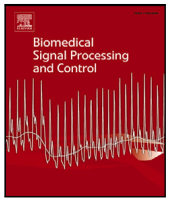
Biomedical Signal Processing and Control 83 (2023) 104720

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



An imbalance-aware nuclei segmentation methodology for H&E stained histopathology images

Emrah Hancer [a,*], Mohamed Traoré [b], Refik Samet [b], Zeynep Yıldırım [b], Nooshin Nemati [b]

[a] Department of Software Engineering, Mehmet Akif Ersoy University, 15030, Burdur, Turkey
[b] Department of Computer Engineering, Ankara University, 06100, Ankara, Turkey

ARTICLE INFO

Keywords:
Nuclei segmentation
Computational pathology
Semantic segmentation
Generalized Dice loss

ABSTRACT

A key step in computational pathology is to automate the laborious process of manual nuclei segmentation in Hematoxylin and Eosin (H&E) stained whole slide images (WSIs). Despite lots of efforts put forward by the researchers to develop automated nuclei segmentation methodologies in the literature, the segmentation performance is still constrained due to several challenges, including overlapping and clumped nuclei, scanners with different resolutions and nuclei with varying sizes and shapes. In this paper, we introduce an imbalance-aware nuclei segmentation methodology to deal with class imbalance problems in H&E stained histopathology images. The introduced methodology involves the following improvements: (1) the design of a preprocessing stage with a variety of resize-split, augmentation and normalization techniques, and (2) an enhanced lightweight U-Net architecture with a generalized Dice loss layer. To prove its effectiveness and efficiency, a comprehensive experimental study is carried out on a well-known benchmark, namely the MonuSeg2018 dataset. According to the results, the proposed methodology outperforms various recently introduced studies in terms of well-known evaluation metrics, such as Aggregated Jaccard Index (AJI) and Intersection of Union (IoU).

1. Introduction

Nuclei segmentation is of importance from a biological perspective. DNA content, chromatin condensation and nucleus morphology are among the valuable information provided by this method [1]. Studies of cell cycles or mutations in cancer-related proteins can be carried out using this information. However, several factors contribute to the difficulty of nuclei segmentation, such as cell overlap, image noise, and non-uniform conditions for the acquisition and preparation of images.

Manual nuclei segmentation requires clinical experience, is time-consuming, and is prone to human error [2]. Consequently, automatic image analysis methods have been developed to reduce the workload of humans and overcome subjective interpretation. There are a lot of conventional methods for detecting nuclei and segmenting them. The most common way is to use thresholding methods in conjunction with watershed algorithms. However, conventional methods are complicated by a variety of factors. For instance, they require manual parameter tuning, become highly specific for certain types of images, and degrade in performance with noise. In recent years, deep learning has become a state-of-the-art method, as it enhances performance in many medical applications. As compared with conventional methods for segmenting nuclei, deep learning methods have a higher generalization capacity, can provide more meaningful features from images and are more robust to noise. As a result, various methods for nuclei segmentation based on deep learning have been proposed over the years. However, the task of nuclei segmentation still remains an active research field [3].

The study of deep learning is advancing rapidly, resulting in the development of new architectures at an accelerated rate. Furthermore, after recognizing the importance of nuclei segmentation, several methods have been introduced to deal with this issue, most of which rely on a simple architecture: U-Net [4]. The U-Net architecture is the most commonly used for segmenting medical images. As part of the Cell Tracking Challenge in 2015, this architecture was developed specifically for biomedical image segmentation. U-Net has similar characteristics with fully convolutional networks (FCN) [5]; however, it has a symmetric architecture with more up-sampling layers. In particular, U-Net gets its name from the U-shaped architecture it has in its design.

1.1. Literature review

The U-Net architecture has received a lot of attention for the task of nuclei segmentation in the literature due to its diverse structure and promising results. In other words, U-Net has a flexible architecture that

* Corresponding author.
E-mail address: ehancer@mehmetakif.edu.tr (E. Hancer).

https://doi.org/10.1016/j.bspc.2023.104720
Received 21 September 2022; Received in revised form 31 January 2023; Accepted 14 February 2023
Available online 20 February 2023

allows for various modification, resulting successful performance in a variety of biomedical applications. Due to the motivation of our study, we will focus on especially the studies of U-Net proposed for nuclei segmentation from histopathology images in this section.

Zhou et al. [6] retrofitted the standard U-Net architecture with a series of nested and dense skip connections for robust biomedical image segmentation. So-called U-Net++, the resulting architecture aimed to maximize the similarity between the feature maps of the encoder and decoder subnetworks by semantically enriching the feature maps of the encoder before merging them with the corresponding ones from the decoder, thus facilitating the learning process in the architecture. Although the proposed architecture outperformed the standard U-Net in biomedical image segmentation, especially in nuclei segmentation, it was observed to have a high computational cost. Li et al. [7] used distance mapping in their study to increase the success rate of overlapping nuclei segmentation. The architecture proposed in the study uses a dual branch decoder: a classification branch for boundary segmentation and a regression branch for distance mapping. In the final, the outputs of these two decoder branches are merged through convolutional fusion layers to produce segmentation masks. As in Li et al. [7], Mahbod et al. [8] used U-Net for both classification and distance mapping. Differently from Li et al. [7], they used separate U-Nets for both classification and distance mapping and applied the Gaussian smoothing and watershed algorithms for the last fusion step. In terms of the F1-Score obtained on the MoNuSeg2018 test subset, the proposed architecture outperformed Li et al. [7]. Kong et al. [9] applied two connected stacked U-Net models, where cross-entropy and focal loss metrics were applied to address overlapping nuclei. Furthermore, four parallel backbones and an attention mechanism were used in their models. Zhang et al. [10] proposed a low-cost U-Net (LCU-Net) architecture for environmental microorganism image segmentation based on the U-Net and Inception architectures. Additionally, they applied a dense conditional random field (CRF) for post-processing. Vahadane [11] introduced an attention U-Net architecture which uses a dual encoder architecture to encode attention prior information and an attention skip module to extract high-quality features. However, the AJI score obtained by this architecture was very low for the MonuSeg2018 dataset. Kiran et al. [12] introduced a modified U-Net architecture based on dense and atrous blocks. While atrous blocks were used to reduce the semantic gap between the encoder and the decoder, dense blocks were used to increase the weight of the information in the last layers at the encoder part. In another study, Yildirim et al. [2] investigated the effect of normalization on nuclei segmentation through the U-Net variants. More information concerning deep nuclei segmentation can be found in [13].

1.2. Contributions

As described in the previous section, a variety of U-Net-based architectures have been proposed to deal with nuclei segmentation in Hematoxylin and Eosin (H&E) stained histopathology images. It is evident from the aforementioned studies that U-Net is capable of segmenting neuronal structures successfully. It uses an encoder–decoder architecture with skip connections that allow information to be sent toward the decoder; furthermore, it works well with the limited training datasets. However, the U-Net-based studies have not come to an end since the segmentation performance is still constrained due to the several challenges, including overlapping and clumped nuclei, scanners with different resolutions and nuclei with varying sizes and shapes.

This study develops a deep nuclei segmentation methodology for H&E stained histopathology images. The main contributions of the methodology are as follows: (1) the design of a preprocessing stage, involving resize-split, stain normalization and data augmentation techniques, (2) the integration of a generalized Dice loss layer in the U-Net architecture, and (3) a simple redesign of the U-Net architecture. Through the newly designed preprocessing stage, it is aimed to enhance the model training, yielding a generalized model for nuclei segmentation. Furthermore, through the generalized Dice loss layer, the methodology can avoid class imbalance problems between foreground and background during the model training without assigning loss weights to patterns of different classes. And lastly, the training process can be completed within a reasonable amount of time due to the simplified U-Net architecture. In order to verify the effectiveness and efficiency of the developed methodology, a comprehensive experimental study is performed on the MonuSeg2018 dataset, known as the most popular benchmark for nuclei segmentation task from the H&E stained images. According to a number of experiments, the developed methodology provides much more better segmentation results than a variety of recent studies in the literature from the perspective of different criteria.

The remainder of the paper is as follows. Section 2 provides a general knowledge concerning the dataset and briefly describes the methods used in the methodology. Section 3 defines the experiment design and presents the experimental results with discussions. Section 4 concludes the study with future trends.

2. Materials and methods

2.1. Dataset description

In this study, we use the MoNuSeg2018 dataset. Originally released by Kumar et al. [14], it was used as the benchmark dataset for the Multi-Organ Nucleus Segmentation Challenge (MoNuSeg2018). In this dataset, whole slide images (WSIs) were collected from patients from 18 different hospitals using crowd-sourcing techniques to ensure heterogeneity and diversity in the nuclei appearance. A further source of diversity to the set is that the WSIs are representative of seven different organs, including the breast, kidney, liver, prostate, bladder, colon and stomach. MoNuSeg2018 provided the boundary coordinates of annotated nuclei in XML format with a MATLAB script for binary mask generation. Additionally, an Aggregated Jaccard Index (AJI) metric source code was provided for better evaluation of segmentation performance on the dataset [15].

The dataset consists of two subsets, a training set and a test set. The training set includes 30 H&E stained histopathology images of 1000×1000 pixels extracted from separate WSIs (each scanned at $40\times$ magnification). This subset contains 21,623 manually annotated nuclei, all of which were checked by an expert pathologist with years of experience in tissue analysis. Collected from the same sources and annotated with the same protocols, the testing set comprises 14 images of 1000×1000 pixels per WSIs with about 7223 hand-annotated nuclei. Table 1 shows the details of the composition of the training and testing subsets of the MoNuSeg2018 dataset.

2.2. Proposed methodology

This section explains the overall flow of the proposed methodology for robust nuclei segmentation in histopathology images. As summarized in Fig. 1, the proposed methodology starts with a data preprocessing stage through which the collected images are first stain-normalized using Macenko's stain normalization technique [16] to reduce artifacts and balance stain colors in the dataset. The stain-normalized images are then split. To achieve this, the images are resized appropriately to avoid any loss of nuclei pixels during splitting. Afterward, to prevent overfitting during model training [17], the MoNuSeg2018 training subset is artificially augmented through an offline data augmentation process. Finally, for nuclei segmentation, the enhanced U-Net model is trained on the generated training set and then evaluated on the initial MoNuSeg2018 testing set in terms of different performance assessment metrics. The following subsections describe the methods used in the stages of the proposed methodology.

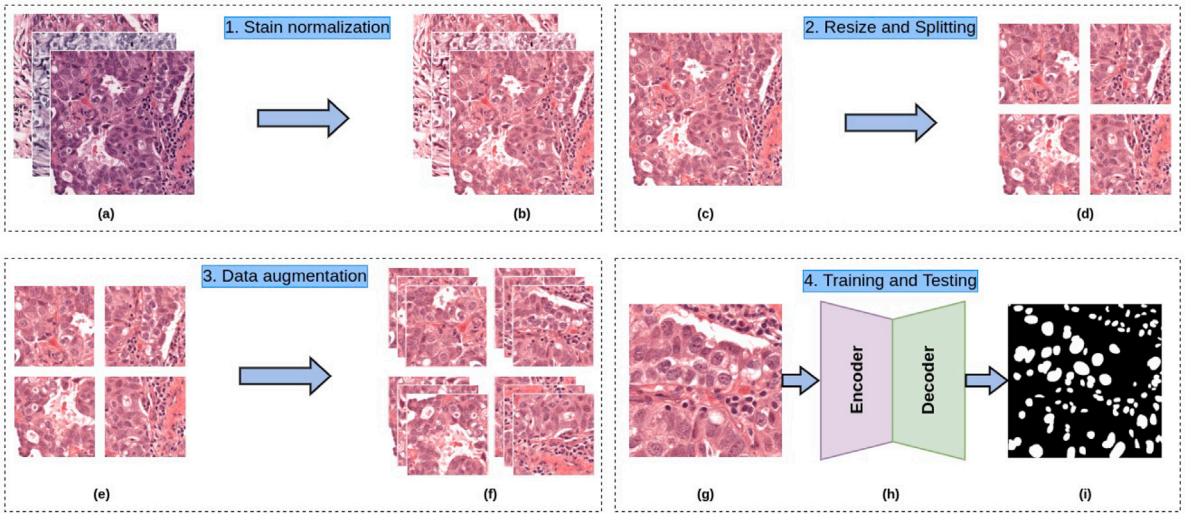


Fig. 1. The flowchart of the proposed methodology for nuclei segmentation. 1. Stain normalization: (a) original images, (b) stain normalized images; 2. Resize and splitting: (c) resized image, (d) split images; 3. Data augmentation: (e) split images, (f) augmented images; 4. Training and testing: (g) preprocessed images, (h) introduced enhanced U-Net architecture, (i) predicted segmentation masks.

Table 1
Composition of MoNuSeg training and testing data subsets.

Data subsets	Images									Total
	Breast	Liver	Kidney	Prostate	Bladder	Colon	Stomach	Lung	Brain	
Training set	6	6	6	6	2	2	2	–	–	30
Test set	2	–	3	2	2	1	–	2	2	14

2.2.1. Data preprocessing

Data preprocessing is an important stage for deep learning architectures since it ensures the diversity, relevancy, consistency, and quality of the dataset in order to build a robust model. The steps of the designed data preprocessing stage are explained as follows.

(a) Stain normalization. Tissue staining is the cornerstone of histopathology. H&E staining is one of the most popular and principal histological staining techniques used by pathologists due to its low cost and ability to provide good insight into tissues and their structures [18]. There are several types of H&E stains and they can give different results. Staining results can also vary significantly between manufacturers depending on the quality of the solutions, the tissue type, the protocols of the prior steps, and even the quality of the water (salinity and acidity) used in the laboratory. Furthermore, tissue images from different sources can also show considerable variations in nuclei appearance and color intensity, which remains one of the significant challenges in tissue segmentation. To address this challenge, we choose Macenko's technique which has been proven effective in many tissue segmentation tasks [2].

The Macenko normalization technique relies on automatically generating correct stain vectors for corresponding H&E stained images in RGB colorspace and then performing the color deconvolution. To achieve this, the input RGB image is first projected into the optical density (OD) space in which the stains are linearly separable. OD, also referred to as absorbance, measures the ability of the stains to absorb light. As shown in Eq. (1), here the OD values are obtained by taking the logarithm to the base 10 of the corresponding images normalized in the range $[0, 1]$. To have more robust results a threshold parameter (β) is employed to threshold pixels with low OD values (almost no stain). Next, the single value decomposition (SVD) is computed on the OD intensities to obtain the SVD directions from which a plane is created. The OD-converted pixels are projected onto this plane and then normalized to unit length. This projection aims to find endpoints that correspond to the stain vectors. The angle of these points are computed with reference to the initial SVD direction, then the αth and $(100-\alpha)$th percentiles are taken to obtain the robust extremes of the angles. As a final step, these extremes are converted back into the OD space to get more adequate stain vectors.

$$OD = -\log_{10}(I) \tag{1}$$

where I is a RGB image with each pixel normalized to $[0, 1]$ range.

Like many other stain normalization techniques [19,20], this technique also requires a target image. In this study, we select the target image randomly from the MoNuSeg2018 training subset. Both the training and testing subsets are stain-normalized, and the hyperparameters (α and β) of the normalizer are respectively set to 1 and 0.5 as suggested in [16]. An illustrative example of the Macenko normalization can be seen in Fig. 1.a.

(b) Image patching. Histopathological image datasets typically provide high-resolution image patches, such as MoNuSeg2018 with 1000 × 1000 pixels. Directly dealing with images at such a resolution requires very deep convolutional neural networks for accurate segmentation. However, deeper convolutional neural networks are computationally demanding [21,22]. Considering this, as in [23,24], researchers prefer to develop architectures with adequate input sizes and then resize or directly create patches from dataset images accordingly. By the same token, to get appropriately sized patches without any nuclei pixel loss, we resize the dataset images to 1200 × 1200 pixels and then split them into 300 × 300 pixel patches. Unlike Shyam et al. [23], who zero-padded images for resizing, we resize images using the nearest neighbor interpolation (NNI) image resizing technique. In the NNI technique, to increase the size of the image, empty spaces are inserted at appropriate points in the source image, and then these spaces are filled with the nearest neighboring pixels while ignoring other neighboring

pixels [25]. In comparison with zero-padding image resizing, NNI does not lead the model to learn additional black pixels, which can create poor nuclei segmentation because they are different from the actual background.

(c) Data augmentation. Data augmentation is one of the most effective regularization techniques in deep learning [26]. It uses artificial data generation based on existing data to increase the size, quality and diversity of a training set [27]. It is crucial for tasks where limited data are available, like histology, which is tedious and time-consuming to prepare slides and annotate nuclei. Data augmentation has become a sine qua non for training deep learning models, especially in computer vision tasks [28]. It reduces the risk of overfitting [17,29,30] and increases the performance and ability of the deep learning model to generalize well [31–34]. Data augmentation also makes it possible to insert the property of equivariance and rotational invariance into CNN-based models [35]. Although many network architectures were proposed to incorporate potential equivariances and invariances into CNN-Based models [35–37], the most efficient way to achieve this without requiring any hardcoding is to use data augmentation [38,39]. Moreover, traditional data augmentation techniques have been demonstrated to be robust in semantic segmentation tasks (e.g. random rotation, random horizontal and vertical flipping, random translation, etc.), see [40,41].

In this work, we employ different data augmentation techniques to increase the MoNuSeg2018 training subset: (1) random rotation in the range of $[0, 360]$ degrees, (2) random reflection on the X-axis, and (3) random translation on the X-axis. This practice allows us to increase the diversity of the MoNuSeg2018 training subset and then improve the generalization ability of the proposed enhanced U-Net architecture. An illustrative example of the augmentation process is presented in Fig. 1.c.

2.2.2. Deep semantic segmentation

The task of locating and categorizing objects in medical images is often considered as a semantic segmentation task [36,42]. Semantic segmentation aims to divide an image into semantically distinct subregions. This task is formulated in many deep learning-based methodologies as a per-pixel classification task [43] due to assigning an object class label to each pixel contained in the image [44–47]. The state-of-the-art (SOTA) deep learning networks for semantic segmentation typically rely on encoder–decoder architecture [6]. Some examples of such SOTA networks are Fully Convolutional Network (FCN) [48], U-Net [4] and Pyramid Scene Parsing Network (PSPNet) [49]. In this work, we adopt U-Net architecture for nuclei segmentation on histology images due to its effectiveness and robustness in biomedical image segmentation [50] and its ability to yield more accurate models on small datasets with some data augmentation [51].

The non-uniform distribution of object classes in a dataset is referred to as class imbalance [52]. Class imbalance is a common issue encountered in semantic segmentation tasks, especially nuclei segmentation [53], where background class may significantly outweigh foreground class depending on the density of nuclei in image patches, or vice versa [54]. Moreover, training deep learning models on unbalanced datasets usually results in poor segmentation since classes with small samples are ignored. Although special architectures have been developed to deal with the class imbalance problem in semantic segmentation tasks, the majority of solutions rely on the use of appropriate loss functions as in [36,52]. Some of these functions are Dice Loss [55], Generalized Dice Loss [56], Focal Loss [57] and Dual Focal Loss [58]. In this study, we integrate a Generalized Dice Loss layer into the proposed U-Net architecture to mitigate the detrimental impact of class imbalance on the segmentation performance.

As the last step of the proposed methodology, we post-processed the output segmentation masks using mathematical morphology to refine the result masks of the proposed enhanced U-Net architecture.

(a) U-Net architecture. U-Net is an encoder–decoder-style fully convolutional neural network designed and implemented by Ronneberger et al. [4] for biomedical image segmentation. The standard U-Net architecture basically consists of two symmetric paths, a contracting path and an expansive path. In the contracting path, feature maps are extracted and downsampled to low-dimensional feature space through a sequence of encoder units. Each of the encoder units consists of 2 consecutive 3×3 convolution layers, each of which is activated with a ReLU non-linearity, and a 2×2 max pooling layer with a stride of 2 for downsampling. At each downsampling step, the spatial resolution of the feature map is halved to expand the receptive field of view in the subnetwork [59], and the number of feature channels is doubled to capture more contextual information [60]. In the expansive path, low-level feature maps are sequentially upsampled through decoder units to reconstruct full-resolution feature maps. Each of the decoder units consists of three subparts which are (1) a 2×2 up-convolution layer that upsamples the spatial resolution of feature maps by a factor of 2 and halves the size of feature map channels, (2) a concatenation layer that merges the upsampled feature maps with the corresponding ones from the contracting path via skip connections to ensure better location accuracy [61] and finally, (3) a sequence of 2 3×3 convolution layers, each activated with a ReLU, to compress the merged feature maps. As the final layer of the network, a 1×1 convolution is applied to produce the segmentation mask. This layer is a linear projection of feature maps to the number of desired object classes. The original U-Net architecture is depth 4 with a total of 23 non-padded convolutional layers.

(b) Generalized Dice loss. Generalized Dice Loss (GDL) is an adaptation of the Generalized Dice Score (GDS) assessment metric [62] for use as a loss function in deep segmentation convolutional neural networks. As formulated in Eq. (2), GDL uses a per-class weighting adjustment to minimize the detrimental impact of class imbalance on the Dice score. The weight of each object class is the inverse-square of its volume, defined by Eq. (3). Accordingly, the contribution of each object class in loss computation is adjusted concerning its volume in the training set.

$$GDL = 1 - 2\frac{\sum_{l=1}^{2} w_l \sum_n r_{ln} p_{ln}}{\sum_{l=1}^{2} w_l \sum_n r_{ln}^2 + p_{ln}^2}, \tag{2}$$

$$w_l = \frac{1}{\left(\sum_{n=1}^{N} r_{ln}\right)^2}, \tag{3}$$

where w_l denotes the weight of the class label, r_{ln} represents the foreground pixel values, p_{ln} represents the foreground prediction maps, and N is the number of pixels.

GDL has been widely used in segmentation studies to overcome class imbalance problems since its introduction in the literature [63–65]. By the way, we would like to emphasize that the integration of GDL in the U-Net architecture is one of the main contributions of our study.

(c) Enhanced U-Net architecture. The proposed enhanced U-Net architecture results from the integration of a GDL-based pixel classification layer into the standard U-Net architecture. The baseline part of the architecture follows the standard U-Net architecture with some retrofitting. First, we reduce the depth of the architecture to 3, yielding in our study a lightweight model without any performance degradation. Second, we apply the same-padding to each convolutional layer within the architecture. Changing the padding scheme to the same-padding ensures that each pixel contained in the image contributes to feature map extraction and allows us to obtain segmentation masks of the same spatial resolution as the input image. As the final layer of the architecture, we integrate a dice pixel classification layer into the base architecture, through which a categorical label is assigned to each pixel using Generalized Dice Loss (GDL). As stated in Section 2.2.2, to minimize class imbalance issues, GDL balances the contribution of each pixel in the loss computation by weighting each class with the inverse-square of its volume in the training set. In summary, the proposed enhanced U-Net architecture consists of three main parts, (1)

Table 2
Layer details of enhanced U-Net architecture.

	Unit level	Layer	Filter	Stride	Output size
Image Input					288 × 288 × 3
Encoder	Level 1	Conv 1	3 × 3/64	1	288 × 288 × 64
		Conv 2	3 × 3/64	1	288 × 288 × 64
		Pooling 1	2 × 2	2	144 × 144 × 64
	Level 2	Conv 3	3 × 3/128	1	144 × 144 × 128
		Conv 4	3 × 3/128	1	144 × 144 × 128
		Pooling 2	2 × 2	2	72 × 72 × 128
	Level 3	Conv 5	3 × 3/256	1	72 × 72 × 256
		Conv 6	3 × 3/256	1	72 × 72 × 256
		Pooling 3	2 × 2	2	36 × 36 × 256
Bridge	Level 4	conv 7	3 × 3/512	1	36 × 36 × 512
		conv 8	3 × 3/512	1	36 × 36 × 512
Decoder	Level 5	Up-conv 1	2 × 2/256	2	72 × 72 × 256
		Conv 9	3 × 3/256	1	72 × 72 × 256
		Conv 10	3 × 3/256	1	72 × 72 × 256
	Level 6	Up-conv 2	2 × 2/128	2	144 × 144 × 128
		Conv 11	3 × 3/128	1	144 × 144 × 128
		Conv 12	3 × 3/128	1	144 × 144 × 128
	Level 7	Up-conv 3	2 × 2/64	2	288 × 288 × 64
		Conv 13	3 × 3/64	1	288 × 288 × 64
		Conv 14	3 × 3/64	1	288 × 288 × 64
Final Conv		Conv 15	1 × 1/2	1	288 × 288 × 2
Softmax		Softmax	–	–	288 × 288 × 2
Generalized Dice Loss		Dice Pixel Classification	–	–	288 × 288 × 2

an encoder part for extracting feature maps, (2) a decoder part for reconstructing feature maps at full spatial resolution and finally, (3) a bridge connecting the encoder and decoder parts, thus ensuring the flow of information. At the end of the decoder part, a dice pixel classification layer is applied to produce segmentation masks. We provide the details of the architecture layers and parameters in Table 2. In total, it contains 18 convolutional and 3 pooling layers. Note that the architecture input size is set at 288 × 288 to ensure that 2 × 2 max-pooling operations are applied on even spatial resolution feature maps, as suggested in [4]. Therefore, the input image patches are resized accordingly at the implementation stage.

(d) Post-processing. Due to the complex texture of histopathological images, masks produced by deep learning models may contain false nuclei. In comparison to the real nuclei, most of these false nuclei are too small in size. In this study, we applied morphological area opening to remove small nuclei that pretend to be false. Additional to an input binary image, morphological area opening requires two more parameters: the number of pixels (p) below which the nucleus is removed and a minimum connectivity degree (cnn). Here these parameters were empirically set as ($p = 100$ and $conn = 8$).

3. Results and discussions

3.1. Experiment design

The experiments are conducted using a PC with an NVDIA RTX 4000 GPU, Intel(R) Xeon(R) W-2245 CPU@3.90 GHz and a 64 GB System RAM on MATLAB platform. The parameter specifications of the deep learning architecture used in the methodology are reported in Table 3. We would like to notify that the parameter values are mostly set to the default values defined in the literature. To prove the superiority of the proposed methodology, we employ the following semantic segmentation architectures:

(a) SegNet [66]: This architecture basically consists of an encoder network and a corresponding decoder network followed by a pixel-based classification layer. There are 13 convolutional layers in the encoder network that are topologically identical to those in the VGG16 network. Decoder networks transform encoder feature maps into full input resolution feature maps to allow pixel-wise classification. In SegNet, the novelty is in the way that feature maps at lower resolutions are upsampled by the decoder. Specifically, the decoder performs non-linear upsampling based on pooling indices calculated during the encoder's max-pooling step.

Table 3
Parameter values of U-Net.

Parameter	Value
Initial learning rate	0.05
Max epochs	60
Mini batch size	2
L2 regularization	0.0001
Optimization	sgdm
Momentum	0.9
Learn rate schedule	piecewise
Gradient threshold	0.05
Encoder depth	3

(b) FCN [5]: FCN consists of a downsampling path for extracting and interpreting context, and an upsampling path for localization. In this architecture, only locally connected layers like convolution, pooling, and upsampling are employed. Accordingly, a training process carried out by FCN can be much more efficient due to avoiding dense layers that have high number of parameters. Another benefit of using only locally connected layers is that FCN can deal with an image dataset with any size.

(c) DeepLabv3+ [67]: This architecture basically consists of an encoder network and a decoder network. While the encoder network is responsible for extracting essential information using a convolutional neural network, the decoder network builds an output by using the information obtained from the encoder network. With the addition of the decoder module, segmentation results along object boundaries are improved. As a CNN architecture, pretrained networks which are ResNet, MobileNetv2 and PSASNet can be used in DeepLabv3+.

To evaluate the performance of nuclei segmentation methods, we use the following evaluation metrics:

1. Mean Accuracy: Accuracy is calculated as the percentage of pixels that have been correctly identified in each class, defined in Eq. (4). Using the accuracy metric, it is possible to determine how well each class identifies pixels correctly. Mean Accuracy is the average accuracy of all classes, defined in Eq. (5).

$$Acc_i = \frac{TP_i}{TP_i + FP_i} \tag{4}$$

$$Mean\ Acc = \frac{1}{C}\sum_{i=1}^{C} Acc_i \tag{5}$$

where TP_i and FP_i respectively denote true and false positives for ith class, and C is the total number of classes.

2. Global Accuracy: It is defined as the ratio of the number of all correctly classified pixels to the total number of pixels. It provides a quick and a computationally efficient calculation to estimate the percentage of correctly classified pixels.

3. Intersection of Union (IoU): The IoU for each class is calculated by dividing the number of correctly classified pixels by the total number of ground truth and predicted pixels. Mean IoU is the average IoU of all classes, defined in Eq. (6), and Weight IoU denotes the average IoU of each class, weighted by the number of pixels in that class, defined in (7).

$$Mean\ IoU = \frac{1}{C}\sum_{i=1}^{C} \frac{TP_i}{TP_i + FP_i + FN_i} \tag{6}$$

$$Weighted\ IoU = \sum_{i=1}^{C} \frac{P_i}{P} \frac{TP_i}{TP_i + FP_i + FN_i} \tag{7}$$

where P_i is the number of pixels for ith class, P is the total number pixels, and FN_i represents false negative for ith class.

4. F1-Score: It is the harmonic average of recall and precision, defined as follows.

$$F1 - Score = 2 * \frac{Recall * Precision}{Recall + Precision} \tag{8}$$

where $Precision$ is the ratio of the number of correctly classified positive instances to the total number of classified positive instances, and $Recall$ is the ratio of the number of correctly classified positive instances to the total number of instances in actual class. Precision and Recall are respectively defined as follows.

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

5. Aggregated Jaccard Index (AJI): AJI is an extension of the Jaccard Index (JI) [10] evaluation metric that uses both object and pixel level information for segmentation performance assessment. It is formulated as follows.

$$AJI = \frac{\sum_{i=1}^{n_\varrho} \left|G_i \cap S\left(G_i\right)\right|}{\sum_{i=1}^{n_\varrho} \left|G_i \cup S\left(G_i\right)\right| + \sum_{k \in K} |S_k|}, \tag{11}$$

where n_ϱ denotes the number of ground truth nuclei, G_i the set of ground truth nuclei, $S\left(G_i\right)$ the set of matching segmented nuclei, and S_k the set of segmented nuclei that did not match any ground truth nucleus. Note that matching two nuclei, a ground truth and a segmented nucleus, involves finding the segmented nucleus with the highest JI score with the relevant ground truth nucleus, defined in Eq. (12).

$$S\left(G_i\right) = argmax_k \frac{\left|G_i \cap S_j\right|}{\left|G_i \cup S_j\right|}, \tag{12}$$

where G_i and S_j are respectively the ground truth and the segmented nuclei. As stated in Section 2.1, AJIs were computed using the source code provided on the MoNuSeg2018 challenge web page [15].

3.2. Experiment results

To prove the superiority of the proposed methodology, we investigate the following points:

- the performance analysis of the proposed methodology through a variety of normalization techniques, demonstrating the impact of normalization on the nuclei segmentation,
- the performance analysis of the proposed methodology through different semantic segmentation architectures, identifying the suitability and robustness of the applied deep learning architecture on the nuclei segmentation,
- the performance analysis of the proposed methodology versus recent studies in the literature, and
- the performance analysis of the proposed methodology from the perspective of encoder depth and loss function.

3.2.1. Comparisons with normalization techniques

The results of the proposed methodology are presented through a variety of different augmentation and normalization techniques in Tables 4 and 5 in terms of global accuracy ('Global Acc'), mean accuracy ('Mean Acc'), mean IoU, weighted IoU and F1-Score. In Table 4, 'Without-Norm' and 'Without-Augment' denote that the proposed methodology is implemented without normalization and augmentation, respectively. Furthermore, the best values are denoted by the bold symbol.

According to Table 4, the proposed methodology with Macenko exhibits significantly better performance than without-Norm. While the mean IoU score obtained by the proposed methodology is nearly 0.82, the mean IoU score obtained by Without-Norm is nearly 0.76. This performance gap can also be illustrated in the other evaluation metrics. It can therefore be suggested that normalization has an indispensable impact on nuclei segmentation. What if without augmentation? From Table 4 it can be seen that the effect of augmentation on nuclei segmentation is as important as the effect of normalization. In other words, it is also possible to significantly improve segmentation performance by applying conventional augmentation techniques. For instance, the mean IoU score was increased from 0.73 to 0.82 thanks to applied augmentation techniques.

When comparing the performance of four normalization techniques in Table 5, Macenko has shown a superior performance than the others, i.e., there exists no such a normalization technique except for Macenko achieves the mean IoU score more than 0.81. Furthermore, the second best segmentation performance is obtained by Reinhard. Moreover, the worst performance is obtained CLAHE. It can therefore be indicated that Macenko is the most appropriate technique for H&E stained histopathology images.

3.2.2. Comparisons with deep segmentation algorithms

The results of the proposed methodology are presented with different semantic segmentation algorithms in Table 6 in terms of the evaluation criteria defined in Table 5. In addition to those criteria, the run time (in minutes) is also used to verify the efficiency of the algorithms. Furthermore, the best values are denoted by the bold symbol.

According to the results, the proposed methodology based on an enhanced version of U-Net achieves far superior segmentation performance than the other segmentation algorithms. For example, the proposed methodology has obtained a mean IoU score of 0.82, whereas the other semantic segmentation algorithms have not been able to achieve a score higher than 0.79. For another instance, the proposed methodology has achieved a mean F1-Score of 0.88, while the other algorithms only obtain a score between 0.78 and 0.83 for the same criterion. Furthermore, the proposed methodology has also proven its outstanding performance in terms of the running time. Specifically, the enhanced U-Net architecture utilized in the proposed methodology and

Table 4
The effect of normalization and augmentation.

Technique	Global Acc	Mean Acc	Mean IoU	Weighted IoU	F1-Score
Without-Norm	0.9201	0.8371	0.7647	0.8546	0.8332
Without-Augment	0.8893	**0.8978**	0.7360	0.8181	0.8075
Macenko&Augment	**0.9381**	0.8941	**0.8203**	**0.8851**	**0.8809**

Table 5
Results of normalization techniques.

Technique	Global Acc	Mean Acc	Mean IoU	Weighted IoU	F1-Score
CLAHE&Augment	0.9234	0.8223	0.7646	0.8575	0.8425
Vahanade&Augment	0.9280	0.8441	0.7822	0.8668	0.8568
Reinhard&Augment	0.9357	0.8602	0.8035	0.8800	0.8695
Macenko&Augment	**0.9381**	**0.8941**	**0.8203**	**0.8851**	**0.8809**

Table 6
Results of semantic segmentation algorithms.

Algorithm	Global Acc	Mean Acc	Mean IoU	Weighted IoU	F1-Score	Run time
SegNet	0.9027	0.9089	0.7662	0.8357	0.7815	**58**
FCN	0.90771	**0.9109**	0.77484	0.84287	0.80431	137
DeepLabv3+	0.9233	0.8767	0.7911	0.8624	0.8336	110
Proposed	**0.9381**	0.8941	**0.8203**	**0.8851**	**0.8809**	63

Table 7
Results of recent studies in terms of F1-Score and IoU.

	F1-Score	IoU
U-Net [4]	0.7943	0.6599
U-Net++ [6]	0.7949	0.6604
Res-UNet [68]	0.7949	0.6607
Axial Attention U-Net [69]	0.7683	0.6249
Gated Axial Atten. [70]	0.7644	0.6201
LoGo [70]	0.7956	0.6617
MedT [70]	0.7955	0.6617
HistoSeg [71]	0.7508	0.7106
Proposed (Class)	0.8418	0.7123
Proposed (Mean)	**0.8809**	**0.8203**

Table 8
Results of recent studies in terms of F1-Score and AJI.

Study	F1-Score	AJI
CDNet [73]	0.8705	0.6331
G-U-Net [65]	0.8469	0.6291
HoVer-Net [74]	–	0.620
Dual U-Net [75]	0.8196	0.6591
Mask R-CNN & ResNet50 [76]	0.7247	0.6217
Mask R-CNN & ResNet101 [76]	0.7519	0.6551
CBA without blur-pooling [76]	0.7640	0.6805
CBA with blur-pooling [76]	0.8247	**0.7985**
Proposed	**0.8809**	0.6895

SegNet can complete the training process in nearly an hour, whereas DeepLabv3+ and FCN need nearly two hours. Despite its efficiency, SegNet has the worst segmentation performance among all the algorithms. It can therefore be suggested that the proposed methodology based on an enhanced version of U-Net is the most robust and efficient algorithm for nuclei segmentation.

3.2.3. Comparisons with recent studies

The results of recent studies with the proposed methodology are presented in Table 7 in terms of the F1-Score and IoU criteria, and in Table 8 in terms of the F1-Score and AJI criteria. In Table 7, we also reported the results of F1-Score and IoU for the nuclei class as well as the results of mean F1-Score and mean IoU. Furthermore, the best values are represented by the bold symbol as in Tables 5 and 6.

According to Table 7, the proposed methodology outperforms all recent studies in terms of F1-Score and IoU. To be specific, it can be illustrated that the performance gap is extremely high between the proposed methodology and most of the recent studies. For instance, the proposed methodology obtains a IoU score of 0.82, while the other studies obtain a score between 0.65 and 0.71. In addition to mean IoU and F1-Score, the proposed methodology is superior in class-based IoU and F1-Score. When considering Table 8, the proposed methodology has generally also shown a superior performance to existing studies in terms of F1-Score and AJI. The proposed methodology only cannot perform better than CBA with blur-pooling in terms of the AJI score, but outperforms CBA with blur-pooling in terms of F1-Score. The reasons why CBA with blur-pooling scored well in the AJI are as follows: (1) blur pooling may have a significant effect on the semantic performance (Notice that we could not make a deep analysis on CBA since the source code was not shared with the researchers); (2) instead of original-sized output images, sliced test images were used in CBA to compute the evaluation metrics. It should be also noted that the AJI score obtained by our proposed methodology is the third best result among 36 results reported in 2018 MICCAI MoNuSeg Challenge [72]. Furthermore, Fig. 2 also demonstrates the effectiveness of the proposed methodology with some visual results. It can therefore be indicated that the proposed methodology can achieve successful and competitive results compared to the recent work.

How effective and efficient is the proposed methodology for nuclei segmentation? Firstly, the proposed methodology includes a well-designed preprocessing stage, which directly impacts the training process. Specifically, artifacts within images are eliminated through normalization, followed by resize-splitting and data augmentation to increase the size of the training set. As the data growth in the training set is not high, the training process can be completed within a reasonable timeframe. A second advantage of the proposed methodology is that it has an enhanced U-Net architecture that is more lightweight than a standard U-Net architecture without any performance degradation. As a final benefit of the enhanced U-Net architecture, it has a generalized Dice loss-based classification layer that prevents class-imbalance problems. As a result of bringing all these relevant factors together, we can conclude that the proposed methodology is well-organized and well-prepared.

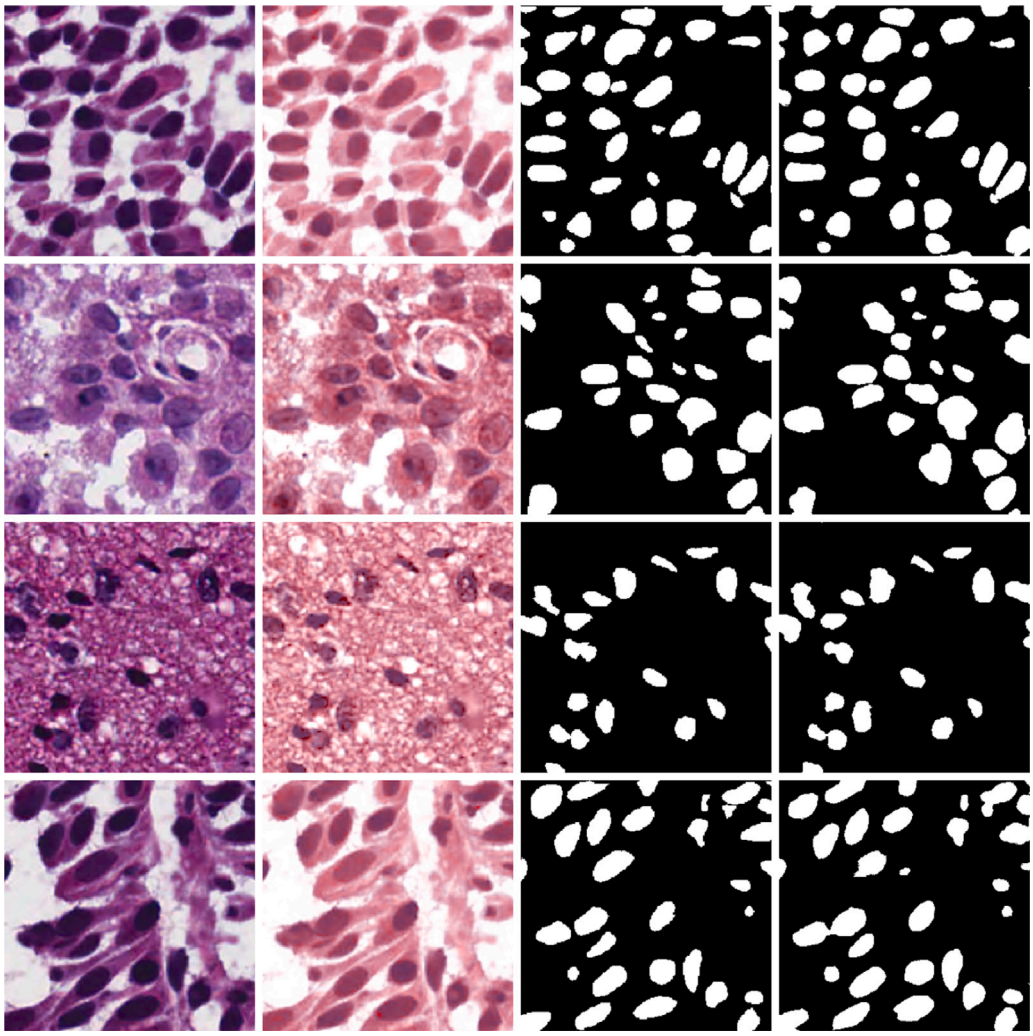


Fig. 2. The visual normalization and segmentation results of the proposed imbalance-aware nuclei segmentation methodology for H&E stained histopathology images. Left to right: sample images from the MoNuSeg2018 test subset, normalized version of the sample images, ground truth masks, and predicted segmentation masks.

3.2.4. Further comparisons

We further compare the proposed methodology from the perspective of the loss function and encoder depth. The results are presented in Table 9, where 'Loss' and 'Depth' respectively represent the loss function and encoder depth. In the 'Loss' column, 'Cross' represents the cross-entropy loss function.

It can be indicated from Table 9 that the encoder depth does not have a significant impact on the performance of nuclei segmentation, i.e., similar results are generally produced even for different values of encoder depth. Only for one case of GDL, the methodology has shown slightly worse performance than the other cases. However, run time tends to increase proportionally to encoder depth. Therefore, the optimal value of encoder depth is determined as 3. When comparing loss functions, it can be inferred that GDL generally performs better than Cross for all depth cases in terms of mean IoU, weighted IoU and F1-Score. It can therefore be indicated that GDL is the most appropriate loss function for U-Net architecture in nuclei segmentation.

4. Conclusion

Due to the development of technology, manual operations are increasingly seen as time-consuming, labor-intensive and fraught with error. A pathologist's diagnostic process is the most obvious example of this. The pathologist examines nuclei under a microscope to determine structures, shapes, colors, densities, etc. Further examinations may be necessary if cancerous areas are detected. In other words, the human-based process involves both a great deal of time-consuming and a higher risk of error. To alleviate this issue, the majority of current studies aim to automate these processes and use quantitative measurements as a basis.

In this study, we introduced a U-Net-based segmentation methodology for H&E stained histopathology images. To improve the effectiveness of the U-Net architecture, we designed a preprocessing stage which applies color normalization to remove artifacts in H&E images and then increases the size of a training set through resize-split and data augmentation techniques. Moreover, we adopted a Generalized Dice Loss in the U-Net architecture to address data imbalance problems since the density of nuclei in image patches may significantly affect the background class over the foreground class. It is pertinent to note that the enhanced U-Net architecture utilized in the proposed methodology is lighter than the standard U-Net architecture. According to a number of experiments, the proposed methodology outperforms a variety of recent works in terms of evaluation metrics. Furthermore, the enhanced U-NET architecture used in the proposed methodology achieved better performance than SegNet, FCN and DeepLabv3+ even used the same increased training set. Even though the proposed methodology achieved promising results, overlapping and clumped nuclei present in H&E images still remain challenges for nuclei segmentation. In the

Table 9
Results for different parameters.

Loss	Depth	Global Acc	Mean Acc	Mean IoU	Weighted IoU	F1-Score	Run time
Cross	2	0.9315	0.9076	0.8107	0.8781	0.8665	54
	3	0.9379	0.8739	0.8131	0.8847	0.8757	62
	4	0.9291	0.9232	0.8102	0.8757	0.8735	75
	5	0.9264	0.925	0.8057	0.8719	0.8681	124
GDL	2	0.9326	0.8568	0.8014	0.8750	0.8985	49
	3	0.9381	0.8941	0.8203	0.8851	0.8809	63
	4	0.9392	0.8886	0.8204	0.8882	0.8811	78
	5	0.9385	0.8946	0.8206	0.8876	0.8814	127

future, we will focus on developing a loss function to deal with this problematic issue. Moreover, we will focus on transformers which is a recently introduced architecture for nuclei segmentation.

CRediT authorship contribution statement

Emrah Hancer: Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Mohamed Traoré:** Conception and design of study, Acquisition of data, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Refik Samet:** Conception and design of study, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Zeynep Yıldırım:** Acquisition of data, Analysis and/or interpretation of data, Writing – review & editing. **Nooshin Nemati:** Conception and design of study, Acquisition of data, Writing – original draft.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Emrah Hancer reports financial support was provided by Scientific and Technological Research Council of Turkey.

Data availability

The data that has been used is confidential.

Acknowledgments

This work is supported by The Scientific and Technological Research Council of Turkey (TUBITAK) (Project number: 121E379). All authors approved the version of the manuscript to be published.

References

[1] H. Irshad, A. Veillard, L. Roux, D. Racoceanu, Methods for nuclei detection, segmentation, and classification in digital histopathology: A review—Current status and future potential, IEEE Rev. Biomed. Eng. 7 (2014) 97–114.
[2] Z. Yildirim, E. Hancer, R. Samet, M.T. Mali, N. Nemati, Effect of color normalization on nuclei segmentation problem in H&E stained histopathology images, in: 30th Signal Processing and Communications Applications Conference (SIU2022), 2022, pp. 1–4.
[3] E. Mentese, E. Hancer, Nucleus segmentation with deep learning approaches on histopathology images, Eur. J. Sci. Technol. Special Issue: ISMSIT (2020) 95–102.
[4] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.
[5] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR2015), 2015, pp. 3431–3440.
[6] Z. Zhou, M.M. Rahman Siddiquee, N. Tajbakhsh, J. Liang, Unet++: A nested U-Net architecture for medical image segmentation, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Springer, 2018, pp. 3–11.
[7] X. Li, Y. Wang, Q. Tang, Z. Fan, J. Yu, Dual U-Net for the segmentation of overlapping glioma nuclei, IEEE Access 7 (2019) 84040–84052.
[8] A. Mahbod, G. Schaefer, I. Ellinger, R. Ecker, Ö. Smedby, C. Wang, A two-stage U-Net algorithm for segmentation of nuclei in H&E-stained tissues, in: European Congress on Digital Pathology, Springer, 2019, pp. 75–82.
[9] Y. Kong, G.Z. Genchev, X. Wang, H. Zhao, H. Lu, Nuclear segmentation in histopathological images using two-stage stacked U-Nets with attention mechanism, Front. Bioeng. Biotechnol. 8 (2020) 573866.
[10] J. Zhang, C. Li, S. Kosov, M. Grzegorzek, K. Shirahama, T. Jiang, C. Sun, Z. Li, H. Li, LCU-Net: A novel low-cost U-Net for environmental microorganism image segmentation, Pattern Recognit. 115 (2021) 107885.
[11] A. Vahadane, B. Atheeth, S. Majumdar, Dual Encoder Attention U-Net for nuclei segmentation, in: 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC2021), IEEE, 2021, pp. 3205–3208.
[12] I. Kiran, B. Raza, A. Ijaz, M.A. Khan, DenseRes-Unet: Segmentation of overlapped/clustered nuclei from multi organ histopathology images, Comput. Biol. Med. 143 (2022) 105267.
[13] T. Hayakawa, V.B.S. Prasath, H. Kawanaka, B.J. Aronow, S. Tsuruoka, Computational nuclei segmentation methods in digital pathology: A survey, Arch. Comput. Methods Eng. 28 (2021) 1–13.
[14] N. Kumar, R. Verma, S. Sharma, S. Bhargava, A. Vahadane, A. Sethi, A dataset and a technique for generalized nuclear segmentation for computational pathology, IEEE Trans. Med. Imaging 36 (7) (2017) 1550–1560.
[15] Multi-Organ Nuclei Segmentation Challenge (MoNuSeg), 2018, Available: https://monuseg.grand-challenge.org/ [Online]. Accessed: May. 24, 2022.
[16] M. Macenko, M. Niethammer, J.S. Marron, D. Borland, J.T. Woosley, X. Guan, C. Schmitt, N.E. Thomas, A method for normalizing histology slides for quantitative analysis, in: IEEE International Symposium on Biomedical Imaging: From Nano To Macro, IEEE, 2009, pp. 1107–1110.
[17] J. Lee, H. Kim, H. Cho, Y. Jo, Y. Song, D. Ahn, K. Lee, Y. Park, S.-J. Ye, Deep-learning-based label-free segmentation of cell nuclei in time-lapse refractive index tomograms, IEEE Access 7 (2019) 83449–83460.
[18] N. Kumar, R. Verma, D. Anand, Y. Zhou, O.F. Onder, E. Tsougenis, H. Chen, P.-A. Heng, J. Li, Z. Hu, et al., A multi-organ nucleus segmentation challenge, IEEE Trans. Med. Imaging 39 (5) (2019) 1380–1391.
[19] E. Reinhard, M. Adhikhmin, B. Gooch, P. Shirley, Color transfer between images, IEEE Comput. Graph. Appl. 21 (5) (2001) 34–41.
[20] A. Vahadane, T. Peng, S. Albarqouni, M. Baust, K. Steiger, A.M. Schlitter, A. Sethi, I. Esposito, N. Navab, Structure-preserved color normalization for histological images, in: 12th IEEE International Symposium on Biomedical Imaging (ISBI2015), IEEE, 2015, pp. 1012–1015.
[21] E. Maggiori, Y. Tarabalka, G. Charpiat, P. Alliez, High-resolution image classification with convolutional networks, in: IEEE International Geoscience and Remote Sensing Symposium (IGARSS2017), IEEE, 2017, pp. 5157–5160.
[22] J. Yamanaka, S. Kuwashima, T. Kurita, Fast and accurate image super resolution by deep CNN with skip connection and network in network, in: International Conference on Neural Information Processing, Springer, 2017, pp. 217–225.
[23] S. Lal, D. Das, K. Alabhya, A. Kanfade, A. Kumar, J. Kini, NucleiSegNet: robust deep learning architecture for the nuclei segmentation of liver cancer histopathology images, Comput. Biol. Med. 128 (2021) 104075.
[24] L. Zhang, B. Li, InvUnet: Inverse the unet for nuclear segmentation in H&E stained images, in: 12th International Conference on Advanced Computational Intelligence (ICACI2020), IEEE, 2020, pp. 251–256.
[25] Nearest neighbor image scaling, 2022.
[26] A. Hernández-García, P. König, Data augmentation instead of explicit regularization, 2018, arXiv preprint arXiv:1806.03852.
[27] Z. Zhong, L. Zheng, G. Kang, S. Li, Y. Yang, Random erasing data augmentation, in: Proceedings of the AAAI Conference on Artificial Intelligence, 34, (07) 2020, pp. 13001–13008.
[28] L. Perez, J. Wang, The effectiveness of data augmentation in image classification using deep learning, 2017, arXiv preprint arXiv:1712.04621.
[29] T. DeVries, G.W. Taylor, Improved regularization of convolutional neural networks with cutout, 2017, arXiv preprint arXiv:1708.04552.
[30] C. Shorten, T.M. Khoshgoftaar, A survey on image data augmentation for deep learning, J. Big Data 6 (1) (2019) 1–48.
[31] Z. Huang, Y. Ding, R. Geng, H. He, X. Huang, J. Chen, Learning erosional probability maps for nuclei instance segmentation, in: 2020 IEEE Conference on Multimedia Information Processing and Retrieval, MIPR, IEEE, 2020, pp. 297–302.

[32] J. Xie, R. Zhu, Z. Wu, J. Ouyang, FFUNet: A novel feature fusion makes strong decoder for medical image segmentation, IET Signal Process. (2022).

[33] E.D. Cubuk, B. Zoph, J. Shlens, Q.V. Le, Randaugment: Practical automated data augmentation with a reduced search space, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 702–703.

[34] S. Lim, I. Kim, T. Kim, C. Kim, S. Kim, Fast autoaugment, Adv. Neural Inf. Process. Syst. 32 (2019).

[35] V. Delchevalerie, A. Bibal, B. Frénay, A. Mayer, Achieving rotational invariance with bessel-convolutional neural networks, Adv. Neural Inf. Process. Syst. 34 (2021) 28772–28783.

[36] B. Chidester, T.-V. Ton, M.-T. Tran, J. Ma, M.N. Do, Enhanced rotation-equivariant U-Net for nuclear segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2019.

[37] B. Chidester, T. Zhou, M.N. Do, J. Ma, Rotation equivariant and invariant neural networks for microscopy image analysis, Bioinformatics 35 (14) (2019) i530–i537.

[38] E.D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, Q.V. Le, Autoaugment: Learning augmentation strategies from data, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 113–123.

[39] F. Quiroga, F. Ronchetti, L. Lanzarini, A.F. Bariviera, Revisiting data augmentation for rotational invariance in convolutional neural networks, in: International Conference on Modelling and Simulation in Management Sciences, Springer, 2018, pp. 127–141.

[40] T. Wan, L. Zhao, H. Feng, D. Li, C. Tong, Z. Qin, Robust nuclei segmentation in histopathology using ASPPU-Net and boundary refinement, Neurocomputing 408 (2020) 144–156.

[41] H. He, C. Zhang, J. Chen, R. Geng, L. Chen, Y. Liang, Y. Lu, J. Wu, Y. Xu, A hybrid-attention nested UNet for nuclear segmentation in histopathological images, Front. Mol. Biosci. 8 (2021) 614174.

[42] P.F. Jaeger, S.A. Kohl, S. Bickelhaupt, F. Isensee, T.A. Kuder, H.-P. Schlemmer, K.H. Maier-Hein, Retina U-Net: Embarrassingly simple exploitation of segmentation supervision for medical object detection, in: Machine Learning for Health Workshop, PMLR, 2020, pp. 171–183.

[43] B. Cheng, A. Schwing, A. Kirillov, Per-pixel classification is not all you need for semantic segmentation, Adv. Neural Inf. Process. Syst. 34 (2021) 17864–17875.

[44] S. Kohl, B. Romera-Paredes, C. Meyer, J. De Fauw, J.R. Ledsam, K. Maier-Hein, S. Eslami, D. Jimenez Rezende, O. Ronneberger, A probabilistic U-Net for segmentation of ambiguous images, Adv. Neural Inf. Process. Syst. 31 (2018).

[45] T. Dhamija, A. Gupta, S. Gupta, R. Katarya, G. Singh, et al., Semantic segmentation in medical images through transfused convolution and transformer networks, Appl. Intell. (2022) 1–17.

[46] M.K. Kar, M.K. Nath, D.R. Neog, A review on progress in semantic image segmentation and its application to medical images, SN Comput. Sci. 2 (5) (2021) 1–30.

[47] P. Wang, P. Chen, Y. Yuan, D. Liu, Z. Huang, X. Hou, G. Cottrell, Understanding convolution for semantic segmentation, in: IEEE Winter Conference on Applications of Computer Vision (WACV2018), Ieee, 2018, pp. 1451–1460.

[48] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.

[49] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, Pyramid scene parsing network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2881–2890.

[50] N. Ibtehaz, M.S. Rahman, MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation, Neural Netw. 121 (2020) 74–87.

[51] A. Sevastopolsky, Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network, Pattern Recognit. Image Anal. 27 (3) (2017) 618–624.

[52] M.S. Hossain, J.M. Betts, A.P. Paplinski, Dual focal loss to address class imbalance in semantic segmentation, Neurocomputing 462 (2021) 69–87.

[53] D. Anand, G. Patel, Y. Dang, A. Sethi, Switching loss for generalized nucleus detection in histopathology, 2020, arXiv preprint arXiv:2008.03750.

[54] P.O. Bressan, J.M. Junior, J.A.C. Martins, M.J. de Melo, D.N. Gonçalves, D.M. Freitas, A.P.M. Ramos, M.T.G. Furuya, L.P. Osco, J. de Andrade Silva, et al., Semantic segmentation with labeling uncertainty and class imbalance applied to vegetation mapping, Int. J. Appl. Earth Obs. Geoinf. 108 (2022) 102690.

[55] F. Milletari, N. Navab, S.-A. Ahmadi, V-Net: Fully convolutional neural networks for volumetric medical image segmentation, in: Fourth International Conference on 3D Vision (3DV), 2016, pp. 565–571.

[56] C.H. Sudre, W. Li, T. Vercauteren, S. Ourselin, M. Jorge Cardoso, Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Springer, 2017, pp. 240–248.

[57] T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, Focal loss for dense object detection, in: IEEE International Conference on Computer Vision (ICCV2017), 2017, pp. 2999–3007.

[58] M.S. Hossain, J.M. Betts, A.P. Paplinski, Dual focal loss to address class imbalance in semantic segmentation, Neurocomputing 462 (2021) 69–87.

[59] Y. Han, J.C. Ye, Framing U-Net via deep convolutional framelets: Application to sparse-view CT, IEEE Trans. Med. Imaging 37 (6) (2018) 1418–1429.

[60] V. Iglovikov, A. Shvets, Ternausnet: U-Net with vgg11 encoder pre-trained on imagenet for image segmentation, 2018, arXiv preprint arXiv:1801.05746.

[61] M.Z. Alom, M. Hasan, C. Yakopcic, T.M. Taha, V.K. Asari, Recurrent residual convolutional neural network based on U-Net (r2U-Net) for medical image segmentation, 2018, arXiv preprint arXiv:1802.06955.

[62] W.R. Crum, O. Camara, D.L. Hill, Generalized overlap measures for evaluation and validation in medical image analysis, IEEE Trans. Med. Imaging 25 (11) (2006) 1451–1461.

[63] X. Xue, S.-I. Kamata, Contextual mixing feature U-net for multi-organ nuclei segmentation, Front. Signal Process. (2022) 2.

[64] M. Lou, J. Meng, Y. Qi, X. Li, Y. Ma, MCRNet: Multi-level context refinement network for semantic segmentation in breast ultrasound imaging, Neurocomputing 470 (2022) 154–169.

[65] B. Chidester, T.-V. Ton, M.-T. Tran, J. Ma, M.N. Do, Enhanced rotation-equivariant U-Net for nuclear segmentation, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR2019), 2019, pp. 1097–1104.

[66] V. Badrinarayanan, A. Kendall, R. Cipolla, SegNet: A deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (12) (2017) 2481–2495.

[67] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-decoder with atrous separable convolution for semantic image segmentation, in: V. Ferrari, M. Hebert, C. Sminchisescu, Y. Weiss (Eds.), Computer Vision – ECCV 2018, Springer International Publishing, Cham, 2018, pp. 833–851.

[68] X. Xiao, S. Lian, Z. Luo, S. Li, Weighted Res-UNet for high-quality retina vessel segmentation, in: 9th International Conference on Information Technology in Medicine and Education (ITME2018), 2018, pp. 327–331.

[69] H. Wang, Y. Zhu, B. Green, H. Adam, A. Yuille, L.-C. Chen, Axial-DeepLab: Stand-alone axial-attention for panoptic segmentation, in: A. Vedaldi, H. Bischof, T. Brox, J.-M. Frahm (Eds.), Computer Vision(ECCV2020), Springer International Publishing, Cham, 2020, pp. 108–126.

[70] J.M.J. Valanarasu, P. Oza, I. Hacihaliloglu, V.M. Patel, Medical transformer: Gated axial-attention for medical image segmentation, in: M. de Bruijne, P.C. Cattin, S. Cotin, N. Padoy, S. Speidel, Y. Zheng, C. Essert (Eds.), Medical Image Computing and Computer Assisted Intervention (MICCAI2021), Springer International Publishing, Cham, 2021, pp. 36–46.

[71] S. Wazir, M.M. Fraz, HistoSeg: Quick attention with multi-loss function for multi-structure segmentation in digital histology images, in: 12th International Conference on Pattern Recognition Systems (ICPRS2022), 2022, pp. 1–7.

[72] N. Kumar, R. Verma, D. Anand, A. Sethi, MONUSEG - Grand challenge, 2022, Accessed: 2022-09-06. https://monuseg.grand-challenge.org/Results/,

[73] H. He, Z. Huang, Y. Ding, G. Song, L. Wang, Q. Ren, P. Wei, Z. Gao, J. Chen, CDNet: Centripetal direction network for nuclear instance segmentation, in: IEEE/CVF International Conference on Computer Vision (ICCV2021), 2021, pp. 4006–4015.

[74] S. Graham, Q.D. Vu, S.E.A. Raza, A. Azam, Y.W. Tsang, J.T. Kwak, N. Rajpoot, Hover-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images, Med. Image Anal. 58 (2019) 101563.

[75] X. Li, Y. Wang, Q. Tang, Z. Fan, J. Yu, Dual U-Net for the segmentation of overlapping glioma nuclei, IEEE Access 7 (2019) 84040–84052.

[76] P. Thi Le, T. Pham, Y.-C. Hsu, J.-C. Wang, Convolutional blur attention network for cell nuclei segmentation, Sensors 22 (4) (2022).