Turkish Journal of Science & Technology
19(2), 485-495, 2024

Research Paper
https://doi.org/10.55525/tjst.1404899

A Deep Learning-based U-Net 3+ Technique for Segmentation Blood Cell

Hasan ULUTAŞ[1*]
[1]Department of Computer Engineering, Yozgat Bozok University, Yozgat
[*1] hasan.ulutas@bozok.edu.tr

(Geliş/Received: 14/12/2023; Kabul/Accepted: 15/07/2024)

Abstract: Segmentation and classification of blood cells are crucial for various medical applications, including disease diagnosis, treatment monitoring, and research purposes. This process allows for accurate identification and quantification of different cell types, aiding in the detection and understanding of various blood-related disorders. The proposed U-Net 3+ architecture incorporates structural modifications, including strengthened connections between convolutional layers, increased filter numbers, and integration of Bayesian optimization for hyperparameter tuning. The model's generalization capability is optimized through the dynamic adjustment of dropout rates and learning rates. Bayesian optimization facilitates the exploration of optimal hyperparameter combinations, allowing the model to adapt effectively to diverse datasets. Advanced training strategies, such as adaptive learning rate adjustment and early stopping, are employed to mitigate overfitting and enhance training efficiency. The proposed model exhibits exceptional performance across multiple folds, achieving low training and validation losses, high accuracy metrics, and robust segmentation indices. Evaluation metrics, including mean IoU (Jaccard Index), dice score, pixel accuracy, and precision, confirm the model's proficiency in accurately delineating blood cell boundaries. The study demonstrates the effectiveness of custom architectures and optimization techniques, achieving an average IoU (Jaccard Index) of 0.9324 and a dice score of 0.9667. The proposed U-Net 3+ model stands as a promising solution for accurate and reliable blood cell segmentation, demonstrating adaptability and robust performance across various datasets. This work sets the stage for future research in the domain of medical image segmentation, emphasizing the potential for continued advancements in precise and efficient segmentation methodologies.

Key words: blood cell, segmentation, U-Net3+, pre-processing, cross validation.

Kan Hücresinin Segmentasyonu için Derin Öğrenmeye Dayalı U-Net 3+ Tekniği

Öz: Kan hücrelerinin segmentasyonu ve sınıflandırılması, hastalık teşhisi, tedavi izleme ve araştırma amaçları dahil olmak üzere çeşitli tıbbi uygulamalar için çok önemlidir. Bu süreç, farklı hücre türlerinin doğru bir şekilde tanımlanmasına ve miktarının belirlenmesine olanak tanıyarak kanla ilgili çeşitli bozuklukların tespit edilmesine ve anlaşılmasına yardımcı olur. Önerilen U-Net 3+ mimarisi, konvolüsyonel katmanlar arasındaki bağlantıları güçlendiren, filtre sayılarını artıran ve hiperparametre ayarlaması için Bayesian optimizasyonu entegre eden yapısal değişiklikler bulunmaktadır. Modelin genelleme yeteneği, dropout oranları ve öğrenme oranlarının dinamik ayarlanması ile optimize edilmiştir. Bayesian optimizasyon, optimal hiperparametre kombinasyonlarını keşfetmeyi sağlayarak modelin çeşitli veri kümelerine etkili bir şekilde uyum sağlamasına imkan tanır. Ayrıca, aşırı uydurmayı azaltmak ve eğitim verimliliğini artırmak için adaptif öğrenme oranı ayarı ve erken durdurma gibi gelişmiş eğitim stratejileri kullanılmıştır. Önerilen model, çoklu katmanlarda düşük eğitim ve doğrulama kayıpları, yüksek doğruluk metrikleri ve güçlü segmentasyon endeksleri elde ederek olağanüstü performans sergilemektedir. Değerlendirme metrikleri, ortalama IoU (Jaccard İndeksi), dice skoru, piksel doğruluğu ve hassasiyet gibi, modelin kan hücre sınırlarını doğru bir şekilde belirleme konusundaki yetkinliğini doğrular. Çalışma, özel mimarilerin ve optimizasyon tekniklerinin etkinliğini 0,9324 ortalama IoU (Jaccard İndeksi) ve 0,9667 dice skoru ile ispatlamaktadır. Önerilen U-Net 3+ modeli, çeşitli veri kümelerinde adaptasyon yeteneği ve güçlü performansıyla umut vadeden bir çözüm olarak ön plana çıkmaktadır. Bu çalışma, medikal görüntü segmentasyonu alanında gelecekteki araştırmalara zemin oluşturarak, hassas ve etkili segmentasyon metodolojilerinde devam eden ilerlemelerin potansiyelini vurgulamaktadır.

Anahtar kelimeler: kan hücresi, segmentasyon, UNet3+, ön işleme, çapraz doğrulama.

1. Introduction

Blood cells traverse the body through the circulatory system, and in mammalian species, they can be categorized into three main types [1]. Firstly, red blood cells specialize in the vital task of transporting oxygen to various tissues and organs. Secondly, leukocytes, which are essential components of the immune system, respond to bacterial infections by concentrating at the infection site, encircling the invading bacteria, and engulfing these bacteria as part of the body's defense mechanism. Lastly, platelets play a pivotal role in hemostasis, contributing to the blood clotting process and preventing excessive bleeding in response to injuries. Together, these three types of blood cells work in harmony to ensure the proper functioning and health of the circulatory system. The shape, number, and type of these cellular components have always been a topic of interest among researchers [2].

* Corresponding author: hasan.ulutas@bozok.edu.tr. ORCID Number of author: 0000-0003-3922-934X

Especially in the field of medicine, the impact of blood components is crucial, addressing conditions ranging from easily treatable ones like anemia to more challenging diseases such as leukemia. Microscopes are used to examine the anatomical and morphological structures of these components. However, manual processes performed on images obtained through microscopes can be tedious and time-consuming due to their reliance on human expertise. To minimize human errors, achieve time efficiency, and facilitate work on a larger number of samples, the use of computer-aided systems in medical studies is increasingly on the rise [3]–[6].

The accurate classification and segmentation of blood cells are crucial for enhancing the accuracy of medical diagnoses and optimizing treatment processes. In this context, automatic or semi-automatic computer-aided methods provide valuable tools for researchers and healthcare professionals. In the literature, a variety of studies focused on the classification and segmentation of blood cells [7]–[10]. Macawile et al. employed pre-trained neural networks, including AlexNet, GoogLeNet, and ResNet-101, to discern and enumerate white blood cells [10]. Comparative analysis of the outcomes from these three networks revealed that AlexNet demonstrated superior performance. The study reported a comprehensive accuracy rate of 96.63%. Yıldırım and Çinar performed the classification of white blood cells for disease diagnosis using established deep learning models such as AlexNet, ResNet-50, DenseNet-201, and GoogLeNet, utilizing a pre-existing dataset [11]. The initial phase involved the classification of the original data with these four networks, followed by the computation of accuracy metrics. The most favorable outcome, featuring an accuracy rate of 83.44%, was noted when employing DenseNet-201 in tandem with Gaussian filtering. In a study conducted by Lu et al., a network named WBC-Net, based on UNet++ and ResNet, was employed for the segmentation of white blood cells [12]. The segmentation was tested and compared both with and without skip connections, yielding precision values of 98.48%, 98.90%, 96.20%, and 94.03% for four respective datasets. In the investigation conducted by Roy et al., the localization and identification of leukocytes in peripheral blood were accomplished through the application of deep learning methodologies. The segmentation of leukocytes utilized the DeepLabv3+ architecture. During the segmentation testing phase, the average accuracy achieved was 98.22%. Subsequently, the overall average accuracy during the validation process was computed as 98.87% ± 1 [13]. Toptas and Hanbay conducted a study on the automatic detection, classification, and segmentation of blood cells [2]. In this research, microscopic blood cell images were segmented using deep learning architectures such as DeepLabv3+, U-Net, and FCN. The highest accuracy, with a value of 95.75%, was achieved with the DeepLabv3+ architecture. These experimental results provide support for the effectiveness of the proposed method. Bozkurt proposed a study on the classification of blood cells using a Dense Convolutional Network [14]. This research was conducted on an open-access BCCD dataset consisting of 12,507 white blood cell images. Experiments demonstrate that the proposed DenseNet121 model achieved an accuracy of 94%. This high accuracy value can expedite diagnosis by automating cell classification, enabling doctors to examine more data. The study presented by Nahzat et al. aims to develop a CNN-based model for the classification of white blood cells and evaluate its performance [15]. Using white blood cell images obtained from Kaggle, the study demonstrates that the RMSprop optimizer yields the best results. Additionally, the proposed model is compared with four pre-trained models including MobileNetV2, DenseNet121, InceptionV3, and ResNet50. The results indicate that despite having the lowest number of trainable parameters and training time, the proposed model exhibits exceptional performance. Khouani et al. proposed an automated method for identifying white blood cells (WBCs) in cytological images, crucial for cancer diagnosis [16]. It introduces a deep learning-based approach to streamline tasks for hematologists. The method involves preprocessing images, applying a deep neural network for localization and segmentation, and refining outputs with combined predictions and corrections. Additionally, a novel algorithm leveraging spatial information enhances segmentation quality. Implemented using Python, TensorFlow, and Keras, the method shows promising results on datasets from Tlemcen Hospital, Algeria, achieving 95.73% accuracy with predictions in less than 1 second, outperforming existing methods. Zhang et al. present a semantic segmentation approach based on deep learning, targeting the complexities arising from variations in cell characteristics and image quality [17]. They introduce the deformable U-Net (dU-Net), which incorporates deformable convolution layers into the U-Net structure to address these challenges. The dU-Net demonstrates precise localization and resilience to diverse cell shapes and image conditions. Experiments on microscopic RBC images from SCD patients show that the dU-Net achieves superior accuracy in both binary and multi-class segmentation tasks compared to unsupervised and state-of-the-art supervised methods.

This study aims to introduce an artificial intelligence model developed to facilitate accurate segmentation of blood cells in medical imaging. The developed model features a novel architecture named U-Net 3+, which incorporates specific features not seen in previous studies. The U-Net 3+ model stands out by enhancing the existing UNet architecture with structural modifications such as reinforced connections between deeper layers, increased filter numbers, and integration of Bayesian optimization for hyperparameter tuning. The primary contribution of this study lies in presenting a new approach to accurately segmenting blood cells in medical imaging. The developed U-Net 3+ model has demonstrated higher accuracy and segmentation performance

compared to other existing methods. This study provides a significant contribution to the development of artificial intelligence-based segmentation techniques in medical imaging.

The article's remaining chapters are organized as follows: Chapter 2 provides information of dataset, image preprocessing, a summary of the deep learning methods such as U-Net, proposed U-Net3+ and evaluation metrics. The findings from U-Net3+ segmentation methods are presented in Chapter 3. The results are discussed in Chapter 4, and conclusions and explanations are provided in Chapter 5.

2. Material and Methods

2.1. Dataset

In this study, a dataset comprising publicly available blood cell images is utilized, totaling 1,328 images [18], [19]. The dataset used originally contains bounding box coordinates for each cell, and the original references of the cells were used to identify individual cells, and detailed information about the dataset is available in [19]. The dataset includes 1328 original blood cell images and an equal number of corresponding real background images. The images vary in size, with some measuring 1600x1200x3 pixels and others 1944x1383x3 pixels. All images in the dataset are in the RGB colour space. Sample images from the dataset are illustrated in Figure 1.

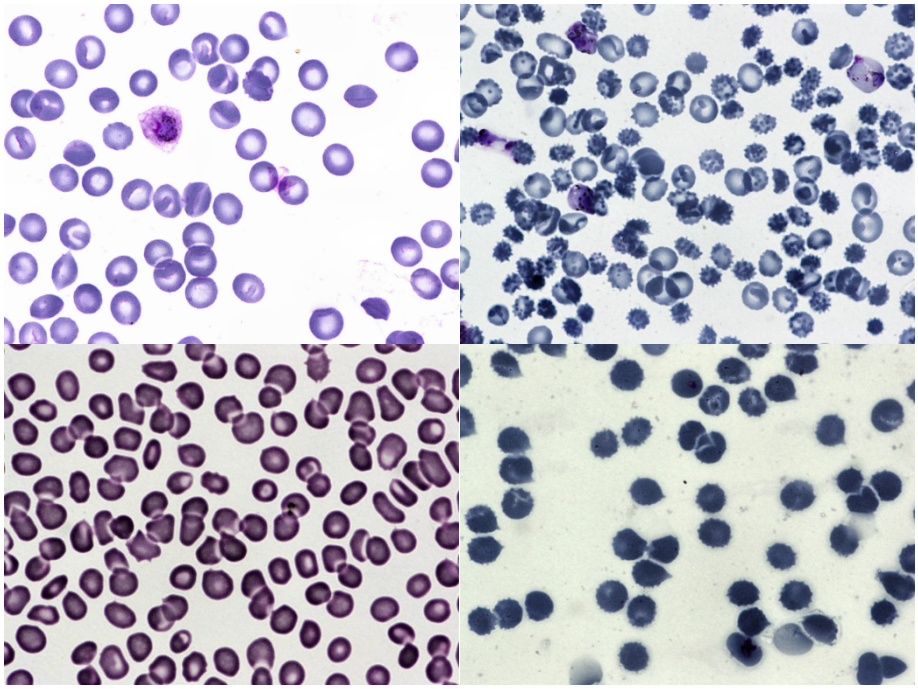

Figure 1. Original blood smear images from the used dataset

2.2. Image Pre-processing

Image pre-processing is a fundamental step in machine learning and image processing applications because these steps prepare the necessary data to obtain more accurate results. Pre-processing also includes operations such as reducing noise, increasing contrast, and adjusting image size, resulting in more reliable and consistent results. The basis of this study is to subject the images in the data set to certain image pre-processing steps before the analysis process.

Image Loading and Resizing: Each image and mask is loaded and resized to a specific target_size (256x256x3). This process standardizes the images to a fixed size, ensuring a consistent input size during the model training.

$$[R'(x,y) = R\left(\frac{x \cdot W_{original}}{W_{target}}, \frac{y \cdot H_{original}}{H_{target}}\right)] \quad (1)$$

In Equation (1), (R) represents the original image, (R') represents the resized image, ($W_{original}$) ve ($H_{original}$) represent the original width and height of the image, and (W_{target}) ve (H_{target}) represent the target width and height dimensions.

Normalization: Image and mask values are scaled from the [0, 255] range to the [0, 1] range. This enables the model to train faster and more stably because smaller input values provide less variance in weight updates and better gradient flow within the network.

$$[R''(i,j) = \frac{R'(i,j)}{255}] \quad (2)$$

In Equation (2), (R'')represents the image after normalization, (R') represents the resized image, and (i) and (j) represent pixel positions.

Grayscale Masking: Masks are loaded in grayscale, meaning they have only one channel without color information. This is because segmentation masks typically need to contain binary values (0 or 1) indicating the presence or absence of an object, and color information is not necessary for this purpose.

2.2.1. U-Net3+ and proposed model

The U-net, proposed by Ronneberger et al. in 2015 [20], is primarily a convolutional neural network architecture designed for image segmentation. The U-Net architecture is a developed image segmentation algorithm, primarily designed for tasks involving segmentation, classification, and image operations [21]. The training duration of U-Net is relatively short, featuring a simple design and fewer parameters. It requires less training data compared to other networks. U-net is symmetric, incorporating skip connections between the downsampling and upsampling paths.

The integration of features across various scales is a pivotal aspect of achieving precise segmentation. U-Net++, a modified iteration of U-Net distinguished by its embedded and densely connected skip connections, has been devised to address this need [22]. However, despite its advancements, this model still lacks in capturing adequate information, signalling a considerable scope for further refinement. In response to this, the proposed U-Net3+ architecture assumes heightened significance, strategically incorporating full-scale skip connections and introducing deep supervisions. These full-scale skip connections adeptly merge low-level details from feature maps across different scales with high-level semantic information. Concurrently, the deep supervisions effectively learn hierarchical representations from fully aggregated feature maps. Beyond the notable improvements in accuracy, U-Net 3+ also exhibits promising potential in enhancing computational efficiency by judiciously reducing network parameters, marking a comprehensive advancement in the field of image segmentation. Details of the architecture is given in Table 1. Figure 2 provides simplified summaries of U-Net, U-Net++, and U-Net 3+.

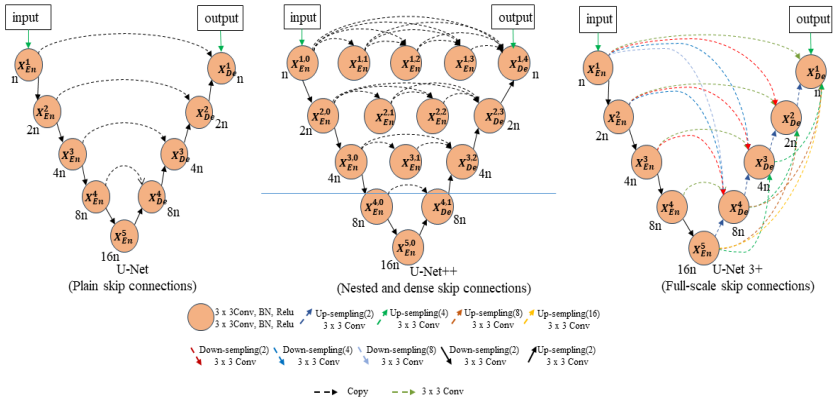

Figure 2. U-Net, U-Net++, and U-Net 3+ architecture [23].

Table 1. Details of the architecture

Layer	Layer Type	Details
Input Layer	Input	The input layer accepts input of size (256, 256, 3), indicating images with a resolution of 256 pixels in width, 256 pixels in height, and 3 color channels (RGB).
Conv Block 1	Conv2D + Activation (ReLU) + (Optional) Dropout	The first convolutional block utilizes a 3x3 kernel with 128 filters. It includes batch normalization, ReLU activation, and dropout.
MaxPooling	MaxPooling2D	A 2x2 MaxPooling operation is applied after the first convolutional block.
Conv Block 2	Conv2D + Activation (ReLU) + (Optional) Dropout	The second convolutional block employs a 3x3 kernel with 256 filters. It includes batch normalization, ReLU activation, and dropout to enhance the model's learning and generalization capabilities.
MaxPooling	MaxPooling2D	A 2x2 MaxPooling operation is applied
Conv Block 3	Conv2D + Activation (ReLU) + (Optional) Dropout	The third convolutional block uses a 3x3 kernel with 512 filters. It incorporates batch normalization, ReLU activation, and dropout to enhance the model's capacity to learn complex patterns and improve its generalization abilities.
MaxPooling	MaxPooling2D	A 2x2 MaxPooling operation is applied
Conv Block 4	Conv2D + Activation (ReLU) + (Optional) Dropout	The fourth convolutional block employs a 3x3 kernel with 1024 filters. It includes batch normalization, ReLU activation, and dropout to capture intricate features and further enhance the model's ability to generalize.
MaxPooling	MaxPooling2D	A 2x2 MaxPooling operation is applied
Bridge	Conv2D + Activation (ReLU)	The bridge layer, or skip connection, uses a 3x3 kernel with 2048 filters and includes ReLU activation. Skip connections are often used to facilitate the flow of information between the encoder and decoder parts of the U-Net architecture, aiding in the preservation of spatial information.
Upsample Block 1	UpSampling2D + Concatenate + Conv Block	An upsampling and concatenation operation is performed, followed by the application of a convolutional block. Upsampling is typically used to increase the spatial resolution of the feature maps.
Upsample Block 2	UpSampling2D + Concatenate + Conv Block	An upsampling and concatenation operation is performed, followed by the application of a convolutional block.
Upsample Block 3	UpSampling2D + Concatenate + Conv Block	An upsampling and concatenation operation is performed, followed by the application of a convolutional block.
Upsample Block 4	UpSampling2D + Concatenate + Conv Block	An upsampling and concatenation operation is performed, followed by the application of a convolutional block.
Output Layer	Conv2D + Activation (Sigmoid/Softmax)	The final layer uses a 1x1 kernel based on the number of classes. It includes either sigmoid or softmax activation, depending on the nature of the task
Total params: 125,547,521, Trainable params: 125,523,969, Non-trainable params: 23,552		

2.3. Evaluation metrics

Rigorous assessment of the U-Net 3+ architecture for blood cell segmentation involves the application of a comprehensive suite of evaluation metrics, addressing considerations of accuracy. The Equations (3)- (6) [24]:

- TP represents the number of pixels that truly belong to class P and are correctly classified as P by the model.
- TN denotes the count of pixels that truly belong to class N and are accurately classified as N by the model.
- FN signifies the number of pixels that actually belong to class P but are misclassified as N by the model.
- FP indicates the count of pixels that truly belong to class N but are misclassified as P by the model.

$$Accuracy = \frac{T_N+T_p}{T_N+T_p+F_N+F_p} \quad (3)$$

$$Precision = \frac{T_p}{T_p+F_p} \quad (4)$$

$$Recall = \frac{T_p}{T_p+F_N} \quad (5)$$

$$F1 - Score = \frac{2xPxR}{P+R} \quad (6)$$

The Jaccard Index, also known as the Intersection over Union (IoU) value, provides the percentage of overlap between predicted masks and ground truth masks. It is a metric that quantifies the similarity of two sets by measuring the intersection divided by the union and given in Equation (7) [25].

$$Jaccard\ (JAC) = \frac{TP}{TP+FN+FP} \quad (7)$$

The Dice Index is a metric that measures the similarity between predicted and underlying ground truth images, as expressed in Equation (8) [25].

$$Dice = \frac{2TP}{2TP+FN+FP} \quad (8)$$

The critical metric "Pixel Accuracy" is employed to evaluate the performance of segmentation models. This metric measures how well the model's predictions align with the ground truth data. Pixel Accuracy is calculated by dividing the number of pixels correctly classified by the model by the total number of pixels [26]. In other words, it quantifies the concordance between the predicted mask generated by the model and the actual mask and given in Equation (9).

$$Pixel\ Accuracy = \frac{pixels\ correctly\ classified}{the\ total\ number\ of\ pixels} \quad (9)$$

3. Experimental Results

In customizing the U-Net3+ architecture, various modifications are implemented to enhance segmentation accuracy and create a structure resistant to overfitting while preserving the fundamental principles of the U-Net3+ architecture. Among the changes made are strengthening the connections between convolutional layers, increasing the number of filters, and integrating hyperparameter optimization mechanisms. Specifically, dropout and learning-rate ratios were adjusted using Bayesian optimization to increase the model's generalization ability. This method provides flexibility in determining the optimal network structure and parameters during the model's adaptation to different datasets and noise profiles. Furthermore, advanced training strategies such as adaptive learning rate adjustment and early stopping are adopted to improve the efficiency of the training process. These strategies serve as mechanisms supporting the model's rapid convergence and protection against overfitting, contributing to more robust and reliable segmentation results. The flowchart of the system is given below:

1. *Data Loading and Pre-processing*
    - Images and masks are loaded.
    - Each image and mask are resized to a specified dimension.
    - Pixel values are normalized to a range between 0 and 1.
2. *Data Splitting*
    - Images and masks are split into training and testing sets.
3. *Model Creation (U-Net 3+)*
    - U-Net 3+ model is constructed.
    - The model is configured with a specific number of layers and filters.
4. *Hyperparameter Optimization (Bayesian Optimization)*
    - Optimal values for certain hyperparameters (e.g., dropout rate and learning rate) are searched.
5. *Model Training*

- The model is trained on different subsets of data using the k-Fold Cross-Validation method.
- Callbacks such as ModelCheckpoint, ReduceLROnPlateau, and EarlyStopping are utilized during training.

6. *Model Evaluation*
   - The trained model is evaluated on the test dataset.
   - Performance metrics (e.g., accuracy, precision, IoU, Dice Score) are computed.
7. *Analysis and Visualization of Results*
   - Training and validation losses, accuracies, and other metrics are visualized on graphs.
   - Model predictions on the test set are compared with actual labels for visual analysis.

Bayesian Optimization is a statistical method used to adjust the hyperparameters of our model, in this case, dropout_rate and learning_rate, to maximize the performance of our deep learning model. The goal of this method is to maximize the accuracy of our model on the validation dataset.
Firstly, the range of values to be explored for the model's hyperparameters is being defined. For example, a range can be set between 0 and 0.5 for the dropout rate, and for the learning rate (10^{-5}) and (10^{-2}). Next, the Bayes theorem is used to predict how the model will perform with these hyperparameter values. This theorem allows us to calculate the probability of the model's performance given certain hyperparameter values. Mathematically, this is expressed using the Bayes theorem as given in Equation (10):

$$[P(performance|hyperparameters) = \frac{P(hyperparameters|performance)\times P(performance)}{P(hyperparameters)} \quad (10)$$

In this formula, each term represents, respectively, the probability of the model's performance given the hyperparameters, the probability of hyperparameters given the model's performance, and the marginal distributions of these two probabilities. Finally, using this probability distribution, the most suitable hyperparameter combinations are determined for our model. This process ensures the adjustment of hyperparameters in a way that maximizes the performance of our model on the validation set. In summary, Bayesian Optimization provides us with the opportunity to fully harness the potential of our model. With this method, we can effectively manage the complex structure of our deep learning model and optimize its performance. Parameters obtained with Bayesian optimization are listed in Table 2.

Table 2. Obtained parameters with Bayesian optimization

	Dropout	**Learning Rate**
Initial Value	0.1	0.00001
Value after optimization	0.184389873055315	0.004261879524120684

ReduceLROnPlateau and EarlyStopping are callback methods utilized to optimize the training process of the U-Net 3+ model in our paper, with the goal of preventing issues like overfitting.

ReduceLROnPlateau (Reducing Learning Rate): ReduceLROnPlateau is employed to dynamically adjust the model's learning rate during training. Throughout the training process, if there is no improvement in a specified metric (e.g., validation loss) for a certain number of epochs, it reduces the learning rate. This method aids the model in making finer adjustments and prevents it from getting stuck in local minima.

The proposed U-Net 3+ model demonstrates excellent performance across multiple folds, as evidenced by the results presented in Table 3. Throughout the training process, the model consistently achieves low training and validation losses, indicative of its ability to minimize errors and generalize well to unseen data. The high values for accuracy, precision, recall, and area under the curve (AUC) further underscore the model's proficiency in accurately segmenting blood cells. Notably, the model maintains a high level of precision and recall, striking a balance between minimizing false positives and false negatives. The consistency of these results across different folds highlights the robustness and reliability of the proposed model. Results of the segmentation are shown in Figure 3.

Table 3. Results of the proposed model

fold	train_loss	val_loss	train_ accuracy	val_ accuracy	train_ precision	val_ precision	train_ recall	val_ recall	train_ auc	val_ auc
1	0.0334	0.0338	0.9798	0.9792	0.96047	0.9623	0.9720	0.9699	0.9814	0.9807
2	0.0331	0.0344	0.9797	0.9802	0.96041	0.9624	0.9732	0.9670	0.9812	0.9799
3	0.0333	0.0342	0.9798	0.9788	0.96005	0.9566	0.9725	0.9748	0.9820	0.9820
4	0.0328	0.0341	0.9799	0.9794	0.96117	0.9607	0.9723	0.9700	0.9817	0.9808
5	0.0321	0.0361	0.9805	0.9779	0.96089	0.9636	0.9740	0.9650	0.9817	0.9776

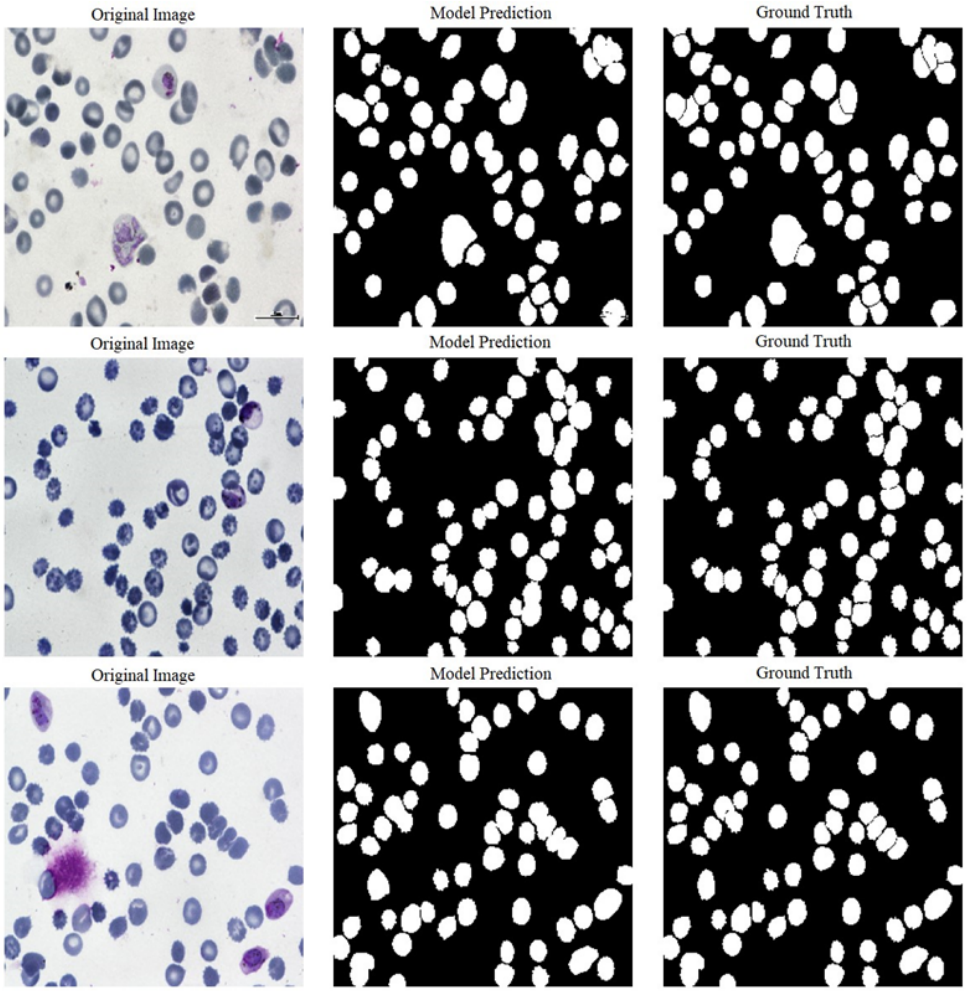

Figure 3. Results of the segmentation

In Table 4, the test results reinforce the model's effectiveness in segmentation tasks. The Mean IoU (Jaccard Index) of 0.9324 indicates a substantial overlap between the predicted and ground truth masks, emphasizing the model's ability to delineate cell boundaries accurately. The Dice score of 0.9667 further corroborates the model's segmentation accuracy. The Pixel Accuracy of 0.9796 signifies a high proportion of correctly classified pixels, attesting to the model's precision at the pixel level. Additionally, the exceptionally high precision value of 0.9968 underscores the model's capability to minimize false positives, crucial in medical image segmentation where accuracy is paramount. These results suggest that the proposed U-Net 3+ model excels in blood cell segmentation, providing accurate and reliable predictions across various evaluation metrics. The combination of low losses, high accuracy, and robust segmentation metrics positions the model as a promising solution for medical image segmentation tasks. In conclusion, the achieved performance metrics, particularly the high accuracy, precision, and spatial overlap metrics, demonstrate the effectiveness of the U-Net 3+ model for blood cell segmentation. The model exhibits robust generalization across different folds, showcasing its potential for accurate and reliable

segmentation of blood cell images. Loss and accuracy graphs for average training and validation are given in Figure 4.

Table 4. Test results of the proposed system

Metrics	Value
Mean IoU (jaccard index)	0.9324
Dice score	0.9667
Pixel accuracy	0.9796
Precision	0.9968

Figure 4. Loss and accuracy graphs for average training and validation

Table 5. Assessment of segmentation performance across various algorithms on the same dataset

Method	Dice-coefficient (%)	Jaccard Index (%)	Reference
Otsu's method	92.60	86.50	[19]
BHT	52.50	49.48	
Watershed	78.21	68.21	
U-Net	93.09	87.16	
U-Net++	88.80	81.44	
TernausNet	93.38	87.65	
R2U-Net	86.76	77.77	
Attention U-Net	91.00	83.73	
Attention R2U-Net	78.52	65.28	
FCN	85.41	75.29	
U-Net 3+	96.67	93.24	Prosed Model

Our model exhibits a significant advantage over other methods in terms of segmentation performance. For instance, in terms of both Dice coefficient and Jaccard index, our U-Net 3+ model outperforms all other methods. Specifically, while our U-Net 3+ model achieves a Dice coefficient of 96.67% and a Jaccard index of 93.24%, its closest competitor, TernausNet, achieves values of 93.38% and 87.65% respectively. Other methods in the comparison include Otsu's method, BHT, Watershed, U-Net++, R2U-Net, Attention U-Net, Attention R2U-Net, and FCN. Each of these methods employs different approaches to fulfill specific segmentation tasks. Additionally, it is crucial to understand why our model outperforms others by evaluating the performance of other methods in comparative analysis. Factors such as an enhanced architecture, more effective feature extraction, or optimized

hyperparameters could contribute to the success of our model. This way, the contribution of the study and the value of our model can be better understood.

4. Discussion

This chapter delves into the findings, providing an in-depth analysis of the performance of the enhanced U-Net 3+ model for blood cell segmentation. The model, which incorporates advancements such as strengthened connections and an increased number of filters, has demonstrated notable improvements in accuracy and precision across multiple folds and the test dataset. These structural enhancements are critical in enabling the model to minimize errors and exhibit robust generalization to novel data. In particular, the bolstered connections facilitate more effective information flow within the network, while the augmented filter numbers enable finer feature extraction, which are both essential for precision in segmenting intricate blood cell structures. The use of Bayesian optimization for hyperparameter tuning is a key aspect of refining the model. This optimization strategy, by systematically adjusting parameters such as dropout and learning rates, has significantly contributed to the adaptability and accuracy of the model. The emphasis on Bayesian methods underscores our commitment to leveraging advanced statistical techniques to fine-tune the model's performance, ensuring optimal generalization across diverse datasets. The achieved mean IoU (Jaccard Index) of 0.9324 and Dice score of 0.9667 serve as testaments to the model's efficacy in accurately predicting cell masks, thereby highlighting its potential in medical image analysis. These metrics reflect not only the model's precision in delineating cell boundaries but also its capability to achieve consistency in segmentation tasks. The exceptional Pixel Accuracy of 0.9796 and high precision value of 0.9968 further highlight the model's adeptness at minimizing false positives, a critical aspect in the realm of medical image segmentation, where accuracy is of paramount importance. This study positions the proposed U-Net 3+ model as a state-of-the-art solution for blood cell segmentation, representing a considerable advancement in the field of deep-learning-based medical image segmentation. The model's ability to accurately and reliably segment images is a significant step forward, and its performance aligns with and extends existing literature, demonstrating the profound impact of advanced architectures and optimization techniques in achieving superior results. However, it is essential to recognize the limitations of this study to gain a comprehensive understanding. The model's performance may be influenced by the specific characteristics of the dataset, highlighting the need for further validation across diverse datasets. Variations in image quality and potential biases within the dataset can impact the model's generalization capabilities. Future research directions should explore transfer-learning approaches to enhance the model's adaptability to different imaging modalities and investigate its applicability in various medical imaging scenarios.

5. Conclusions

This study introduces a novel and customized U-Net 3+ model for blood cell segmentation, integrating substantial structural improvements and advanced optimization techniques. The model has demonstrated exceptional performance, as evidenced by consistently low training and validation losses, high accuracy metrics, and robust segmentation indices. The use of Bayesian optimization for hyperparameter tuning has been instrumental in enhancing the generalization capabilities of the model. The achieved Mean Intersection over Union (IoU) of 0.9324 and Dice score of 0.9667 highlight the model's proficiency in accurately predicting cell masks, emphasizing its potential utility in medical image analysis. The pixel accuracy of 0.9796 and the precision of 0.9968 further illustrate the model's precision at both the pixel and class levels, which is crucial in medical scenarios where accuracy is paramount.

The discussion provides a nuanced analysis, recognizing the model's strengths while acknowledging potential limitations and suggesting directions for future research. This study contributes significantly to the field of deep learning-based medical image segmentation, demonstrating the effectiveness of customized architectures and optimization strategies. In conclusion, the U-Net 3+ model, with its high accuracy and reliability for segmenting blood cells, has emerged as an innovative and effective solution in this domain, setting a benchmark for future advancements.

References

[1] Zhu Z, et al. RETRACTED: BCNet: A Novel Network for Blood Cell Classification. Front Cell Dev Biol. 2002; (9): 813996.

[2] Toptaş M and Hanbay D. Mikroskobik Kan Hücre Görüntülerinin Güncel Derin Öğrenme Mimarileri ile Bölütlemesi. Mühendislik Bilimleri ve Araştırmaları Dergisi, 2023; 5(1): 135-141.

[3] Habibzadeh M, Jannesari M, Rezaei Z, Baharvand H, Totonchi M. Automatic white blood cell classification using pre-trained deep learning models: Resnet and inception. InTenth international conference on machine vision (ICMV 2017) 13 April 2018: SPIE. Vol. 10696, pp. 274-281.

[4] Sahin ME. Image processing and machine learning-based bone fracture detection and classification using X-ray images. Int J Imaging Syst Technol. May 2023; 33(3): 853-65.

[5] Şahin ME. A Deep Learning-Based Technique for Diagnosing Retinal Disease by Using Optical Coherence Tomography (OCT) Images. Turkish Journal of Science and Technology. 1 July 2022; 17(2): 417-26.

[6] Ulutas H, Sahin ME, Karakus MO. Application of a novel deep learning technique using CT images for COVID-19 diagnosis on embedded systems. Alexandria Eng J. 1 July 2023; 74: 345-58.

[7] Alam MM, Islam MT. Machine learning approach of automatic identification and counting of blood cells. Healthcare Technol Lett. August 2019; 6(4): 103-8.

[8] Shahin AI, Guo Y, Amin KM, Sharawi AA. White blood cells identification system based on convolutional deep neural learning networks. Comput Methods Programs Biomed. 1 January 2019; 168: 69-80.

[9] Banik PP, Saha R, Kim KD. An automatic nucleus segmentation and CNN model based classification method of white blood cell. Expert Syst. Appl. 1 July 2020; 149: 113211.

[10] Macawile MJ, Quiñones VV, Ballado A, Cruz JD, Caya MV. White blood cell classification and counting using convolutional neural network. In2018 3rd International conference on control and robotics engineering (ICCRE) 20 April 2018: IEEE. pp. 259-263.

[11] Yildirim M, Çinar A. Classification of white blood cells by deep learning methods for diagnosing disease. Rev d'Intelligence Artif. November 2019; 33(5): 335-40.

[12] Lu Y, Qin X, Fan H, Lai T, Li Z. WBC-Net: A white blood cell segmentation network based on UNet++ and ResNet. Applied Soft Computing. 1 March 2021; 101: 107006.

[13] Reena MR, Ameer PM. Localization and recognition of leukocytes in peripheral blood: A deep learning approach. Comput. Biol. Med. 1 November 2020; 126: 104034.

[14] Bozkurt F. Classification of blood cells from blood cell images using dense convolutional network. Journal of Science, Technology and Engineering Research. November 2021; 2(2): 81-8.

[15] Nahzat S, Bozkurt F, Yağanoğlu M. White blood cell classification using convolutional neural network. Journal of Science, Technology and Engineering Research. 2022; 3(1): 32-41.

[16] Khouani A, El Habib Daho M, Mahmoudi SA, Chikh MA, Benzineb B. Automated recognition of white blood cells using deep learning. Biomed Eng Lett. August 2020; 10: 359-67.

[17] Zhang M, Li X, Xu M, Li Q. Automated semantic segmentation of red blood cells for sickle cell disease. IEEE J Biomed Health Inf. 22 June 2020; 24(11): 3095-102.

[18] Dataset, https://github.com/Deponker/Blood-cell-segmentation-dataset.

[19] Depto DS, Rahman S, Hosen MM, Akter MS, Reme TR, Rahman A, Zunair H, Rahman MS, Mahdy MR. Automatic segmentation of blood cells from microscopic slides: a comparative analysis. Tissue and Cell. 1 December 2021; 73: 101653.

[20] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. InMedical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, 5-9 October 2015; Munich, Germany: proceedings, part III, Springer International Publishing. pp. 234-241.

[21] Yan X, Tang H, Sun S, Ma H, Kong D, Xie X. After-unet: Axial fusion transformer unet for medical image segmentation. InProceedings of the IEEE/CVF winter conference on applications of computer vision 2022 pp. 3971-3981.

[22] Huang H, Lin L, Tong R, Hu H, Zhang Q, Iwamoto Y, Han X, Chen YW, Wu J. Unet 3+: A full-scale connected unet for medical image segmentation. InICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP) 4 May 2020: IEEE. pp. 1055-1059.

[23] Deng Y, Hou Y, Yan J, Zeng D. ELU-net: An efficient and lightweight U-net for medical image segmentation. IEEE Access. 31 March 2022; 10: 35932-41.

[24] Sahin ME. Deep learning-based approach for detecting COVID-19 in chest X-rays. Biomed Signal Process Control. 1 September 2022; 78: 103977.

[25] Vasconcelos FF, Medeiros AG, Peixoto SA, Reboucas Filho PP. Automatic skin lesions segmentation based on a new morphological approach via geodesic active contour. Cognit Syst Res. 1 June 2019; 55: 44-59.

[26] Liu J, Yildirim O, Akin O, Tian Y. AI-driven robust kidney and renal mass segmentation and classification on 3D CT images. Bioengineering. 13 January 2023; 10(1): 116.