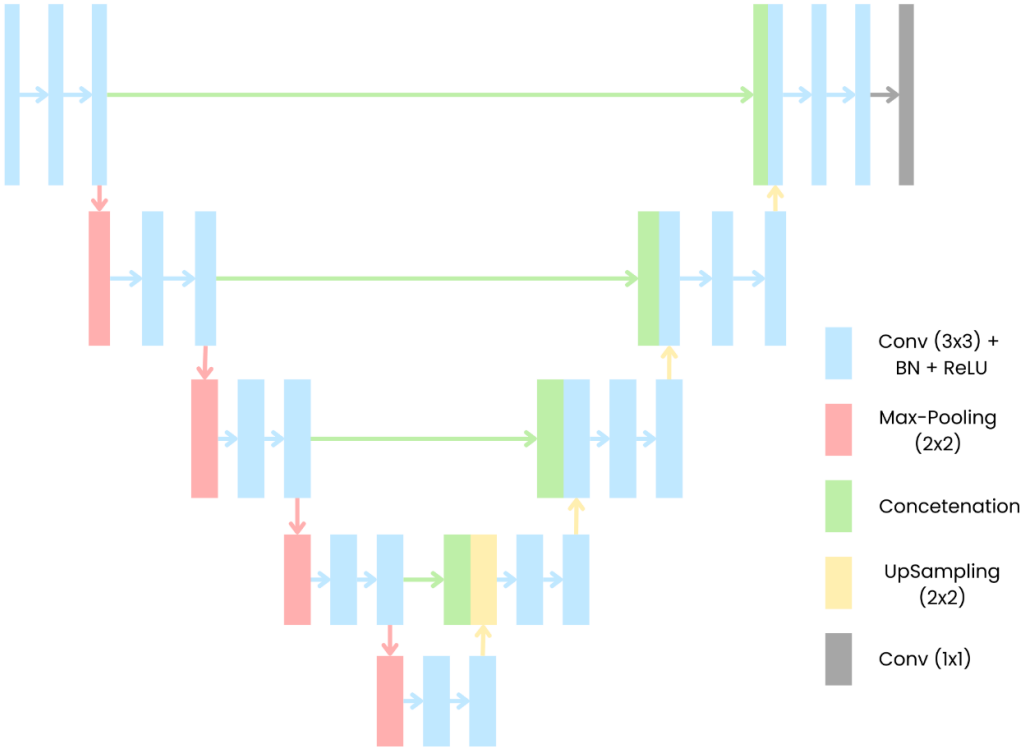

Figure 3. U-Net Architecture

2.4 Two-Stage Segmentation and Classification Framework

The two-stage segmentation and classification framework separates localization and identification into sequential steps to leverage specialized deep learning models. This structure enables precise extraction of cellular regions followed by accurate classification of individual cells.

The workflow begins by utilizing a binary U-Net to extract all cell regions from the background. During this phase, the network is trained solely to distinguish between these two categories, producing binary masks without attempting to classify cell types. This simplifies the learning task and ensures consistent detection of all potential cell regions.

After binary segmentation, post-processing is applied to isolate individual cell instances. A contour detection algorithm identifies the boundaries of each segmented object, and a minimum area threshold of 250 pixels is used to filter out noise and small artifacts. Bounding boxes are then generated around valid contours, and the corresponding regions are cropped from the original high-resolution images to create a dataset of single-cell patches.

In the final stage, the cropped cell patches are passed to a CNN classifier to determine whether each cell belongs to the Positive or Negative class. Two transfer learning models, MobileNetV2 and VGG16, were evaluated for this classification task. MobileNetV2 was selected because of its low computational requirements. To enhance generalization and reduce overfitting, the network was equipped with Global Average Pooling and a dropout rate of 0.3 [15]. VGG16, selected for its deeper architecture, was integrated with Global Average Pooling followed by a fully connected layer containing 256 units with ReLU activation. To effectively reduce overfitting, a dropout rate of 0.5 was applied [16]. The MobileNetV2 and VGG16 architectures used in the classification stage are presented in Figure 4 and Figure 5, respectively.

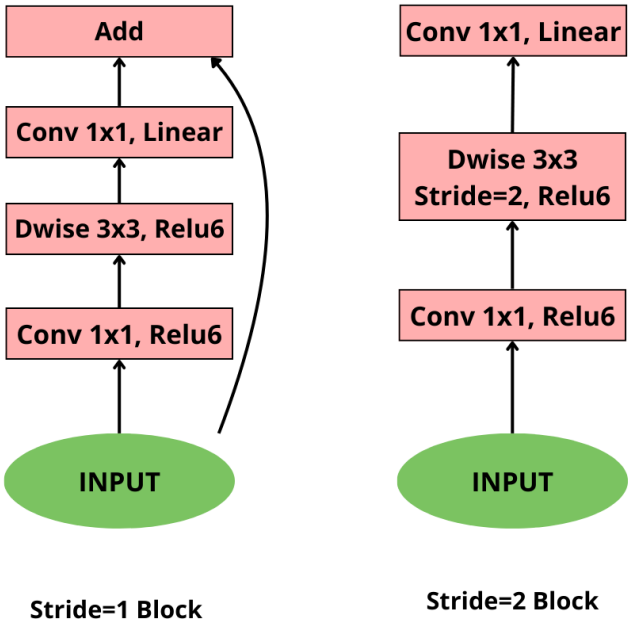

Figure 4. MobileNetV2 Architecture

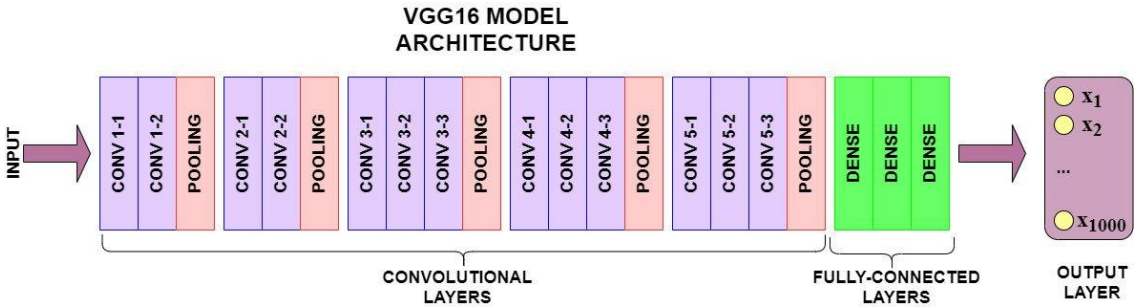

Figure 5. VGG16 Architecture

2.5 Experimental Setup

For model development and training, Python and TensorFlow were employed. The experiments were conducted on a local machine equipped with 32 GB of system memory and an NVIDIA GeForce RTX 3060 graphics card. Six experimental scenarios were designed to assess the impact of different preprocessing strategies (Resizing vs. Tiling) and hyperparameter configurations. The models were optimized using the Adam optimizer [17]. To optimize the performance of each method, specific loss functions were employed during training. Class imbalance was addressed using a weighted combination of Focal Loss and Dice Loss. This approach is particularly effective given the dominance of background pixels (majority class) over the smaller cell regions (minority class) [18]. In the two-stage framework, the classification models (MobileNetV2 and VGG16) were trained using Categorical Crossentropy loss to distinguish between positive and negative cell patches.

2.6 Evaluation Metrics

To quantify the quality of the segmentation output, standard evaluation metrics, namely Intersection over Union (IoU) and the Dice Coefficient were employed [19], [20]. For the classification task, Accuracy, Precision, and Recall were used. The final cell counting performance was evaluated using the Mean Absolute Error (MAE), defined as the mean absolute deviation between the predicted cell counts (Pi) and the ground truth cell counts (Gi) across N images, as formulated in Equation 1:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|Pi - Gi| \quad (1)$$

This metric directly reflects the clinical utility of the system for estimating tumor proliferation indices.

3. Results and Discussion

The experimental results are presented below, benchmarking the effectiveness of the proposed single-stage approach against the conventional two-stage segmentation and classification workflow. The discussion focuses on the impact of the data preprocessing strategies (Resizing vs. Tiling) and the effectiveness of each architecture for automated cell counting.

3.1 Performance of Single-Stage Multi-Class Segmentation

The performance of the U-Net model was evaluated through six experimental scenarios, each employing different preprocessing strategies and hyperparameter configurations. Table 1 summarizes the segmentation performance on the test set.

Table 1. Performance comparison of Multi-class Segmentation Scenarios

Scenario	Data Preprocessing Strategy	Epochs	Mean IoU	Dice Coefficient	Accuracy
1	Resize 512 × 512, Batch Size 4, Learning Rate 0.0001, Dropout 0.2	50	0.6031	0.8445	0.9146
2	Resize 512 × 512, Batch Size 8, Learning Rate 0.00005, Dropout 0.3 + L2	50	0.5943	0.8102	0.9085
3	Resize 512 × 512, Batch Size 4, Learning Rate 0.0001, Dropout 0.2	100	0.6292	0.8653	0.9212
4	Resize 512 × 512, Batch Size 8, Learning Rate 0.00005, Dropout 0.3 + L2	100	0.6138	0.8562	0.9234
5	Tiling 500 × 500, Batch Size 4,	50	0.6603	0.8740	0.9241

6	Learning Rate 0.0001, Dropout 0.2 Tiling 500 × 500, Batch Size 4, Learning Rate 0.00005, Dropout 0.3 + L2	50	0.6619	0.8664	0.9221

As shown in Table 1, the data preprocessing strategy significantly influenced model performance. Scenarios utilizing Image Tiling strategy (Scenarios 5 and 6) consistently outperformed those using Image Resizing (Scenarios 1-4). Scenario 5 achieved the best performance, recording a Dice Coefficient of 0.8740 and a Mean IoU of 0.6603. Figure 6 illustrates the convergence behavior of the best-performing model (Scenario 5), showing the training and validation trajectories over 50 epochs. The model demonstrates stable learning without significant overfitting, as evidenced by the close alignment between the training and validation curves throughout the training process.

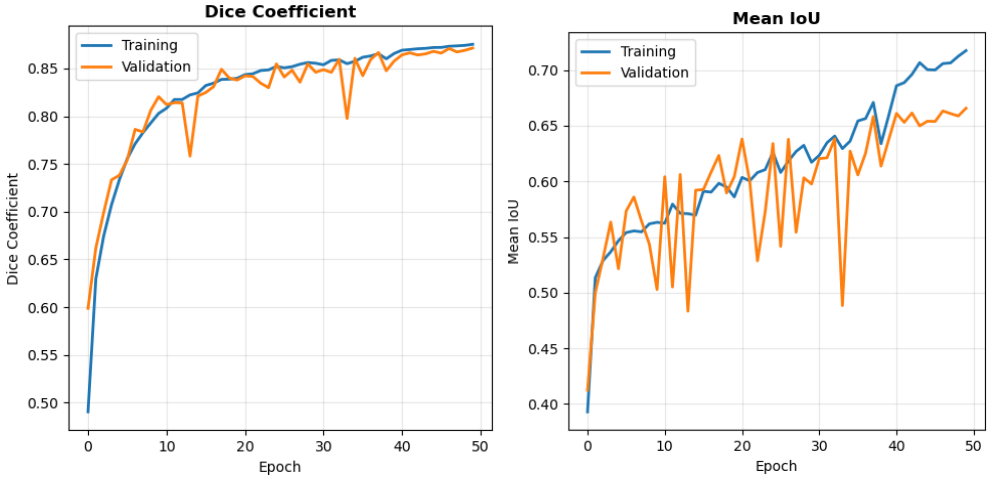

Figure 6. Training and Validation Metrics History for the Optimal Multi-class U-Net model (Scenario 5)

The superiority of the Tiling strategy can be attributed to its preservation of native image resolution (1000 × 1000 pixels). Resizing images to 512 × 512 pixels results in the loss of high-frequency spatial details, blurring the boundaries between adjacent cells [13]. In contrast, the Tiling strategy preserves the fine morphological features required for the model to distinguish touching cells, resulting in sharper segmentation masks, as visualized in Figure 7. From a clinical perspective, the improvement in Mean IoU from 0.60 (Resizing) to 0.66 (Tiling) may appear modest numerically, but it reflects a substantial qualitative gain in boundary delineation between adjacent cells—a critical factor when estimating Ki-67 proliferation index, where misclassifying a cluster of cells as a single object directly inflates or deflates the reported index. Furthermore, Scenario 5 (Dropout 0.2) slightly outperformed Scenario 6 (Dropout 0.3 + L2), suggesting that excessive regularization on high-resolution data may suppress subtle features of the minority class (negative cells), leading to lower sensitivity [6].

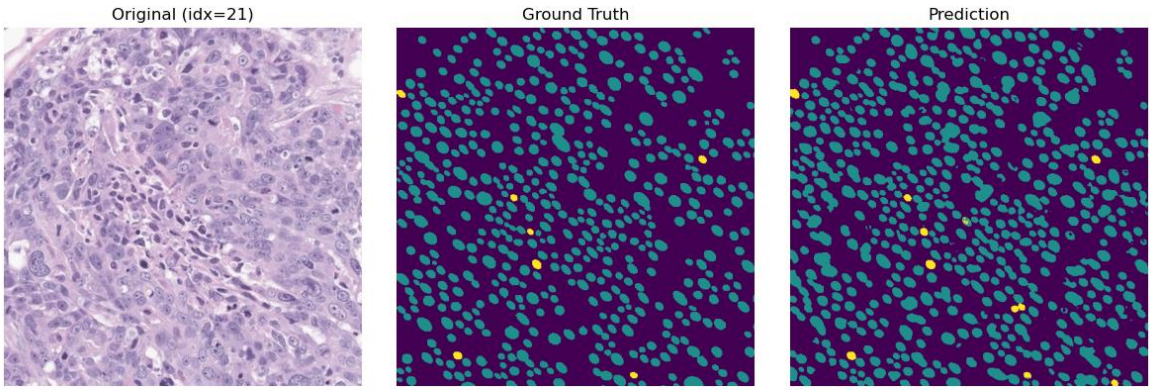

Figure 7. Visualization of Multi-class Segmentation Results using the Tiling Strategy

3.2 Performance of the Two-Stage Classification Framework

The performance of the two-stage framework was evaluated sequentially, beginning with binary segmentation to isolate cellular regions, followed by the classification of these regions.

The ability of the U-Net architecture to isolate cellular regions from the background was evaluated under three experimental scenarios, each defined by different epoch and batch size configurations. The results presented here were obtained from the evaluation on the test dataset.

Table 2. Performance Comparison of Binary Segmentation Scenarios for Region of Interest (ROI) Extraction

Scenario	Configuration	Epochs	Mean IoU	Dice Coefficient	Accuracy
1	Tiling 500 × 500 Batch Size 4, Learning Rate 0.0001	50	0.6602	0.7950	0.9337
2	1000 × 1000, Batch Size 4, Learning Rate 0.0001	100	0.5693	0.7253	0.9197
3	Resize 1024 × 1024, Batch Size 4, Learning Rate 0.0001	100	0.6519	0.7890	0.9330

As presented in Table 2, the binary segmentation performance was evaluated across three different configurations to determine the optimal method for Region of Interest (ROI) extraction. The best performance was achieved by Scenario 1, which employed a 500 × 500 tiling strategy trained for 50 epochs, achieving a Dice Coefficient of 0.7950 and a Mean IoU of 0.6602. However, Scenario 3, which employed a Resizing strategy (1024 × 1024) was trained for 100 epochs, achieved highly comparable performance with a Mean IoU of 0.6519. Despite the slightly higher IoU achieved by Scenario 1, Scenario 3 was selected as the segmentation model for the two-stage framework. This decision was driven by the model's stability and robustness resulting from the longer training duration (100 epochs) compared with Scenario 1, ensuring more consistent convergence for detecting cellular regions. This trade-off reflects a deliberate design decision: in a two-stage pipeline, the segmentation model serves as the sole source of object proposals for the downstream classifier. Therefore, a model with a slightly lower IoU but more stable convergence is preferred over a higher-IoU model with less stable training dynamics, as inconsistent segmentation outputs would propagate errors throughout the entire counting pipeline. The qualitative binary segmentation results produced by Scenario 3 are presented in Figure 8.

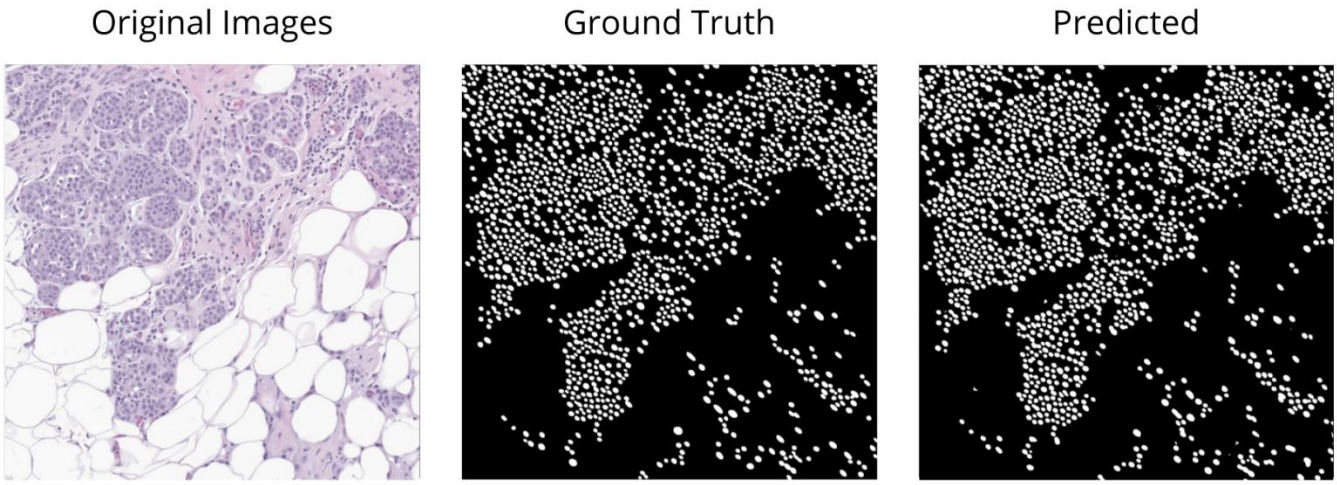

Figure 8. Visualization of Binary Segmentation Results (Scenario 3)

Following segmentation, the isolated cellular regions were cropped and fed into the classification models. Table 3 presents the performance comparison between MobileNetV2 and VGG16 using these single-cell patches.

Table 3. Evaluation of Classifier Performance Using Individual Cell Patches

Model	Accuracy	Precision	Recall	Training Time
MobileNetV2	0.9880	0.9880	0.9880	16 hours
VGG16	0.9880	0.9880	0.9880	38 hours

MobileNetV2 achieved a testing accuracy of 98.80%, identical to that of VGG16, but with significantly higher computational efficiency (approximately 16 hours training time compared with 38 hours for VGG16). This finding aligns with previous studies highlighting that MobileNet is efficient for medical image analysis due to its depthwise separable convolutions [15]. The high precision and recall (both >98%) indicate that both models are highly reliable in distinguishing Negative (benign) from Positive (malignant) cells when provided with correctly isolated cell patches.

3.3 Comparative Analysis of Counting Accuracy

While the two-stage classifiers showed high individual classification accuracy, the final system-level performance revealed a critical discrepancy when applied to whole-slide counting. Table 4 compares the counting error (MAE) between the best Single-Stage model (Scenario 5) and the Two-Stage framework.

Table 4. Comparison of Cell Counting Error (MAE) Between the Single-Stage and Two-Stage Approaches

Model	Single-Stage (Multi-Class Segmentation)	Two-Stage (Segmentation + Classification)
MAE (Positive)	18.46	119.87
MAE (Negative)	1.66	1.39

The results indicate a substantial difference in counting accuracy. The Two-Stage method resulted in a high MAE of 119.87 for positive cells, despite achieving a classification accuracy of 98.80%. This error stems from the bottleneck in the initial binary segmentation stage. The binary U-Net failed to distinguish overlapping cells, causing them to merge into a single region and resulting in under-segmentation. Consequently, the classifier treated a cluster of multiple cells as a single instance, leading to significant undercounting [21]. This under-segmentation behavior is illustrated in Figure 9, where multiple adjacent cells are merged in the predicted output.

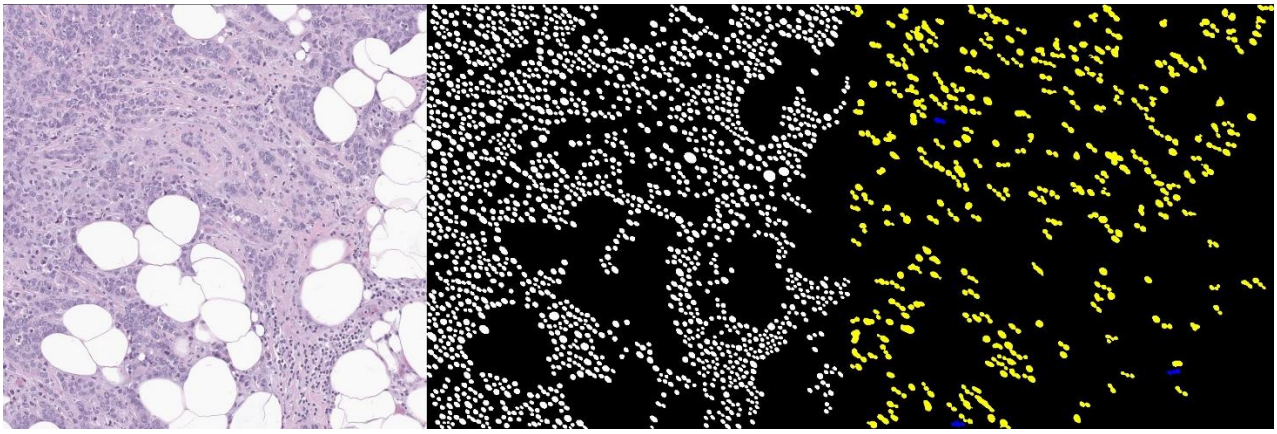

Figure 9. Visualization of the Two-Stage Method Results: (1) Original Microscopic Image, (2) Ground Truth, and (3) Prediction

However, the Single-Stage Multi-class approach (Scenario 5) achieved a significantly lower MAE of 18.46. By learning semantic classes and spatial boundaries simultaneously from high-resolution tiled images, the model successfully delineated individual cells within dense clusters. Although the Two-Stage method showed a slightly lower MAE for negative cells (1.39 vs. 1.66), this difference is marginal and is likely influenced by the class imbalance, where negative cells are relatively sparse. Overall, the Single-Stage approach proves to be more robust for clinical applications requiring accurate proliferation indexing, overcoming the limitations of cascade systems reported in previous studies [22]. To further illustrate the clinical significance of these findings, an MAE of 119.87 for positive cells implies that, on average, the two-stage system miscounts approximately 120 proliferating cells per image—an error magnitude that would render the resulting Ki-67 index clinically unreliable. In contrast, an MAE of 18.46 achieved by the single-stage approach represents a substantially more acceptable margin of error for supporting pathologists in tumor grading.

Table 5 presents a comparison between the proposed method and related studies in the field of automated cell segmentation and counting in histopathological images. Since prior studies employed different datasets and evaluation metrics, direct numerical comparisons are not always feasible. However, the table illustrates the progression of methodological approaches and highlights the unique contribution of this work in terms of end-to-end counting accuracy, measured by the MAE.

Table 5. Comparison of the Proposed Method with Related Studies

Author (Year)	Method	Dataset	Metric	Result
Veta et al. (2013) [21]	Multi-scale watershed segmentation	H&E Breast Cancer Histopathology (multi-organ)	Sensitivity / PPV / Dice	0.853 / 0.886 / ~0.9
Hatipoglu & Bilgin (2017) [22]	Deep learning + spatial relationships		Accuracy / Dice / IoU	90.75% / 83.74% / 83.2%
Chandranegara et al. (2023) [10]	CNN + Transfer Learning (VGG-7)	Breast histopathology	Accuracy / Precision / Recall / F1	98% / 99% / 98% / 98%
Proposed Method (2025)	Multi-class U-Net + Tiling (Scenario 5)	IHC4BC & RSUA	MAE (Positive) / MAE (Negative) / Dice / Mean IoU	18.46 / 1.66 / 0.8740 / 0.6603

4. Conclusion

This study successfully evaluated and compared two deep learning strategies for automated breast cancer cell counting in H&E histopathology images. The experimental results demonstrate that the Single-Stage Multi-Class Segmentation method utilizing the Image Tiling strategy (Scenario 5) is significantly superior to the Two-Stage Segmentation-Classification framework. The single-stage model achieved the most accurate counting performance, with a Mean Absolute Error (MAE) of 18.46 for positive cells and 1.66 for negative cells, effectively preserving high-resolution details and accurately separating overlapping cells.

In contrast, although the Two-Stage approach achieved a high classification accuracy of 98.80% using MobileNetV2 and VGG16, it failed to translate this performance into accurate cell counting. The system exhibited a high counting error (MAE of 119.87 for positive cells), primarily due to the under-segmentation bottleneck in the initial binary segmentation step, where clustered cells were merged and counted as single objects. Additionally, MobileNetV2 proved to be more computationally efficient than VGG16, requiring less than half the training time for equivalent classification accuracy.

For future work, it is recommended to explore instance segmentation architectures such as Mask R-CNN or StarDist to explicitly handle overlapping cells [23]. Furthermore, implementing an overlap-tile strategy during inference may further reduce errors at the patch boundaries. It is also recommended to address class imbalance using advanced augmentation techniques, such as Generative Adversarial Networks (GANs), to enhance the detection sensitivity of minority negative cells [24]. Finally, future research could investigate other lightweight architectures, such as EfficientNet, to further optimize deployment on resource-constrained devices [25].

Acknowledgement

The authors sincerely thank the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia (Kemdiktisaintek) for supporting this research through the BIMA Funding Scheme for the 2025 fiscal year under Contract No. 017/C3/DT.05.00/PL/2025.

References

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," CA: A Cancer Journal for Clinicians, vol. 71, no. 3, pp. 209–249, 2021. https://doi.org/10.3322/caac.21660

[2] N. Harbeck, F. Penault-Llorca, J. Cortés, M. Gnant, N. Houssami, P. Poortmans, K. Ruddy, J. Tsang, and F. Cardoso, "Breast cancer," Nature Reviews Disease Primers, vol. 5, no. 1, pp. 1–31, 2019. https://doi.org/10.1038/s41572-019-0111-2

[3] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," CA: A Cancer Journal for Clinicians, vol. 70, no. 1, pp. 7–30, 2020. https://doi.org/10.3322/caac.21590
[4] K. Tehrani, J. Park, E. Chaney, H. Tu, and S. Boppart, "Nonlinear imaging histopathology: A pipeline to correlate gold-standard hematoxylin and eosin staining with modern nonlinear microscopy," IEEE Journal of Selected Topics in Quantum Electronics, vol. 29, no. 4, pp. 1–12, 2023. https://doi.org/10.1109/JSTQE.2023.3241234
[5] M. Feng, Y. Deng, L. Yang, Q. Jing, Z. Zhang, L. Xu, and H. Bu, "Automated quantitative analysis of Ki-67 staining and HE images recognition and registration based on whole tissue sections in breast carcinoma," Diagnostic Pathology, vol. 15, no. 1, p. 55, 2020. https://doi.org/10.1186/s13000-020-00971-8
[6] J. Nunes, D. Montezuma, D. Oliveira, T. Pereira, and J. Cardoso, "A survey on cell nuclei instance segmentation and classification: Leveraging context and attention," Medical Image Analysis, vol. 99, p. 103360, 2024. https://doi.org/10.1016/j.media.2023.103360
[7] P. Wang, X. Hu, Y. Li, Q. Liu, and X. Zhu, "Automatic cell nuclei segmentation and classification of breast cancer histopathology images," Signal Processing, vol. 122, pp. 1–13, 2016. https://doi.org/10.1016/j.sigpro.2015.11.019
[8] H. Yu, L. Yang, Q. Zhang, D. Armstrong, and M. Deen, "Convolutional neural networks for medical image analysis: State-of-the-art, comparisons, improvement and perspectives," Neurocomputing, vol. 444, pp. 92–110, 2021. https://doi.org/10.1016/j.neucom.2021.03.014
[9] S. Dong, P. Wang, and K. Abbas, "A survey on deep learning and its applications," Computer Science Review, vol. 40, p. 100379, 2021. https://doi.org/10.1016/j.cosrev.2021.100379
[10] D. R. Chandranegara, F. H. Pratama, S. Fajrianur, M. R. E. Putra, and Z. Sari, "Automated Detection of Breast Cancer Histopathology Image Using Convolutional Neural Network and Transfer Learning," MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer, vol. 22, no. 3, pp. 455–468, 2023. https://doi.org/10.30812/matrik.v22i3.2876
[11] D. Albashish, R. Al-Sayyed, A. Abdullah, M. H. Ryalat, and N. A. Almansour, "Deep CNN Model Based on VGG16 for Breast Cancer Classification," in 2021 International Conference on Information Technology (ICIT), Amman, Jordan, 2021, pp. 805–810. https://doi.org/10.1109/ICIT52682.2021.9491631
[12] A. Basu, P. Senapati, M. Deb, R. Rai, and K. G. Dhal, "A survey on recent trends in deep learning for nucleus segmentation from histopathology images," Evolving Systems, pp. 1–46, Mar. 2023. https://doi.org/10.1007/s12530-023-09491-3
[13] K. AnbuDevi and K. Suganthi, "Review of Semantic Segmentation of Medical Images Using Modified Architectures of UNet," Diagnostics, vol. 12, no. 10, p. 2309, 2022. https://doi.org/10.3390/diagnostics12102309
[14] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, Germany, 2015, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
[15] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv preprint arXiv:1704.04861, 2017. https://doi.org/10.48550/arXiv.1704.04861
[16] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv preprint arXiv:1409.1556, 2014. https://doi.org/10.48550/arXiv.1409.1556
[17] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," arXiv preprint arXiv:1412.6980, 2014. https://doi.org/10.48550/arXiv.1412.6980
[18] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 2980–2988. https://doi.org/10.1109/ICCV.2017.324
[19] T. Eelbode, J. Bertels, M. Berman, D. Vandermeulen, F. Maes, R. Bisschops, and M. Blaschko, "Optimization for Medical Image Segmentation: Theory and Practice When Evaluating With Dice Score or Jaccard Index," IEEE Transactions on Medical Imaging, vol. 39, no. 11, pp. 3679–3690, 2020. https://doi.org/10.1109/TMI.2020.2989417
[20] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," BMC Medical Imaging, vol. 15, no. 1, p. 29, 2015. https://doi.org/10.1186/s12880-015-0068-x
[21] M. Veta, P. J. van Diest, R. Kornegoor, A. Huisman, M. A. Viergever, and J. P. W. Pluim, "Automatic Nuclei Segmentation in H&E Stained Breast Cancer Histopathology Images," PLoS ONE, vol. 8, no. 7, p. e70221, 2013. https://doi.org/10.1371/journal.pone.0070221
[22] N. Hatipoglu and G. Bilgin, "Cell segmentation in histopathological images with deep learning algorithms by utilizing spatial relationships," Medical & Biological Engineering & Computing, vol. 55, no. 10, pp. 1829–1848, 2017. https://doi.org/10.1007/s11517-017-1632-3
[23] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 2961–2969. https://doi.org/10.1109/ICCV.2017.322
[24] J. L. Ruiz-Casado, M. A. Molina-Cabello, and R. M. Luque-Baena, "Enhancing Histopathological Image Classification Performance through Synthetic Data Generation with Generative Adversarial Networks," *Sensors*, vol. 24, no. 12, p. 3777, 2024. https://doi.org/10.3390/s24123777
[25] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Proceedings of the 36th International Conference on Machine Learning (ICML), vol. 97, pp. 6105–6114, 2019.