

Article

Improved Segmentation of Cellular Nuclei Using UNET Architectures for Enhanced Pathology Imaging

Simão Castro , Vitor Pereira  and Rui Silva * 

The Center for Research in Organizations Markets and Industrial Management (COMEGI), Universidade Lusíada, 1349-001 Lisboa, Portugal; d1364@fam.ulusiada.pt (S.C.)

* Correspondence: rsilva@fam.ulusiada.pt

Abstract: Medical imaging is essential for pathology diagnosis and treatment, enhancing decision making and reducing costs, but despite various computational methodologies proposed to improve imaging modalities, further optimization is needed for broader acceptance. This study explores deep learning (DL) methodologies for classifying and segmenting pathological imaging data, optimizing models to accurately predict and generalize from training to new data. Different CNN and U-Net architectures are implemented for segmentation tasks, with their performance evaluated on histological image datasets using enhanced pre-processing techniques such as resizing, normalization, and data augmentation. These are trained, parameterized, and optimized using metrics such as accuracy, the DICE coefficient, and intersection over union (IoU). The experimental results show that the proposed method improves the efficiency of cell segmentation compared to networks, such as U-NET and W-UNET. The results show that the proposed pre-processing has improved the IoU from 0.9077 to 0.9675, about 7% better results; also, the values of the DICE coefficient obtained improved from 0.9215 to 0.9916, about 7% better results, surpassing the results reported in the literature.

Keywords: medical imaging; computer-aided diagnostics; machine learning; convolutional neural networks



Citation: Castro, S.; Pereira, V.; Silva, R. Improved Segmentation of Cellular Nuclei Using UNET Architectures for Enhanced Pathology Imaging.

Electronics **2024**, *13*, 3335. <https://doi.org/10.3390/electronics13163335>

Academic Editor: Beiwen Li

Received: 22 July 2024

Revised: 13 August 2024

Accepted: 21 August 2024

Published: 22 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Medical imaging plays an extremely important role in the diagnosis and treatment of pathologies. The quality of these images as well as their processing improves medical decision making, allowing for an early detection of pathologies and, thus, contributing to a reduction in the costs of subsequent treatments [1–5]. Several studies have been proposed to improve the diagnosis and solve specific problems in different medical imaging modalities using computational studies [6–9]; however, their optimization is required so that their results are generally accepted [10,11]. The rapid advancement of image processing techniques leveraging neural networks (NNs) has revolutionized medical diagnostics and treatment planning [6–9]. From identifying intricate patterns in medical images to segmenting diseased regions with remarkable accuracy, NN-based methodologies have emerged as powerful tools in modern healthcare. This intersection of artificial intelligence and medical imaging not only enhances the precision of diagnosis but also streamlines treatment pathways, ultimately leading to improved patient outcomes.

Convolutional neural networks (CNNs) represent a specific type of neural network specialized in data processing and have a known topology, such as time-series tables with a 1D topology and images with data in topology [12]. Two-dimensional CNNs represent a neural network architecture, which uses the mathematical operation convolution instead of multiplication in at least one of its layers [13]. Within the different families of neural network architectures in DL, CNNs are a unique family of specific architectures widely used for the identification and characterization of patterns in images [14–17]. The development of these emerges mainly from studies of the visual cortex found in living beings [18,19]. One

of the problems found in some neural networks (for example, fully connected feed-forward models) is the increasing number of parameters, even in shallow architectures where there are numerous neurons, making these models impractical in imaging applications [13]. Using a model based on a CNN, it is possible to develop more “deep” architectures, even with a reduced number of parameters. Specific CNNs, such as U-NET architecture developed by [20,21], were used in biomedical imaging data segmentation applications. U-NET shows promising results in quantification and segmentation tasks, such as cell detection and measurement of data geometry in medical images [22]. U-NET family architectures, such as W-UNET, which structurally has two U-NETs (encoder–decoder) [23,24], and UNET++, which is also made up of two U-NETs (different geometries and the integration of escape routes) [15], were implemented, allowing for a comparative study regarding the performance in segmentation of cellular nuclei.

In the field of medical image processing, different imaging modalities present unique challenges that require tailored segmentation approaches. U-NET architecture, known for its simplicity and effectiveness, has been widely adopted for a variety of tasks due to its robust performance across diverse image types [20]. On the other hand, W-UNET, a more recent enhancement, incorporates additional layers and parameters designed to capture more intricate features, which can be particularly beneficial in complex medical imaging scenarios [23]. While the differences in image features and segmentation challenges between these architectures are indeed significant, analyzing them together offers a comprehensive perspective on their applicability and sustainability in real-world medical imaging. By comparing U-NET and W-UNET across a range of tasks, we can better understand how each architecture adapts to varying levels of image complexity and feature diversity. This comparison not only highlights the strengths of each model in specific contexts but also provides a broader understanding of their generalizability, which is crucial for advancing segmentation techniques in the medical field. Moreover, the integration of both approaches within a single study can provide valuable insights into their complementary roles. For instance, U-NET may excel in processing simpler, high-contrast images, where speed and resource efficiency are paramount [23]. In contrast, W-UNET’s ability to delve deeper into image details makes it indispensable for more challenging cases, such as detecting subtle pathological changes in low-contrast or noisy images [21]. This dual approach allows for a more nuanced application of deep learning techniques in medical imaging, ensuring that practitioners can select the most appropriate model based on the specific needs of the task at hand. Ultimately, the comparison of U-NET and W-UNET architectures within this study underscores the importance of versatility in medical image segmentation. By demonstrating the conditions under which each model performs optimally, we contribute to the development of more reliable and accurate diagnostic tools. This not only enhances the quality of patient care but also supports the ongoing innovation in medical imaging technologies. Therefore, the integrated analysis of these architectures is not only significant but also essential for advancing the field and ensuring that segmentation methods remain adaptable to the ever-evolving challenges of medical diagnostics.

Ronneberger and co-workers [20] proposed a CNN-based U-NET architecture to be used in biomedical imaging data segmentation applications. This model is used in quantification tasks, such as cell detection and data geometry measurement in medical images [22]. The architecture is composed of two paths, one of contraction to capture the context of the data located on the left side and the symmetrical expansive path on the right side, with the purpose of locating the data with a high degree of precision (Figure 1) [20]. The strategy used in the U-NET learning process depends on the use of augmentation for effective training from a few samples accompanied by annotation [20]. U-NET, a U-shaped architecture, consists of a specific encoder–decoder scheme. The encoder reduces the spatial dimensions in each of the layers while increasing the channels. On the other hand, the decoding component increases the spatial dimensions while reducing the channels. In the end, the spatial dimensions are restored to make a prediction for each pixel present in the input image (Figure 1).

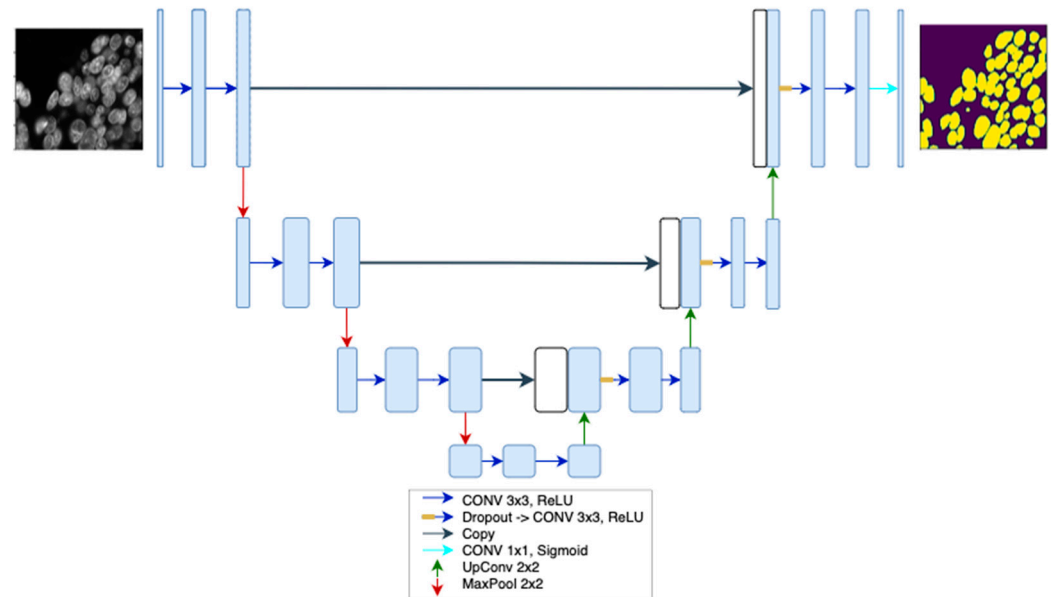


Figure 1. U-NET architecture: the feature maps on each layer are represented by the blue boxes, the arrows represent the different operations, and the white boxes represent the cropped feature maps from the contraction path.

In this study, we delve into the realm of medical imaging with a keen focus on advancing the classification and segmentation of pathological data through deep learning (DL) methodologies. By leveraging the powerful capabilities of convolutional neural networks (CNNs) and specialized architectures, such as U-NET, our research endeavors to push the boundaries of pathology diagnosis. The core of our work primarily revolves around image segmentation, complemented by thorough calibration and simulation tests. Following this preparatory phase, we meticulously conduct a series of tests, culminating in a critical and comparative analysis of the results across various datasets and implemented models. Through this comprehensive approach, we aim to underscore the transformative potential of NN-driven image processing within the medical domain, paving the way for enhanced diagnostic accuracy and treatment efficacy.

2. Materials and Methods

This section delineates the methodological procedures undertaken to investigate the classification and segmentation models of medical imaging data in this study. Deep learning (DL) methodologies, specifically convolutional neural networks (CNNs) and various U-NET family architectures, were deployed for segmentation tasks on histological image datasets. The pre-processing phase involved several critical techniques: resizing ensured uniform input dimensions across the dataset, which is essential for consistent processing by the neural networks. Normalization adjusted pixel values to a standard range, improving convergence during training by stabilizing gradients. Data augmentation, such as random rotations and inversions, was applied to artificially increase the diversity of the training data, thereby reducing the risk of overfitting and enhancing the model's ability to generalize to unseen data. During training, each model was carefully parameterized, with hyperparameters such as the learning rate, batch size, and number of epochs optimized through grid search and validation. We employed Adam and stochastic gradient descent (SGD) [20,25] as optimization algorithms, which were chosen for their effectiveness in deep learning tasks. The models were evaluated using metrics, like accuracy, the DICE coefficient, and IoU, which were selected based on their relevance to medical image segmentation, with a focus on balancing precision and recall in order to optimize segmentation accuracy. The primary focus of this study was to assess the performance of these models in accurately predicting pathology-related features.

For each segmentation task, a dataset was carefully selected to match the specific requirements of the experiment. An automated processing pipeline was then developed using Python scripts to standardize the sequence of pre-processing, model training, and evaluation. This pipeline included steps, such as data loading, augmentation, and model fitting, all of which were managed through a modular approach that allowed for easy modification and testing of different architectures and parameters. Figure 2 illustrates the detailed workflow, highlighting each step from data pre-processing to final model evaluation. In the initial phase, the data underwent pre-processing techniques, including image resizing, normalization, and random image inversions and rotations. Subsequently, the dataset was subdivided into three components and subjected to the defined processing techniques. Following this, for each of the models implemented across the tasks, the training, parameterization, and optimization processes were meticulously initiated. Finally, the trained models underwent rigorous testing, evaluation, and analysis to ascertain their efficacy and performance.

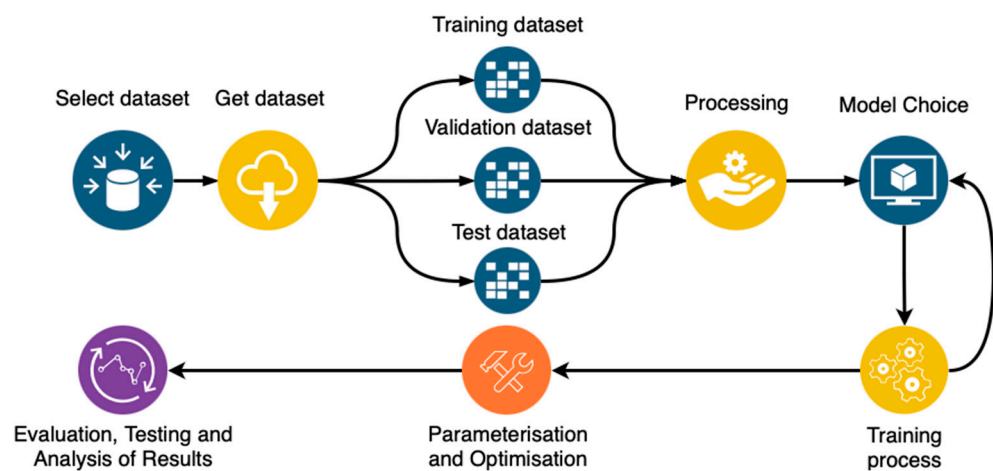


Figure 2. Workflow depicting the steps developed and executed in this study.

The dataset used in the implementation of detection and segmentation methodologies in medical images was the set of core images from the competition initiative, known as DSB2018 (or BBBC038) [26,27]. The set of images is provided with the training data, and the test was already properly separated. It consists of 670 image pairs and the respective mask/annotation that will be used in the training process of the implemented models. For testing, 65 images are provided. The training image set was divided into two data groups, one for the learning process of the models with 536 images and a second for the inherent validation process of the training process with 134 images. The dataset consists of a diverse array of histological images, including variations such as purple-stained tissues, pink and purple combinations, grayscale images, and fluorescent tissues. This diversity posed significant challenges for the segmentation models, as each type of tissue presented unique features that required the model to adapt its learned representations. We addressed these challenges by applying tailored pre-processing techniques and data augmentation strategies, which helped the models to effectively learn and generalize across different tissue types. The original size of the images vary between 256×256 and 1040×1388 . The first transformation performed on the DSB2018 dataset consisted of resizing all the images to a uniform dimension of 256×256 . The remaining processing methods applied to the dataset include normalization, inversion, and random image rotations. Simple techniques and transformations in data augmentation, such as random rotations and image inversions, have shown improvements in the results obtained, especially in datasets where feature encoding is required for tasks such as image recognition [28].

To evaluate the performance of different architectures in nuclei segmentation, we implemented a series of deep learning models, each configured with specific hyperparameters. For each model, the learning rate, batch size, and number of epochs were selected based on

preliminary experiments aimed at optimizing performance. We also implemented early stopping and cross-validation techniques to prevent overfitting and ensure robustness in our results [29]. We selected a CNN as a baseline model due to its well-established effectiveness in image classification tasks, providing a comparative benchmark. Models from the U-NET family, including the original U-NET, W-UNET, and UNET++, were chosen for their superior performance in image segmentation, particularly in medical imaging contexts. U-NET is renowned for its encoder–decoder structure that captures both local and global context, while W-UNET, with its dual U-NET configuration, is designed to further enhance segmentation accuracy by refining feature encoding and decoding. UNET++, a more recent variant, offers additional layers and dense connections, which improve performance on complex segmentation tasks (Figure 1).

In W-UNET, the first U-NET processes the input image I through its encoding and decoding paths [30]. During encoding, the image I is passed through a series of convolutional layers f_1 , ReLU activations σ , and max-pooling operations P , producing encoded feature maps $E_{1,l+1} = P(\sigma(f_1(E_{1,l})))$. In the decoding stage, these encoded features are upsampled using transposed convolutions T_1 and concatenated with corresponding features from the encoding path using skip connections \oplus . This process, followed by additional convolutions f_2 , generates the intermediate segmentation map S_1 , represented as $D_{1,l-1} = \sigma(f_2(T_1(E_{1,l}) \oplus E_{1,l}))$ and $S_1 = D_{1,0}$. The second U-NET then takes this intermediate segmentation map S_1 as input and further processes it. The encoding in this second U-NET is similar to the first, with the map S_1 being transformed into refined encoded features $E_{2,l+1} = P(\sigma(f_3(E_{2,l})))$. The decoding stage follows, where these refined features are upsampled and combined with skip connections, yielding the final segmentation map $S_2 = D_{2,0}$ after undergoing the transformation $D_{2,l-1} = \sigma(f_4(T_2(E_{2,l}) \oplus E_{2,l}))$.

In each of the architectures implemented in the study, the DICE function was used to calculate the losses and the metrics of accuracy (ACC), intersection over union (IoU), recall, DICE coefficient, and precision as measures of performance. We selected accuracy (ACC), DICE coefficient, IoU, recall, and precision as performance metrics based on their relevance to segmentation quality. Accuracy provided a general measure of performance, while the DICE coefficient and IoU offered more nuanced insights into the overlap between predicted and actual segmentation masks, which is critical in medical imaging where precision is paramount. Recall and precision were analyzed to understand the trade-offs between false positives and false negatives, with particular attention to ensuring that the models did not miss critical pathology-related features. This balanced approach to metric selection allowed us to thoroughly assess each model's performance in both typical and challenging scenarios. For the functions used to calculate the losses, the binary cross entropy and the DICE coefficient were used [31]. The use of the DICE coefficient is relevant in segmentation given that in medical imaging, the regions of interest are often small compared to the entire sample, causing the learning process to stall at the local minimum of the loss function.

U-NET architecture, along with its variants, U-NET++ and W-UNET, were developed to address core segmentation tasks in medical imaging, leveraging Python (version 3.8) and the PyTorch library [32]. These architectures are built on a modular design, with each model comprising distinct layers organized into functional blocks, namely, Downsampling and Upsampling blocks, which are critical to their performance in segmentation tasks. The Downsampling block, fundamental to the encoding phase, consists of two sequential convolutional layers followed by a max-pooling layer. The convolutional layers are responsible for extracting hierarchical features from the input images, progressively learning more complex representations. The max-pooling layer then reduces the spatial dimensions of the feature maps, effectively condensing the information and allowing the model to capture more abstract features with reduced computational complexity. Conversely, the Upsampling block is pivotal during the decoding phase. It comprises a transposed convolution layer (also known as a deconvolution operation) that performs the inverse operation of max pooling, increasing the spatial resolution of the feature maps. This step is crucial for reconstructing the image details lost during Downsampling. To enhance the precision of the

reconstruction, the output of the Upsampling block is concatenated with the corresponding Downsampling block's output from the encoding path. This skip connection ensures that the model retains high-resolution features, improving the accuracy of the segmentation boundaries. Throughout the development and evaluation process, various optimization techniques were employed to enhance model performance. Specifically, we experimented with optimization functions, such as Adam and stochastic gradient descent (SGD) [25]. Adam was selected for its adaptive learning rate capabilities, which provide efficient convergence and robust performance across different training scenarios. In contrast, SGD, with its momentum-based updates, was also tested to gauge its effectiveness in refining the model's parameters, particularly in scenarios where Adam's adaptiveness might not yield optimal results. These architectural and optimization choices were crucial in fine tuning the models for high-performance segmentation tasks, allowing us to effectively balance computational efficiency with segmentation accuracy.

To evaluate segmentation performance, we implemented the W-UNET variant, also known as W-NET, which consists of two U-NET models. The first U-NET acts as an encoder, processing the input images, while the second U-NET performs the decoding and reconstruction. W-UNET architecture includes two types of modules: Downsampling and Upsampling. The Downsampling block consists of two convolutional layers followed by a max pooling layer, while the Upsampling block features a transposed convolution layer (deconvolution) paired with a Downsampling module. This setup includes a total of four Downsampling modules and four Upsampling modules. Each convolutional layer is followed by a Rectified Linear Unit (ReLU) activation function and a normalization layer. We evaluated W-UNET's performance using metrics such as the DICE coefficient, accuracy (ACC), intersection over union (IoU), recall, and precision.

Additionally, we implemented the UNET++ variant, also known as Nested-UNET, to further investigate segmentation capabilities. UNET++ comprises two U-NET networks of different depths, with their decoders densely connected at the same resolution through escape routes. The architecture starts with four Downsampling modules, followed by ten Upsampling modules. Each Upsampling module includes a transposed convolution layer, concatenation with outputs from Downsampling modules, and an additional Downsampling unit. The network's final output is produced through four convolutional layers with sigmoid activation.

In the last phase of this work, different preparatory tests were carried out to calibrate and identify the simulation limits so that it was possible to sustain a comparative discussion of the results. After completing all tests, we conducted a comprehensive analysis of the results across different datasets and models. U-NET++ consistently outperformed other models on more complex segmentation tasks, particularly those involving diverse tissue types, due to its enhanced network architecture. W-UNET demonstrated robust performance in cases requiring precise boundary delineation, while the baseline CNN struggled with more intricate features but provided a useful benchmark for comparison. These findings contribute valuable insights into the practical application of these architectures in medical image segmentation, highlighting specific conditions where each model excels or faces challenges.

3. Results and Discussion

As previously mentioned, for the detection and segmentation of images, the dataset of histological images DSB2018 was selected. The main objective of this work was the detection of cell nuclei. For the analysis of the DSB2018 dataset, four architectures were implemented, namely, the CNN, U-NET, W-UNET, and UNET++. The implementation of each of the architectures included two phases: the learning phase, that is, the training and validation phase, and the test and inference phase. Images were pre-processed so that they could be corrected or highlighted properly, improving their contrast, correcting defective pixels, or reducing noise. This treatment applied to the images that make up the dataset is important since the direct visualization of digital medical images is generally not adequate.

Figure 3 shows the graphs obtained from the different metrics applied, depending on the number of iterations (Epochs) during the learning phase of W-UNET architecture. Directing attention to the primary metrics of interest in this study, namely, the DICE coefficient and the intersection over union (IoU) (Figure 3C and Figure 3D, respectively), the trend indicates a progressive increase in both metrics throughout the iterations.

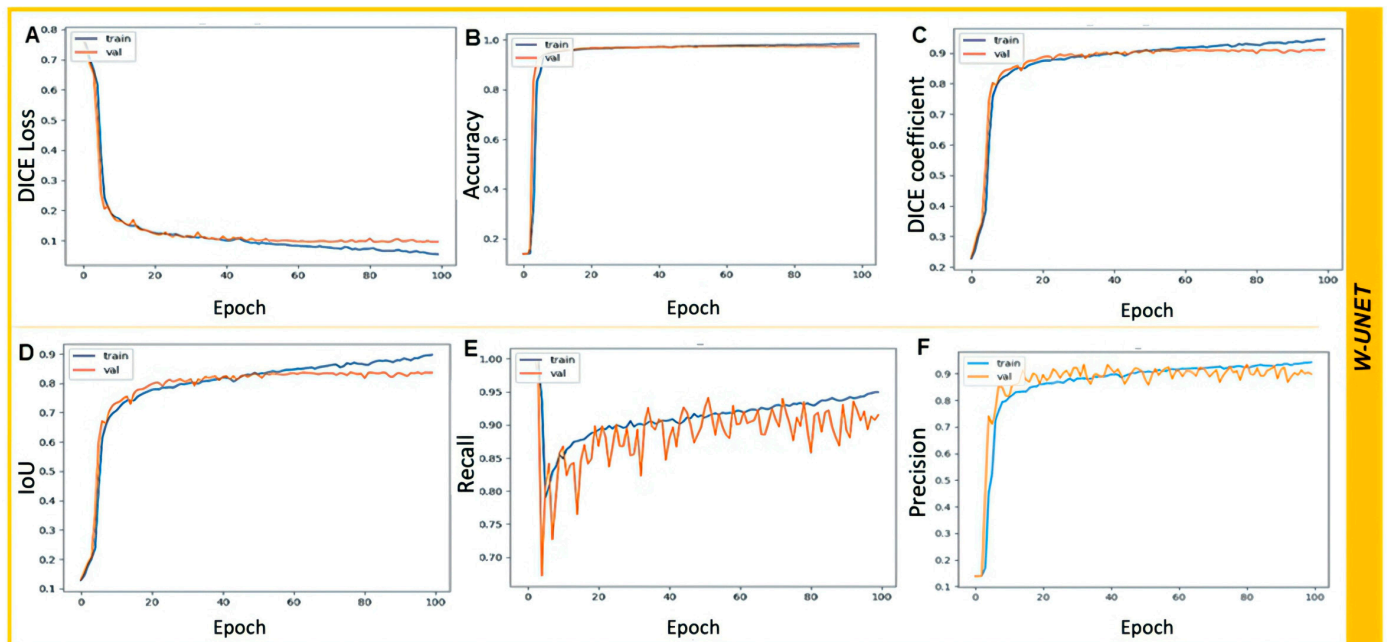


Figure 3. Graphs of the metrics obtained during the learning phase in image detection and segmentation for W-UNET architectures. (A) Graph of loss values vs. number of iterations (Epochs); (B) Graph of accuracy values vs. number of iterations (Epochs); (C) Graph of DICE coefficient values vs. number of iterations (Epochs); (D) Graph of IoU values vs. number of iterations (Epochs); (E) Graph of recall values vs. number of iterations (Epochs); (F) Graph of precision values vs. number of iterations (Epochs) obtained with the W-UNET and UNET++ architectures, respectively.

After completing the training process, the model was tested with the images intended for the training and validation process, and more importantly, with the unknown images, that is, those intended only for the test. The choice of images to be segmented was random. Some of these results, by way of demonstration, are shown in Figure 4, where the results obtained with training images and images only to be used in the test phase with W-UNET architecture can be analyzed.

In the tests with the training and validation data, four pieces of information are presented: the original image of the dataset, the respective mask, the prediction made with the implemented model, and, also, an overlapping figure of the previous three images. This overlap is important for detecting the veracity of the results. For the test process with unknown images (test images), the original image used and the image resulting from the prediction made by the model are presented. Focusing on the image of the overlay, it is possible through the variation of colors to interpret the veracity of the results obtained in the segmentation. The green, purple, red, and white colors represent true positives, false positives, false negatives, and true negatives, respectively.

Table 1 shows the values of the metrics resulting from the test phase, with each of the implemented architectures. To complement the study of core segmentation, metrics related to training and validation processes are also presented. Additionally, Table 1 presents some metrics reported for this dataset in the literature with different architectures, as well as the results obtained [30,33].

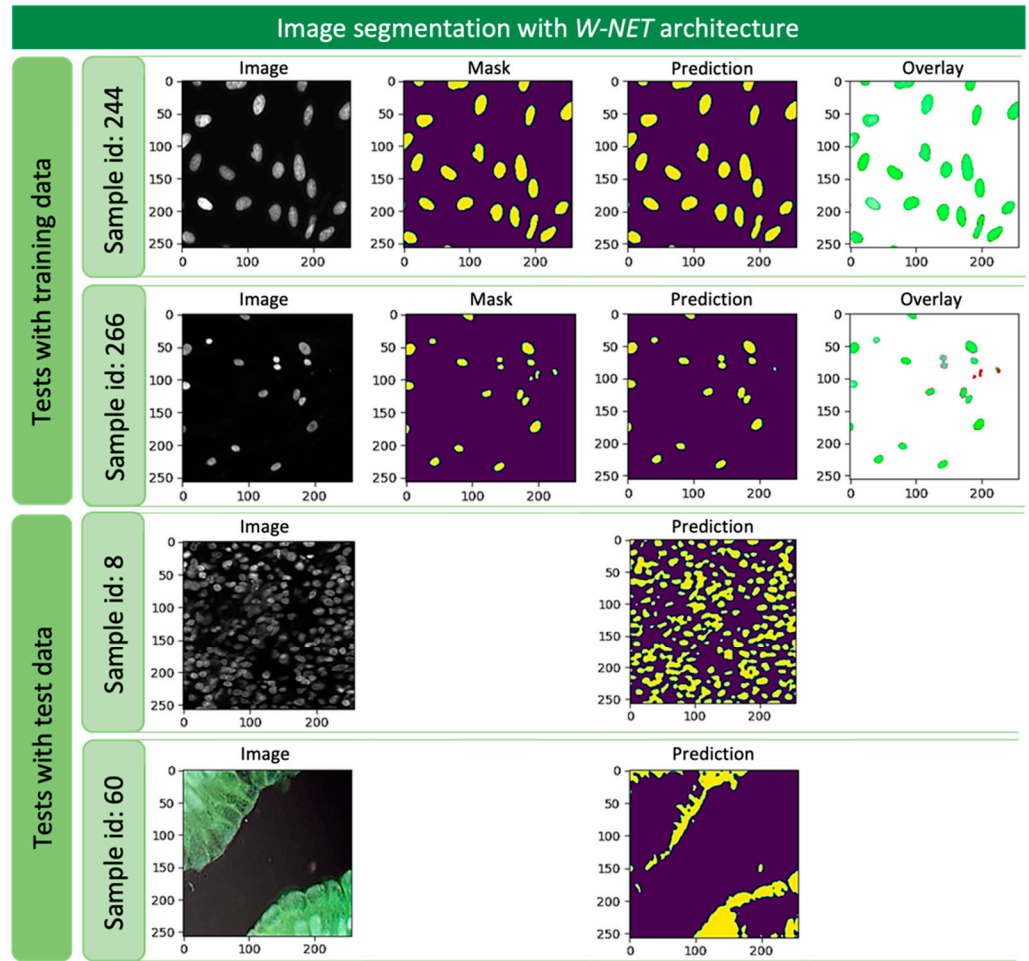


Figure 4. Images resulting from segmentation tests obtained with W-UNET architecture for the DSB2018 dataset.

Table 1. The results obtained for the image segmentation process using the DSB2018 dataset for the four developed architectures and the results reported in the literature.

Phase	Loss DICE	Accuracy	Coeff DICE	IoU	Recall	Precision	Model
Training	0.1553	0.9566	0.8449	0.7324	0.8433	0.8454	CNN
Validation	0.1531	0.9587	0.8459	0.7339	0.8640	0.8399	
Test	0.0161	1.000	0.9835	0.9675	1.0000	1.0000	
Training	0.1670	0.9479	0.8326	0.7138	0.9446	0.7457	UNET
Validation	0.1662	0.9459	0.845	0.7335	0.9411	0.7469	
Test	0.0084	1.0000	0.9916	0.9832	1.0000	1.0000	
Training	0.0542	0.9850	0.9459	0.8974	0.9564	0.9369	W_UNET
Validation	0.0960	0.9738	0.9104	0.8363	0.9151	0.8970	
Test	0.0079	1.0000	0.9921	0.9843	1.0000	1.0000	
Training	0.0907	0.9744	0.9094	0.8341	0.9249	0.8954	UNET++
Validation	0.1116	0.9689	0.8918	0.8052	0.9148	0.8647	
Test	0.0159	1.0000	0.9848	0.9701	1.0000	1.0000	
N.A.	-	-	-	0.9077	-	-	U-NET [27]
	-	-	-	0.9263	-	-	UNET++ [27]
	-	-	-	0.9092	-	-	Wide UNET [27]
	-	-	0.9215	-	-	-	R2U-NET [28]

N.A., Not Applicable.

Comparing the results obtained in the test phase with those in the learning phase, they improve considerably in all metrics for all implemented architectures. Despite the implementation of accuracy, precision, and recall metrics, only the DICE coefficient and IoU metrics will be considered for this study, as they are the most appropriate metrics for segmentation studies. Thus, analyzing the most appropriate metrics for this study, the DICE coefficient and the IoU, it is evident that the best results are obtained with U-NET and W-UNET architectures, followed by UNET++ architecture and, finally, the CNN.

In addition, it is important to highlight that the values obtained for the IoU (0.9675–0.9843) are higher than those reported in [27] (0.9077–0.9263). The same is verified when comparing the DICE coefficient values (0.9835–0.9916) with one reported value (0.9215) [28], thus reinforcing that the results obtained are quite promising. It is noteworthy that despite yielding inferior results, the traditional CNN architecture demonstrates faster processing and training times compared to other architectures, given an equal number of iterations.

The proposed method addresses some of the common limitations of general U-Net architectures, particularly in terms of imprecise segmentation and dependence on data quality, while also introducing its own considerations. To mitigate the models' dependency on data quality and improve segmentation outcomes, significant pre-processing was applied to the images in the DSB2018 dataset, as mentioned earlier. However, the method still faces limitations. Despite the enhancements provided by W-UNET and UNET++, challenges with segmentation precision may persist, particularly in complex or noisy images. While W-UNET's dual structure helps refine features, it may not fully resolve imprecision, especially when segmenting small or irregular structures. Although it shows promise for the DSB2018 histological image dataset, its effectiveness may vary when applied to different types of medical images or datasets with different characteristics, such as varying resolutions, contrasts, or noise types. In summary, while the proposed method introduces advanced techniques to address some limitations of general U-NET architectures, it is not without challenges. Precision in segmentation and dependency on high-quality data remain concerns, and further investigation may be required to ensure its applicability across different datasets.

4. Conclusions

The results obtained demonstrate the importance of artificial intelligence techniques in supporting medical diagnosis, allowing the process of medical imaging analysis to become an autonomous and automatic task and allowing the medical community to achieve greater speed and efficiency in diagnosis. In greater detail, pre-processing techniques including resizing, normalization, and data augmentation can have a significant impact on classification and thus overall performance when using neural networks. It can be concluded that the methodologies developed for segmentation demonstrate great potential, having obtained superior results when compared to those reported in the literature that were used for the performance study. For image segmentation with the DSB2018 dataset, it is important to note that the values obtained for the IoU (0.9675–0.9843) are better than those reported in the literature (0.9077–0.9263) by Shorten and Khoshgoftaar [27], as well as the values of the DICE coefficient obtained (0.9835–0.9916) relative to the reported value (0.9215) by Durkee et al. [28].

This study's findings suggest promising avenues for future research aimed at refining and extending the work undertaken. To augment the current scope, it is proposed to delve deeper into the identification and refinement of specific cellular structures, accompanied by the utilization of additional evaluative metrics to gauge performance more comprehensively. Furthermore, within the segmentation framework, there is a call to integrate the delineation of various stages of mitosis, such as the prophase, metaphase, anaphase, and telophase. This holistic approach promises to provide invaluable insights into the broader applicability and effectiveness of the methodologies employed, thereby advancing the field of medical image analysis. Additionally, the development of mechanisms for the automatic

selection of models and processing parameters in complex datasets would enhance the applicability and efficiency of DL methodologies in medical imaging. Future work should focus on identifying specific cellular structures, evaluating segmentation performance, and quantifying cores in samples, enhancing the applicability and performance of DL methodologies in medical imaging.

Author Contributions: Conceptualization, S.C. and R.S.; methodology, S.C. and R.S.; software, S.C.; validation, S.C. and R.S.; formal analysis, S.C.; investigation, S.C. and R.S.; resources, S.C.; writing—original draft preparation, S.C. and R.S.; writing—review and editing, S.C., R.S. and V.P.; supervision, R.S.; funding acquisition, R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by Fundação para a Ciência e a Tecnologia, Portugal, Project Reference UIDB/04005/2020.

Data Availability Statement: The original data presented in the study are openly available at <https://bbbc.broadinstitute.org/bbbc/BBBC038>, accessed on 10 January 2023.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Metter, R.L.V.; Beutel, J.; Kundel, H.L. *Handbook of Medical Imaging*; John Wiley & Sons: Hoboken, NJ, USA, 2000; Volume 1, ISBN 978-0-8194-7772-9.
2. Krupinski, E.A. The Importance of Perception Research in Medical Imaging. *Radiat. Med.* **2000**, *18*, 329–334. [[PubMed](#)]
3. Bergmeir, C.; García Silvente, M.; Benítez, J.M. Segmentation of Cervical Cell Nuclei in High-Resolution Microscopic Images: A New Algorithm and a Web-Based Software Framework. *Comput. Methods Programs Biomed.* **2012**, *107*, 497–512. [[CrossRef](#)]
4. Rguibi, Z.; Hajami, A.; Zitouni, D.; Elqaraoui, A.; Bedraoui, A. CXAI: Explaining Convolutional Neural Networks for Medical Imaging Diagnostic. *Electronics* **2022**, *11*, 1775. [[CrossRef](#)]
5. Galić, I.; Habijan, M.; Leventić, H.; Romić, K. Machine Learning Empowering Personalized Medicine: A Comprehensive Review of Medical Image Analysis Methods. *Electronics* **2023**, *12*, 4411. [[CrossRef](#)]
6. Wan, T.; Zhao, L.; Feng, H.; Li, D.; Tong, C.; Qin, Z. Robust Nuclei Segmentation in Histopathology Using ASPPU-Net and Boundary Refinement. *Neurocomputing* **2020**, *408*, 144–156. [[CrossRef](#)]
7. Jia, D.; Zhang, C.; Wu, N.; Guo, Z.; Ge, H. Multi-Layer Segmentation Framework for Cell Nuclei Using Improved GVF Snake Model, Watershed, and Ellipse Fitting. *Biomed. Signal Process. Control* **2021**, *67*, 102516. [[CrossRef](#)]
8. Aswath, A.; Alsahaf, A.; Giepmans, B.N.G.; Azzopardi, G. Segmentation in Large-Scale Cellular Electron Microscopy with Deep Learning: A Literature Survey. *Med. Image Anal.* **2023**, *89*, 102920. [[CrossRef](#)]
9. Xu, Z.; Lim, S.; Lu, Y.; Jung, S.-W. Reversed Domain Adaptation for Nuclei Segmentation-Based Pathological Image Classification. *Comput. Biol. Med.* **2024**, *168*, 107726. [[CrossRef](#)]
10. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A Survey on Deep Learning in Medical Image Analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [[CrossRef](#)]
11. Chen, Y.; Yin, M.; Li, Y.; Cai, Q. CSU-Net: A CNN-Transformer Parallel Network for Multimodal Brain Tumour Segmentation. *Electronics* **2022**, *11*, 2226. [[CrossRef](#)]
12. Park, Y.; Park, J.; Jang, G.-J. Efficient Perineural Invasion Detection of Histopathological Images Using U-Net. *Electronics* **2022**, *11*, 1649. [[CrossRef](#)]
13. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016; ISBN 978-0-262-03561-3.
14. Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Mohamed, N.A.; Arshad, H. State-of-the-Art in Artificial Neural Network Applications: A Survey. *Heliyon* **2018**, *4*, e00938. [[CrossRef](#)] [[PubMed](#)]
15. Kiran, I.; Raza, B.; Ijaz, A.; Khan, M.A. DenseRes-Unet: Segmentation of Overlapped/Clustered Nuclei from Multi Organ Histopathology Images. *Comput. Biol. Med.* **2022**, *143*, 105267. [[CrossRef](#)]
16. Wang, J.; Zhang, Z.; Wu, M.; Ye, Y.; Wang, S.; Cao, Y.; Yang, H. Nuclei Instance Segmentation Using a Transformer-Based Graph Convolutional Network and Contextual Information Augmentation. *Comput. Biol. Med.* **2023**, *167*, 107622. [[CrossRef](#)]
17. Zhao, B.; Chen, X.; Li, Z.; Yu, Z.; Yao, S.; Yan, L.; Wang, Y.; Liu, Z.; Liang, C.; Han, C. Triple U-Net: Hematoxylin-Aware Nuclei Segmentation with Progressive Dense Feature Aggregation. *Med. Image Anal.* **2020**, *65*, 101786. [[CrossRef](#)] [[PubMed](#)]
18. Hubel, D.H.; Wiesel, T.N. Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex. *J. Physiol.* **1962**, *160*, 106–154. [[CrossRef](#)] [[PubMed](#)]
19. Poggio, T.; Serre, T. Models of Visual Cortex. *Scholarpedia* **2013**, *8*, 3516. [[CrossRef](#)]
20. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lect. Notes Comput. Sci.* **2015**, *9351*, 234–241. [[CrossRef](#)]
21. Kanavos, A.; Papadimitriou, O.; Al-Hussaini, K.; Maragoudakis, M.; Karamitsos, I. Advanced Convolutional Neural Networks for Precise White Blood Cell Subtype Classification in Medical Diagnostics. *Electronics* **2024**, *13*, 2818. [[CrossRef](#)]

22. Falk, T.; Mai, D.; Bensch, R.; Çiçek, Ö.; Abdulkadir, A.; Marrakchi, Y.; Böhm, A.; Deubner, J.; Jäckel, Z.; Seiwald, K.; et al. U-Net: Deep Learning for Cell Counting, Detection, and Morphometry. *Nat. Methods* **2018**, *16*, 67–70. [[CrossRef](#)]
23. Xia, X.; Kulis, B. W-Net: A Deep Model for Fully Unsupervised Image Segmentation. *arXiv* **2017**, arXiv:1711.08506.
24. Waqas, N.; Safie, S.I.; Kadir, K.A.; Khan, S. Knee Cartilage Segmentation Using Improved U-Net. *Int. J. Adv. Comput. Sci. Appl.* **2023**, *14*. [[CrossRef](#)]
25. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
26. Kaggle 2018 Data Science Bowl | Broad Bioimage Benchmark Collection. Available online: <https://bbbc.broadinstitute.org/bbbc/BBBC038> (accessed on 10 January 2023).
27. Caicedo, J.C.; Goodman, A.; Karhohs, K.W.; Cimini, B.A.; Ackerman, J.; Haghghi, M.; Heng, C.K.; Becker, T.; Doan, M.; McQuin, C.; et al. Nucleus Segmentation across Imaging Experiments: The 2018 Data Science Bowl. *Nat. Methods* **2019**, *16*, 1247–1253. [[CrossRef](#)] [[PubMed](#)]
28. Shorten, C.; Khoshgoftaar, T.M. A Survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*, 60. [[CrossRef](#)]
29. Durkee, M.S.; Abraham, R.; Clark, M.R.; Giger, M.L. Artificial Intelligence and Cellular Segmentation in Tissue Microscopy Images. *Am. J. Pathol.* **2021**, *191*, 1693–1701. [[CrossRef](#)]
30. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*; Springer International Publishing: Berlin/Heidelberg, Germany, 2018; Volume 11045, pp. 3–11. [[CrossRef](#)]
31. Jadon, S. A Survey of Loss Functions for Semantic Segmentation. In Proceedings of the 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Via del Mar, Chile, 27–29 October 2020; pp. 1–7.
32. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Proceedings of the 33rd Annual Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019. Available online: https://proceedings.neurips.cc/paper_files/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf (accessed on 10 January 2023).
33. Alom, M.Z.; Yakopcic, C.; Taha, T.M.; Asari, V.K. Nuclei Segmentation with Recurrent Residual Convolutional Neural Networks Based U-Net (R2U-Net). In Proceedings of the NAECON 2018-IEEE National Aerospace and Electronics Conference, Dayton, OH, USA, 23–26 July 2018; pp. 228–233. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.