

## Article

# GSN-HVNET: A Lightweight, Multi-Task Deep Learning Framework for Nuclei Segmentation and Classification

Tengfei Zhao <sup>1</sup>, Chong Fu <sup>1,2,3,\*</sup>, Yunjia Tian <sup>4</sup>, Wei Song <sup>1</sup> and Chiu-Wing Sham <sup>5</sup><sup>1</sup> School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China<sup>2</sup> Engineering Research Center of Security Technology of Complex Network System, Ministry of Education, Shenyang 110819, China<sup>3</sup> Key Laboratory of Intelligent Computing in Medical Image, Ministry of Education, Northeastern University, Shenyang 110819, China<sup>4</sup> State Grid Liaoning Information and Communication Company, Shenyang 110006, China<sup>5</sup> School of Computer Science, The University of Auckland, Auckland 1142, New Zealand

\* Correspondence: fuchong@mail.neu.edu.cn

**Abstract:** Nuclei segmentation and classification are two basic and essential tasks in computer-aided diagnosis of digital pathology images, and those deep-learning-based methods have achieved significant success. Unfortunately, most of the existing studies accomplish the two tasks by splicing two related neural networks directly, resulting in repetitive computation efforts and a redundant-and-large neural network. Thus, this paper proposes a lightweight deep learning framework (GSN-HVNET) with an encoder–decoder structure for simultaneous segmentation and classification of nuclei. The decoder consists of three branches outputting the semantic segmentation of nuclei, the horizontal and vertical (HV) distances of nuclei pixels to their mass centers, and the class of each nucleus, respectively. The instance segmentation results are obtained by combing the outputs of the first and second branches. To reduce the computational cost and improve the network stability under small batch sizes, we propose two newly designed blocks, Residual-Ghost-SN (RGS) and Dense-Ghost-SN (DGS). Furthermore, considering the practical usage in pathological diagnosis, we redefine the classification principle of the CoNSeP dataset. Experimental results demonstrate that the proposed model outperforms other state-of-the-art models in terms of segmentation and classification accuracy by a significant margin while maintaining high computational efficiency.

**Keywords:** joint nuclei segmentation and classification; lightweight, multi-task deep learning framework; Residual-Ghost-SN; Dense-Ghost-SN



**Citation:** Zhao, T.; Fu, C.; Tian, Y.; Song, W.; Sham, C.-W. GSN-HVNET: A Lightweight, Multi-Task Deep Learning Framework for Nuclei Segmentation and Classification. *Bioengineering* **2023**, *10*, 393. <https://doi.org/10.3390/bioengineering10030393>

Academic Editors: Yan Pei, Jijiang Yang and Liang Luo

Received: 12 February 2023

Revised: 13 March 2023

Accepted: 20 March 2023

Published: 22 March 2023



**Copyright:** © 2023 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

Over the past several years, deep-learning-based computer vision techniques have been extensively applied to computer-aided diagnosis (CAD). In computational pathology, pathological image analysis based on the deep learning method has proven powerful in improving efficiency and accuracy in cancer detection [1]. The morphology of the nuclei is the essential feature used by pathologists in cancer diagnosis and further cancer prognoses, such as predicting survival [2] and pathological grading of tumors [3]. Accurate nuclei segmentation and classification can advance the quality of tissue segmentation [4,5]. Nuclei segmentation is the crucial first step to obtaining the morphological features used in the downstream analysis. However, the morphological heterogeneity of nuclei makes studies challenging. The karyomorphism shows variability, while different diseases may cause chromatin abnormalities to exhibit variable size and shape patterns. Another severe problem is that the cells in a cancerous tumor are usually densely packed and even have more than one nucleus, causing overlapping nuclei. This overlapping brings difficulty for further research on separating neighboring instances via automatic segmentation.

Extracting each nucleus and distinguishing its type can promote the diagnostic potential in present-day digital pathology pipelines. For instance, precisely distinguishing each nucleus from tumors or lymphocytes can significantly facilitate downstream analysis of tumor-infiltrating lymphocytes (TIL), which has been proven effective in predicting cancer recurrence [6]. The nucleus-by-nucleus classification has become another problem researchers have been interested in recently due to the high variability and diversity of nuclei appearance in a whole slide image.

The current deep models for histopathology image diagnosis are mainly based on single-task learning. Single-task learning is designing a model for a specific task and then optimizing iteratively. In this case, the nuclei segmentation and classification tasks require two independent models, one for detecting the location of each nucleus and the other for classifying the type of nuclei [7,8]. For more complicated tasks, we are accustomed to modeling each part of the task by disassembling. However, there exists an obvious problem in this way. When modeling each sub-task, it is easy to ignore the relationships, conflicts, and constraints between different sub-tasks, resulting in the downgrading of the overall performance of the entire task.

To address the above issue, multi-task models have drawn much attention [9–12]. The multi-task models have the following advantages: (1) multiple tasks share the same model, reducing the amount of memory; (2) multiple tasks obtain results through a forward calculation at one time, and the inference speed increases; (3) associated tasks share information and complement each other, improving each tasks' performance.

Recently, several multi-task deep models for histopathology image diagnosis have been suggested and achieved promising results [13–15]. Unfortunately, these approaches still suffer from efficiency issues, such as dealing with a cumbersome model with a huge amount of parameters. In addition, the classification on the CoNSeP dataset [13] seems hard to meet the needs of practical pathological diagnosis.

The present paper proposes a lightweight, multi-task deep learning framework for segmenting and classifying nuclei simultaneously. To address the problem of network stability encountered by batch normalization (BN) when dealing with small batch sizes, we introduce two newly designed blocks. We devise an efficient encoder–decoder architecture, where the encoder adopts our proposed RGS for down-sampling, while the decoder uses Dense-Ghost-Module (DGM) and convolution for up-sampling. By encoding the HV distance of nuclei pixels, we can obtain more representative features on the instance with fewer layers. Here, HV distance can be used to segment overlapping nuclei instances accurately. Later, the decoder using the output features of the encoder predicts nuclei types. According to the above characteristics, we call the proposed network GSN-HVNET. Our experimental results show that the proposed model can retain shallow features on nuclei to improve segmentation and classification accuracy. Our main contributions are outlined below:

- We propose a novel, lightweight, multi-task deep learning framework containing a unified model for segmentation and classification of nuclei instances simultaneously with superior efficiency and accuracy.
- We propose the newly designed RGS and DGS to improve accuracy and compress the training model.
- We redefine the classification principle of the CoNSeP dataset so that the auxiliary diagnostic results have practical significance in pathological diagnosis.
- Our experiments on the CoNSeP, Kumar, and CPM-17 datasets confirm the improvements to existing works [13,14]. Compared with the state-of-the-art HoVer-Net [13], the number of parameters is reduced by 64%. In addition, we try different batch sizes in our experiments and prove that batch size is no longer a strict limitation on the proposed network; even when a small batch is presented, the proposed network can maintain a high performance.

The remainder of this paper is organized as follows: Section 2 introduces the current research on applying learning algorithms in histopathology image analysis. Our new

network architecture is presented in Section 3. We conduct experiments and show desirable results in Section 4. Finally, Section 5 concludes our work and gives a brief discussion of future work.

## 2. Related Work

### 2.1. Nuclei Segmentation

Nuclei segmentation is the crucial first step in computer-aided systems for cancer detection. Low level information analysis of histopathology images, such as histograms analysis [16–20], were often used for early nuclei segmentation algorithms. There was an obvious shortcoming that occurred to those algorithms. A certain threshold was hard to be determined to adapt to all scenarios. In [21], the authors proposed a fast and flexible segmentation algorithm based on computing the watersheds in digital grayscale images. Unfortunately, related experiments reported several false-positive segmentation cases. In [22], the authors proposed a novel, marker-controlled watershed based on mathematical morphology to segment clustered cells with less oversegmentation, designing a tracking method based on modified mean shift algorithm to segment undersegmented cells or merge oversegmented cells. In [23], the authors proposed a method combining region growing and machine learning to segment touching nuclei and classify them. In [24], the authors proposed an improved method, which used a joint optimization of a multiple-level set function to segment the cytoplasm and nuclei from clumps of overlapping cervical cells. In [25], the authors proposed using the graph theory technique to segment glands and computed a gland-score for estimating how similar a segmented region is to a gland. In [26], the authors proposed a superpixel-based segmentation technique with different morphological and clustering algorithms. Unfortunately, these existing segmentation algorithms cannot provide utterly reliable results because they need to manually extract nucleus features, which, thus, are inflexible and laborious to extend to a complex scenario.

Rather than manual feature extraction in traditional algorithms, deep learning methods can automatically extract a distinct set of features, and have been widely applied to nuclei segmentation [27]. For instance, U-Net has presented an outstanding performance in biomedical image segmentation [28]. In [29], the authors proposed a deep multi-scale neural network for accurately segmenting nuclei by improving sensitivity to hematoxylin intensity. In [30], to meet the challenge of segmenting overlapping or touching nuclei, the authors formulated the segmentation problem as a regression task of the distance map, and the nuclei boundary information was used as prior knowledge for a segmentation network. In [31], the authors proposed a contour-aware informative aggregation network with a multi-level information aggregation module between two task decoders: one of these segments the nuclei, and the other segments the contours.

### 2.2. Nuclei Classification

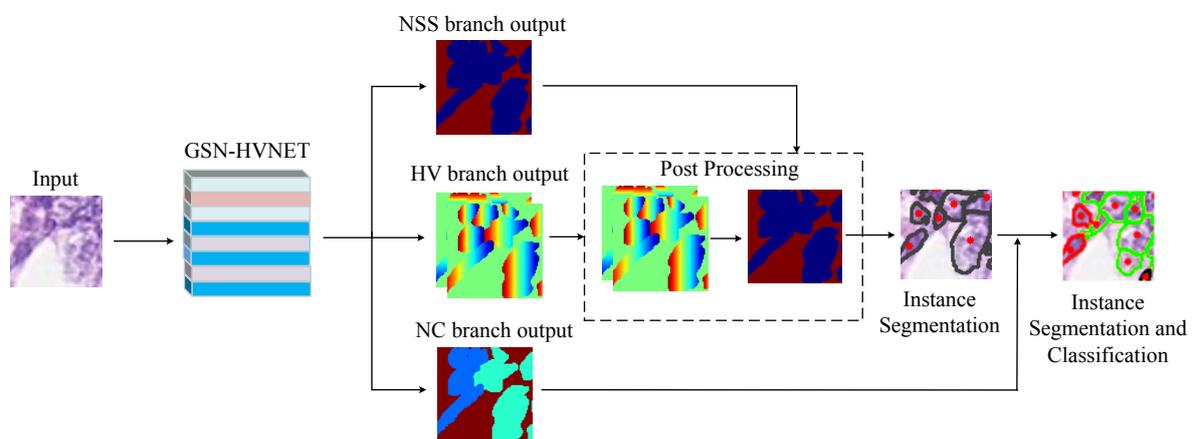
Nuclei classification is a vital step in histopathology image analysis, promoting downstream analysis such as evaluating cancer progression. Early studies utilize manually extracted features to classify the nuclei automatically. Typically, an SVM-based method [32] applied iterative feedback to obtain subtle and complex features of cellular morphology. Albeit showing good performance in high-penetrant phenotypes, it can hardly achieve a satisfying performance in lower-penetrant phenotypes. In [33], Ada-boost was used as the classifier to classify the nuclei after segmentation. The classifier was constructed based on intensity, texture, and morphology features. However, these machine learning methods manually extract features, and their representation ability and stability can still be affected by subjective factors to some extent.

Generally, a deep-learning-based nuclei classification model consists of two main phases. Firstly, each nucleus is segmented or detected using a deep model; then, those features are fed into a classifier to confirm nuclei types. For instance, in [34], the nuclei in colon cancer histology images were firstly detected using a spatially constrained CNN.

Then, each nucleus with associated patches was fed into the convolution network to predict its type, i.e., inflammatory, healthy, or malignant epithelium.

### 3. Proposed Method

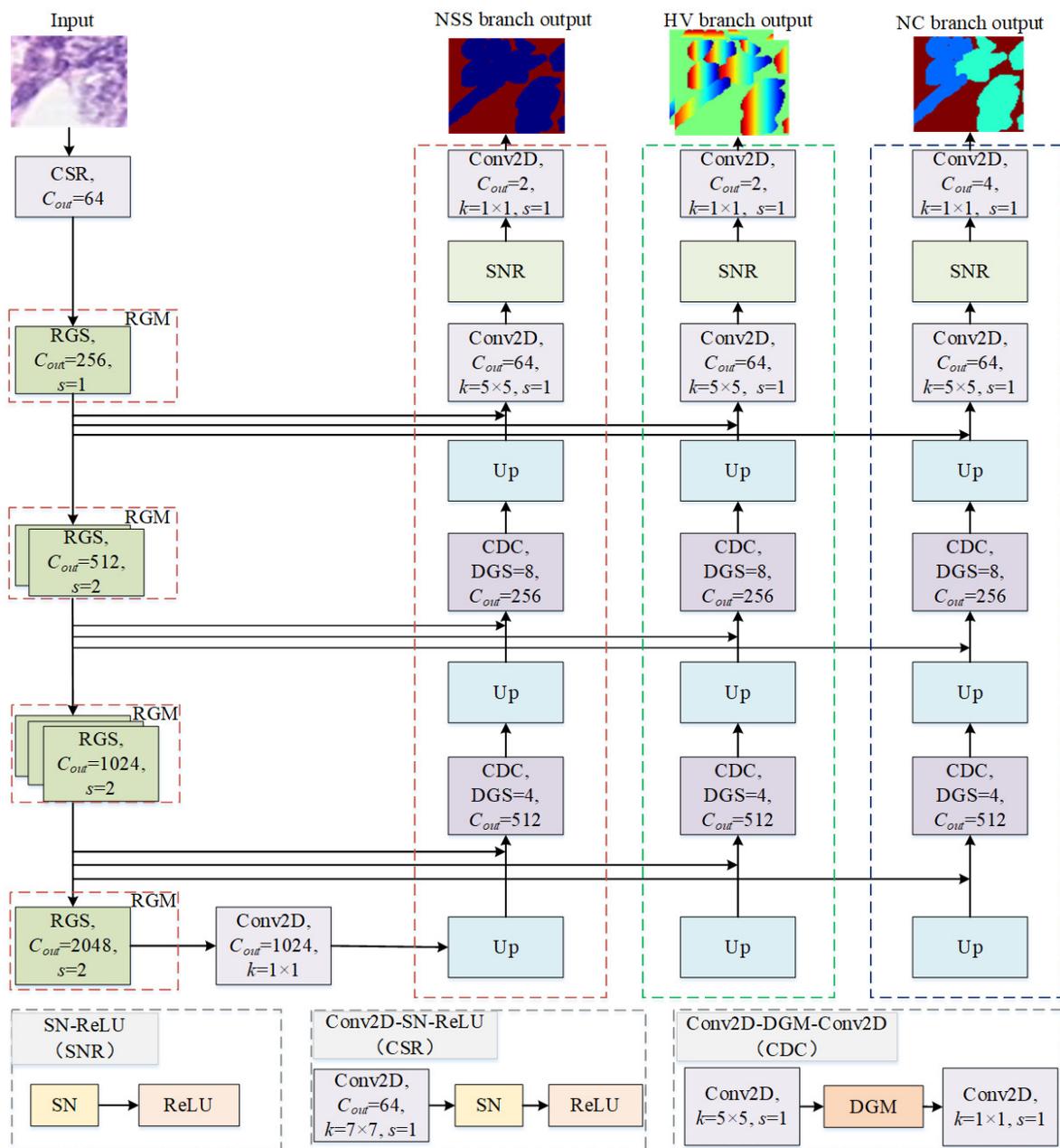
Figure 1 shows an overview of the GSN-HVNET for simultaneous nuclei instance segmentation and classification. The network input starts with  $80 \times 80 \times 3$  images, which are center patches cropped out from the sample images of size  $270 \times 270 \times 3$ . The model can simultaneously segment the nuclei and predict nuclei types and HV-Maps (horizontal and vertical maps). After a post-processing procedure, the nuclei instance can be obtained using HV-Map and nuclei pixel predictions. The final output results can be obtained by combining the segmentation results with the nuclei-type predictions. In other words, the network can complete the segmentation and classification of nuclei instances at the last step.



**Figure 1.** An overview of the GSN-HVNET for simultaneous nuclei instance segmentation and classification. The NSS branch achieves nuclei semantic segmentation, and the HV branch predicts the HV distances of nuclei pixels to their mass centers. Nuclei types are predicted in the NC branch. The nuclei instance segmentation can be accomplished by combining the output of the NSS and HV branches.

#### 3.1. Network Architecture

Figure 2 illustrates the detailed structure of the proposed GSN-HVNET. The proposed network consists of an encoder and a decoder for automatic segmentation and classification of nuclei instances. The encoder can extract an effective set of features; then, the output result of the encoder is used as the decoder input. The decoder contains three branches. Branch I (NSS) is used in nuclei semantic segmentation, and branch II (HV) predicts the HV distances of nuclei pixels to their mass centers. Nuclei types are predicted in branch III (NC). We combine the output of branch I and branch II to accomplish the instance segmentation. Then, the instance segmentation result combines the branch III output to accomplish automatic segmentation and classification of the nuclei instance. The encoder employs the proposed RGS, as discussed in Section 3.1.1. The details of GBS and RGS will be introduced in Sections 3.1.2 and 3.1.3, respectively. In Section 3.1.4, the decoder designed with DGS will be described. The details of DGS will be presented in Section 3.1.5.



**Figure 2.** The structure of GSN-HVNET. Our proposed network contains an encoder and a decoder. The encoder, which extracts an effective set of features, is composed of a CSR block, four RGMs, and a Conv2D. The decoder is composed of three branches to achieve accurate nuclei instance segmentation and classification simultaneously.

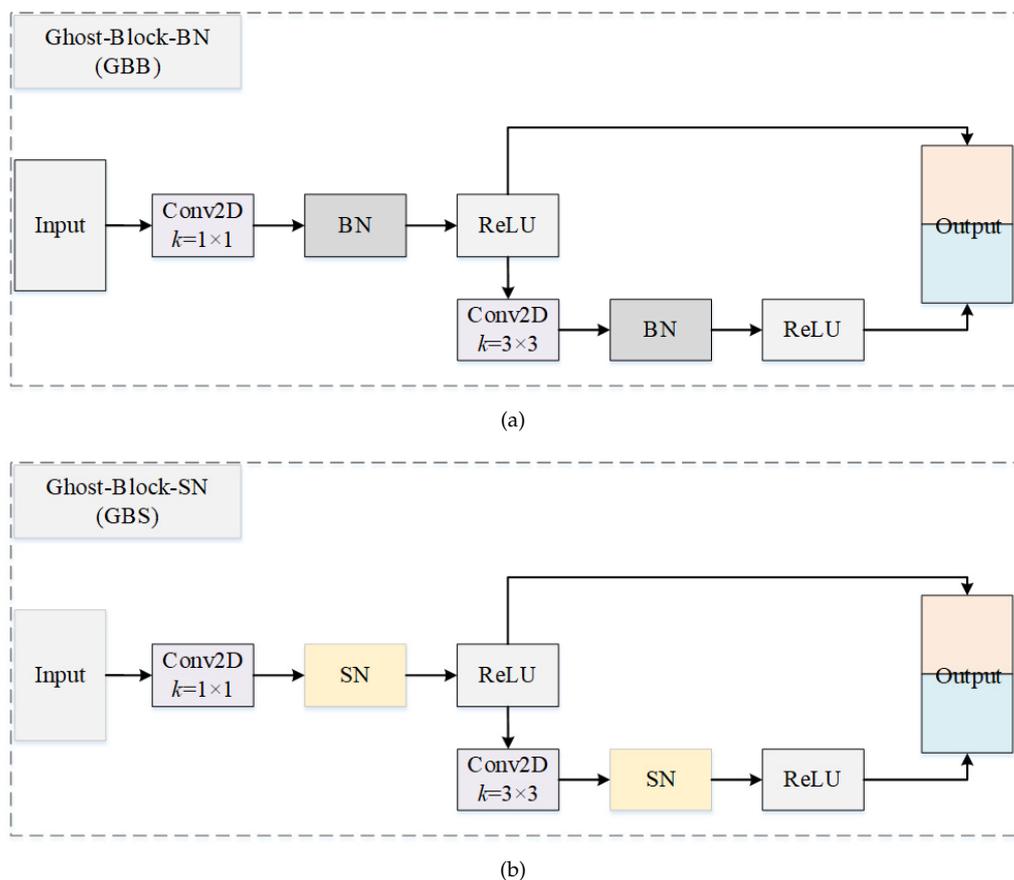
### 3.1.1. Encoder

To extract a practical set of features, we design a novel residual ghost network as part of the encoder in the overall network. The network employs a Conv2D-SN-ReLU (CSR) and a series of 4 Residual-Ghost-Modules (RGMs) for down-sampling. Here, the CSR block is composed of a Conv2D, SN, and ReLU. An RGM consists of multiple instances of our improved Ghost-Block—Residual-Ghost-Block with switchable normalization (RGS) [35]. Benefiting from ghost convolution, our network requires much fewer parameters to generate abundant feature maps compared with using ordinary convolution, resulting in an improved computational efficiency of our encoder. Moreover, the SN can select an optimal combination of different normalizers for different normalization layers, improving the

network stability, i.e., the accuracy is not affected by the batch size. Each RGM is used as a down-sampling level of 2, which means that the spatial resolution of the input is reduced by a factor of 2. We will give a detailed discussion on RGS in the two subsequent subsections.

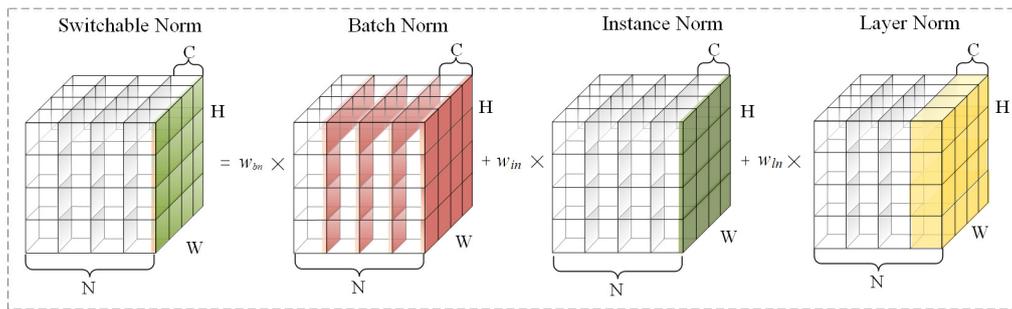
### 3.1.2. Ghost Block with Switchable Normalization

Figure 3 compares the structure of Ghost-Block-BN (GBB) [36] and our suggested Ghost-Block-SN (GBS). As is known, Ghost-Block can help a convolutional neural network to generate more features at a much lower cost. To do that, a Ghost-Block first generates several intrinsic feature maps using ordinary convolution operation and then uses cheap linear operations to expand the features and increase the channels. The computational cost of the linear operations on feature maps is much lower than traditional convolution and transcends other existing efficient works. We can customize the kernel size of the primary convolution in a Ghost-Block, and  $1 \times 1$  point-wise convolution is employed in this paper. In the Residual-Ghost-Block (RGB), each Ghost-Block is followed by a BN layer, which offers several advantages, including stabilizing and speeding up the training procedure. However, the performance of GBB is severely restricted by the batch size. This is because BN only utilizes a single normalizer in the entire network, which can be unstable and hurt the accuracy in the case of a small batch size.



**Figure 3.** An illustration of the ghost block and the improved ghost block with switchable normalization. (a) Ghost block with batch normalization. (b) Ghost block with switchable normalization.

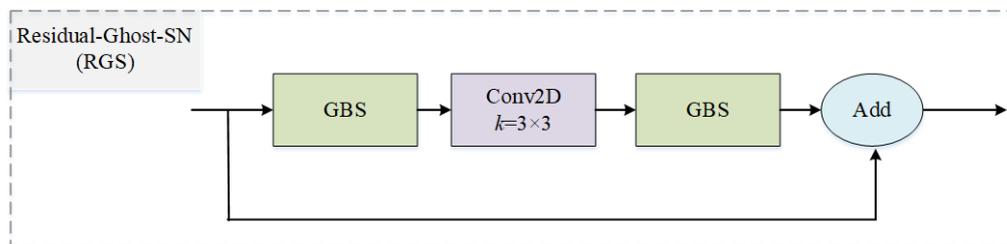
To solve the above problem, we apply switchable normalization (SN), which is robust to a wide range of batch sizes, whether a small batch size or not. As shown in Figure 4, SN measures channel-wise, layer-wise, and minibatch-wise statistics by using instance normalization (IN) [37], layer normalization (LN) [38], and batch normalization (BN) [39], respectively, and tries to find an optimal combination by learning their important weights, ensuring the stability and accuracy of the network in the case of small batch size.



**Figure 4.** Switchable normalization. It learns to select different normalizers for different normalization layers of a deep neural network.

### 3.1.3. Residual Ghost Block with Switchable Normalization

Our RGS adopts the structure of residual block—the essential building unit of residual neural network (ResNet) [40]—owing to its outstanding performance. The key idea behind residual block is to reformulate the layers as learning residual functions with reference to the layer inputs, instead of learning unreferenced functions. As shown in Figure 5, we embed the proposed GBS in a residual block as RGS. Later, several RGSs are stacked to form the RGM. Our network contains of four stacked RGMs with 1, 2, 3, and 1 RGS, respectively. Compared with original ResNet-50, our network employs fewer building units to extract feature maps and reduce redundant features, leading to a reduction in model size. In addition, our proposed RGS is generic and can be used in the construction of other lightweight deep learning architectures.



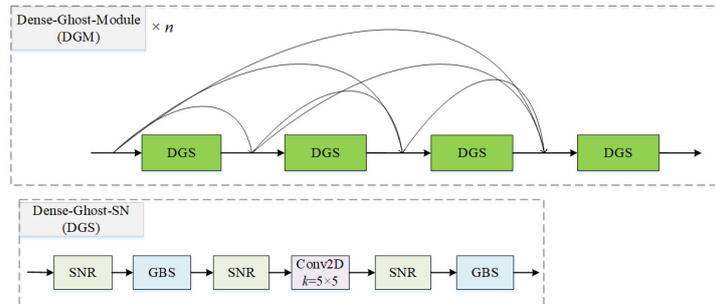
**Figure 5.** Residual ghost block with switchable normalization. The GBS denotes the ghost block with switchable normalization.

### 3.1.4. Decoder

As aforementioned, the decoder contains three branches to obtain accurate nuclei instance segmentation and classification simultaneously. These three branches adopt the same architecture consisting of a series of up-sampling operations and two Dense-Ghost-Modules. A DGM contains a series of cascading DGSs. Through stacking multiple DGSs, we can enrich the receptive field with relatively fewer parameters compared with the most popular Dense-Block, resulting in increased computational efficiency. As is known, low-level information is critical in segmentation tasks because it precisely helps to determine object boundaries. To make use of it, we adopt the skip connections to merge feature from each RGS in the encoder via the concatenation operation. The DGM follows the first and second up-sampling operations. There are eight and four DGSs in the first and second DGM, respectively. Each of the three branches contains three up-sampling steps, making the output feature the same dimension as the input image, i.e.,  $80 \times 80 \times 3$ . By combining the results of the two up-sampling branches, NSS and HV, we can obtain accurate boundaries of each individual cell nucleus, and thereby accomplish the nuclei instance segmentation. Compared with independent networks for different tasks, the proposed network is a unified model to simultaneously accomplish nuclei segmentation and classification, thus reducing the total training time.

### 3.1.5. Dense Ghost Module with Switchable Normalization

In this part, we propose a novel module applied to the decoder of GSN-HVNET. An example of the proposed module is shown in Figure 6, in which  $n = 4$ . Each DGS connects to other DGSs with forwarding feedback and employs GBS to extract feature maps. The feature maps from all preceding layers are utilized as current inputs, and the feature maps output by a DGS are used as inputs for all subsequent layers.



**Figure 6.** Dense ghost module with switchable normalization. The GBS and SNR denote the ghost block with switchable normalization and switchable normalization with ReLU, respectively.

Thus, the proposed module can retain more abundant features as inputs of subsequent layers.

Similarly, benefiting from the lightweight nature of GBS, our proposed DGS utilizes fewer parameters to generate abundant feature maps and valid features compared with Dense-Block [41]. Moreover, it helps to avoid unnecessary calculations by reducing redundant feature maps. Particularly, the DGM can maintain its performance under a small mini-batch size.

### 3.1.6. Joint Loss Function of GSN-HVNET

We design different loss functions for each different task. In Table 1, we define the notations for our works. The joint loss function  $L_{Join}$  is defined by

$$L_{Join} = L_{NSS} + L_{HV} + L_{NC}. \tag{1}$$

The NSS branch corresponds to a semantic segmentation task, and its loss function is designed using BCE loss and dice loss. It is defined by

$$L_{NSS} = \lambda_a L_{BCE} + \lambda_b L_{DICE}, \tag{2}$$

where  $L_{BCE}$  and  $L_{DICE}$  represent the binary cross-entropy loss function and dice loss function for the output of the NSS branch, respectively. The  $\lambda_a$  and  $\lambda_b$  are scalars that give weights to their associated loss function. The above two functions are defined by

$$L_{BCE} = -\left[\frac{1}{n} \sum_{i=1}^N \sum_{k=1}^K X_{i,k}(I) \log Y_{i,k}(I) + \sum_{i=1}^N \sum_{k=1}^K (1 - X_{i,k}(I)) \log(1 - Y_{i,k}(I))\right] \tag{3}$$

and

$$L_{DICE} = 1 - \frac{2 \times \sum_{i=1}^N (Y_i(I) \times X_i(I)) + \epsilon}{\sum_{i=1}^N Y_i(I) + \sum_{i=1}^N X_i(I) + \epsilon}, \tag{4}$$

where  $X$  represents the ground truth,  $Y$  denotes the prediction, and  $K$  represents the number of categories. In order to avoid zero denominators, we set  $\epsilon$  to  $1.0 \times e^{-4}$  in the numerator and denominator.

The loss function for the HV branch is defined by

$$L_{HV} = \lambda_c L_{MSE} + \lambda_d L_{MSGEL}, \tag{5}$$

where  $L_{MSE}$  represents the mean squared error measuring the difference between the HV distances prediction and the ground truth,  $\lambda_c$  and  $\lambda_d$  are the weights of their associated loss function. The loss function  $L_{MSGE}$  is used to calculate the gradients of the mean squared error between HV maps and ground truth.  $L_{MSE}$  and  $L_{MSGE}$  are defined by

$$L_{MSE} = \frac{1}{n} \sum_{n=1}^n (p_i(I) - \Gamma_i(I))^2 \tag{6}$$

and

$$L_{MSGE} = \frac{1}{m} \sum_{i \in M} (\nabla_x(p_{i,x}(I) - \Gamma_{i,x}(I)))^2 + \frac{1}{m} \sum_{i \in M} (\nabla_y(p_{i,y}(I) - \Gamma_{i,y}(I)))^2, \tag{7}$$

where  $I$  represents the input image and  $p_i(I)$  is defined as the regression output of HV branch. The pixel-wise softmax predictions of NSS and NC branches are represented by  $q_i(I)$  and  $r_i(I)$ , respectively.  $\Gamma_i(I)$  denotes the ground truth of the HV distance of nuclei pixels to their mass centers.

The loss function of  $L_{NC}$  is defined by

$$L_{NC} = \lambda_e L_{BCE} + \lambda_f L_{DICE}. \tag{8}$$

Similarly,  $\lambda_e$  and  $\lambda_f$  are used to balance the two loss functions  $L_{BCE}$  and  $L_{DICE}$ .

**Table 1.** The definition of notations.

Notation	Definition
$h_{ncij}$	The value of a pixel before normalization.
$\hat{h}_{ncij}$	The value of a pixel after normalization.
$\gamma, \beta$	Scale and shift parameter
$I_k,  I_k $	A set of pixels, and the number of pixels in $I_k$ .
$L, \lambda$	$L$ denotes the loss function and $\lambda$ represents its parameters.
$I$	The input image.
$\Gamma_i(I)$	The HV distance of nuclei pixels to their mass centers.
$p_i(I)$	The regression output of HV branch.
$q_i(I)$	The pixel-wise and softmax predictions of NSS branch.
$r_i(I)$	The pixel-wise and softmax predictions of NC branch.
$E$	The energy landscape.
$F_c^t$	The whole measurement for nuclei type classification and nuclei instance segmentation.
$FP, FN$	False-positive, false-negative.
$TP, TN$	True-positive, true-negative.

### 3.2. Post-Processing

The proposed network produces three outputs. To obtain the nuclei location and separate overlapping or clustered nuclei, we need to post-process the output of NSS and HV. Within each HV map, there are significant differences between pixels in adjacent instances. Using this property, we can calculate the gradient so as to separate the clustered nuclei. To do that, we have

$$S_m = \max(Hor(p_{hor}), Ver(p_{ver})), \tag{9}$$

where  $p_{hor}$  and  $p_{ver}$  represent the horizontal and vertical predictions produced by the HV branch, and  $Hor$  and  $Ver$  refer to the horizontal and vertical components, respectively, of the Sobel operator, which calculates the horizontal and vertical derivative approximations.

In Figure 1,  $S_m$  highlights the regions where pixels in adjacent regions of two instances differ significantly in the horizontal and vertical maps.

We compute the marker  $M$  according to

$$M = \sigma(\tau(q, h) - \tau(S_m, k)), \quad (10)$$

where  $q$  is the output probability map of the NSS branch and  $\tau(q, h)$  is a threshold function acting on  $q$  and sets values above  $h$  to 1 or 0; otherwise,  $\sigma$  is a rectifier setting all negative value to 0 and  $M$  is the output marker. We can obtain desired segmentation results by choosing appropriate  $h$  and  $k$ .

Next, we compute the energy landscape  $E$  according to

$$E = [1 - \tau(S_m, k)] * \tau(q, h). \quad (11)$$

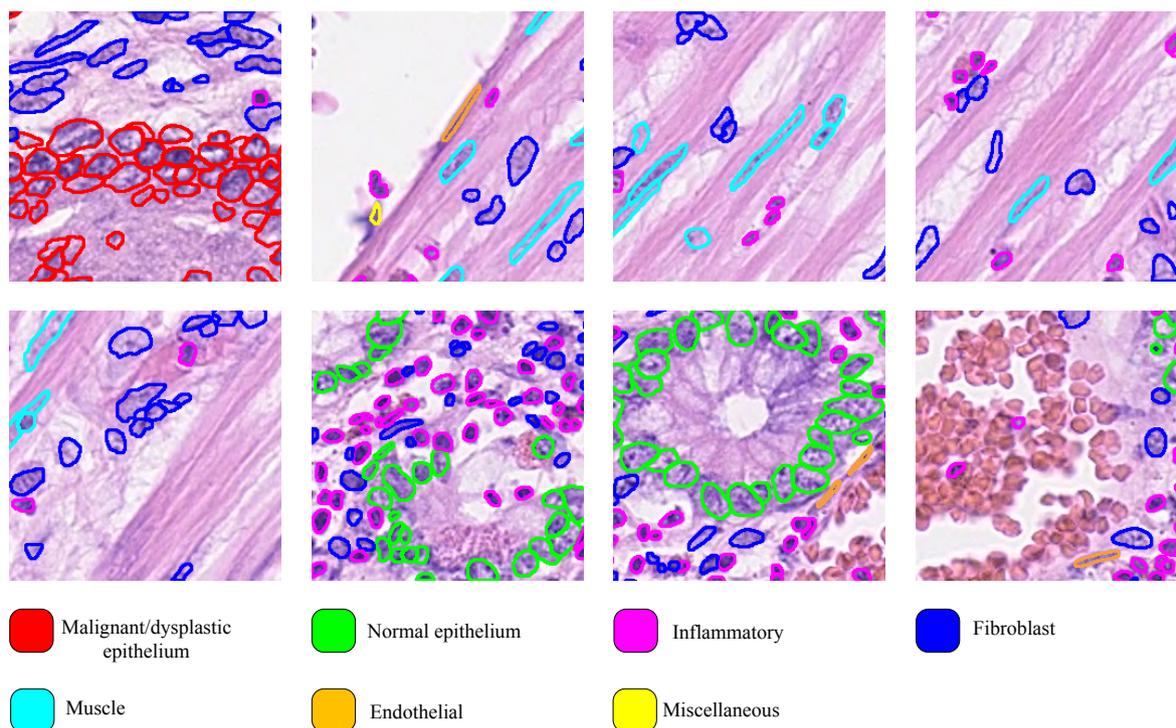
Finally, given the energy landscape  $E$ , a marker-controlled watershed is carried out using  $M$  as the marker to determine how to split  $\tau(q, h)$ , given the energy landscape  $E$ . The task of joint segmentation and classification of nuclei requires converting per-pixel nuclei type prediction in the NSS branch to the prediction of the type of nuclei instances. To do that, we combine the post-processing result with NC branch output.

## 4. Experiment

### 4.1. Datasets and Implementation

In our experiment, we adopt three authoritative nuclei datasets: CoNSeP [13], Kumar [42], and CPM-17 [43]. Table 2 describes these datasets used in our experiment. The CoNSeP dataset, extracted from 16 colorectal adenocarcinoma (CRA) WSIs, consists of 41 hematoxylineosin (H&E) staining images, each of size  $1000 \times 1000$  at  $40 \times$  objective magnification. In CoNSeP dataset, tumor regions, stroma, muscular, fat, glandular, and collagen can be observed. In addition to containing different tissue components, seven nuclei types are provided, including malignant/dysplastic epithelial nuclei, normal epithelium, inflammatory, fibroblast, muscle, endothelial, and miscellaneous. In [13], the authors combined the original seven categories into four categories, of which malignant/dysplastic epithelial and normal epithelial were combined into a single type corresponding to the epithelial class, and fibroblast, muscle, and endothelial were combined into a single type corresponding to the spindle-shaped class. However, in practical clinical diagnosis, a CAD system should mainly focus on the identification of lesion area. To address this issue, we reclassified this dataset in our experiment. Specially, the normal epithelium, fibroblast, muscle, endothelial, and miscellaneous were combined into a single type corresponding to normal region, and the malignant/dysplastic epithelial and inflammatory are considered as two separate types—that is, the reclassified contain three nuclei categories as well as the background category. With this classification rule, our model can directly report the types of nuclei in lesion areas.

Kumar is an annotated dataset containing over 13,000 segmented nuclei from four different organs—breast, kidney, liver, and prostate—of 16 patients. The CPM-17 dataset provides the tissue image with labels for nuclei segmentation and classification. It is obtained from patients with head and neck squamous cell (HNSCC), glioblastoma multiforme (GBM), non-small cell lung cancer (NSCLC), and lower-grade glioma tumors (LGG). Some examples taken from these datasets are shown in Figure 7.



**Figure 7.** The sample clipping region is extracted from the CoNSeP dataset, where the color of each nuclear boundary indicates its category.

**Table 2.** Description of the dataset used in our experiment. The Seg denotes the dataset with segmentation labels and the Class denotes the dataset with classification labels.

	CoNSeP	CPM-17	Kumar
Total numbers of nuclei	24,319	7570	21,623
Labeled nuclei	24,319	0	0
Number of images	41	32	30
Origin	UHCW	TCGA	TCGA
Magnification	40×	40× & 20×	40×
Size of images	1000×1000	500×500 to 600×600	1000×1000
Seg/Class	Seg&Class	Seg	Seg
Number of cancer types	1	4	8

We run our code on a server equipped with an NVIDIA Geforce RTX 3090 GPU and Intel(R) Xeon(R) Gold 5118 CPU. During the training phase, we performed data augmentation to augment the training data. We randomly combined zooming, channel shifting, shearing, rotating, and horizontal/vertical flipping, which cropped the original image into  $270 \times 270$  sub-images. We used Kaiming normalization [44] to initialize weights and set initial bias as false. We used Adam [45] as the optimizer, with a trainable batch size of 4. We set an initial learning rate as  $1.0 \times e^{-4}$  and weight decay as 0.1. The six hyper-parameters  $\lambda_a, \lambda_b, \lambda_c, \lambda_d, \lambda_e,$  and  $\lambda_f$  used for balancing the joint loss function are tuned to be  $\{1, 1, 1, 1, 2, 1\}$  on the validation set.

#### 4.2. Evaluation Metrics

##### 4.2.1. Nuclei Instance Segmentation Evaluation

The segmentation of the nuclei instances can be divided into three sub-tasks; these three sub-tasks are the separation of the nuclei from the background, the detection of individual nuclei instances, and the segmentation of each detected instance. The Ensemble

Dice [43] and Aggregated Jaccard Index [42] are two popular metrics used to measure the performance of nuclei instance segmentation. To better investigate the proposed method, we need to measure the performance of each sub-task. The dice coefficient ( $F_1$  score) is defined by

$$Dice\_coef = \frac{|TP|}{|TP| + \frac{1}{2}|FP| + \frac{1}{2}|FN|} = \frac{2 \times (X \cap Y)}{(|X| + |Y|)}, \tag{12}$$

where  $TP$  represents the true-positive rate,  $FP$  represents the false-positive rate, and  $FN$  represents the false-negative rate.  $X$  and  $Y$  represent the ground truth and prediction, respectively.

The  $AJI$  calculates the ratio of an aggregated intersection cardinality to an aggregated union cardinality between the ground truth and prediction. It is defined by

$$AJI = \frac{\sum_{i=1}^N |G_i \cap P_M^i|}{\sum_{i=1}^N |G_i \cup P_M^i| + \sum_{F \in U} |P_F|}, \tag{13}$$

where  $G_i$  is the  $i$ th nucleus from the ground truth with  $N$  nuclei.  $P_M^i$  represents the  $M$ th connected component in prediction, which has the largest Jaccard Index with  $G_i$ , and where each  $M$  cannot be utilized more than once.  $U$  is a set representing the connected component in the prediction without the corresponding ground truth.

Unfortunately,  $F_1$  score and  $AJI$  only calculate an overall score for the instance segmentation quality. In addition, the two metrics suffer from a limitation that they will produce excessive penalization and result in an abnormal score for overlapping regions.

To take a measurement of each sub-task, we take advantage of panoptic quality [46] with accurate quantification and interpretability to measure the performance of nuclei instance segmentation. The panoptic quality for nuclei instance segmentation is defined by

$$PQ = DQ \times SQ = Dice\_coef \times \frac{\sum_{(x,y) \in TP} IoU(x,y)}{|TP|}, \tag{14}$$

where  $x$  and  $y$  denote a ground truth component and a prediction component, respectively. The  $IoU$  represents the intersection over union. Each  $(x, y)$  must be unique over the whole set of prediction and ground truth segments, if their  $IoU(x, y) > 0.5$ .  $DQ$  and  $SQ$  help to give a direct insight into detecting individual nuclear instances and segmenting each detected instance. Therefore,  $PQ$  can serve as the objective evaluation criteria for measuring the performance of the nuclei instance segmentation task.

To demonstrate the effectiveness of the proposed method, we use the following three metrics. Dice coefficient and  $PQ$  are used to measure the separation of all nuclei from the background and serve as a unified score for comparison, respectively. The  $AJI$  is used for the comparison with other methods. In this study, these three metrics serve as objective evaluation criteria. As the most reliable assessment of the segmentation quality, the subjective evaluation can also be carried out in practical applications.

#### 4.2.2. Nuclei Classification Evaluation

Nuclei classification is influenced by nuclei instance segmentation. The whole measurement for nuclei type classification should include nuclei instance segmentation. HoVer-Net [13] defines an efficient evaluation, which is defined by

$$F_c^t = \frac{2(TP_c + TN_c)}{2(TP_c + TN_c) + 2(FP_c + FN_c) + (FP_d + FN_d)}, \tag{15}$$

where  $FP_d$  and  $FN_d$  are false-positive and false-negative in detecting ground truth instances, respectively.  $TP_c$ ,  $TN_c$ ,  $FP_c$ , and  $FN_c$  denote true-positive, true-negative, false-positive, and false-negative, respectively.

### 4.3. Experimental Results

Table 3 compares the number of trainable parameters of the proposed and other popular models. As can be seen from this table, our model gives the smallest size among all others in terms of the nuclei segmentation task and the second smallest size in term of the joint segmentation and classification task. Consequently, our model offers a high degree of computational efficiency. Table 4 compares the Dice scores of the proposed and two state-of-the-art models working with small mini-batch sizes. The results indicate that the proposed model appears more stable, i.e., our model can work well on small-memory-capacity GPUs, such as the NVIDIA 1080ti or 2080ti, thus reducing the hardware cost. The model size is significantly smaller than other compared networks.

**Table 3.** Comparative results for the number of trainable parameters of different networks for nuclei segmentation and classification. The Seg denotes the single-task network for segmentation. The Seg&Class denotes the multi-tasking network for simultaneous segmentation and classification.

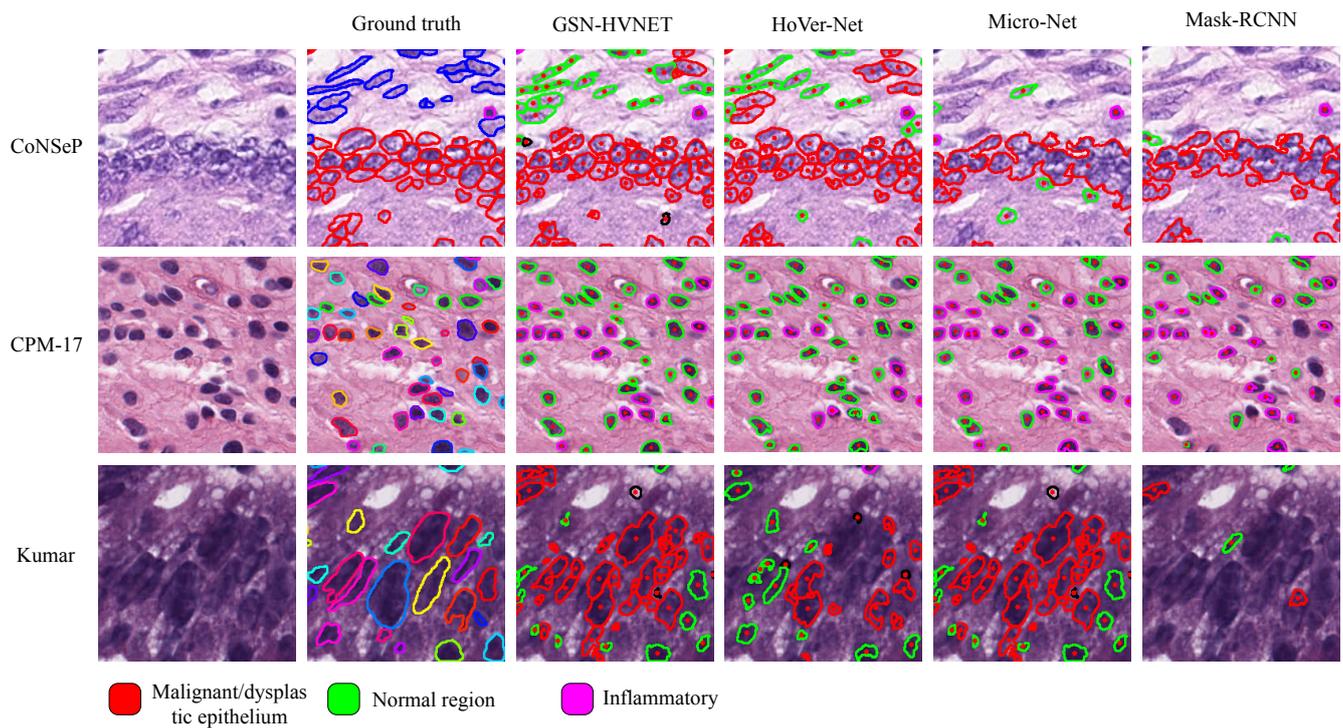
Method	Seg/Class	Parameters
HoVer-Net [13]	Seg	42.94M
HoVer-Net [13]	Seg&Class	52.20M
Micro-Net [14]	Seg&Class	183.67M
DIST [30]	Seg&Class	8.81M
DCAN [47]	Seg	39.54M
SegNet [48]	Seg	28.07M
FCN8 [49]	Seg	128.05M
U-Net [28]	Seg&Class	35.23M
Mask-RCNN [15]	Seg&Class	44.17M
Our proposed	Seg	15.03M
Our proposed	Seg&Class	32.52M

**Table 4.** Comparative results for different mini-batch sizes presenting in three multi-tasking networks. The Dice coefficient is used to evaluate the segmentation performance on the CoNSeP, Kumar, and CPM-17 datasets.

Batch Size	Our Proposed			HoVer-Net			Micro-Net		
	Dice			Dice			Dice		
	CoNSeP	Kumar	CPM-17	CoNSeP	Kumar	CPM-17	CoNSeP	Kumar	CPM-17
1	0.821	0.851	0.865	0.816	0.794	0.843	0.752	0.759	0.828
2	0.830	0.844	0.870	0.806	0.804	0.875	0.764	0.785	0.857
3	0.839	0.842	0.870	0.835	0.819	0.879	0.758	0.794	0.859

The proposed network is measured by the three kinds of metrics discussed above, compared with baselines and other state-of-the-art networks, and the results are reported in Table 5. The results indicate that our proposed network achieves the highest accuracy among all the others. Moreover, even though the DIST model has fewer parameters than ours on joint nuclei segmentation classification task, its segmentation performance is worse than ours by a large margin on all three datasets. Therefore, our network offers an optimal trade-off between accuracy and efficiency.

As aforementioned, the 4-class nuclei classification carried out in HoVer-Net is impractical for use in practical pathological diagnosis. Accordingly, we have reclassified the data. Table 6 lists the comparative results for 3-class nuclei classification on the CoNSeP dataset. Here,  $F_d$  denotes the  $F_1$  score for nuclei detection.  $F_c^1$ ,  $F_c^2$ , and  $F_c^3$  denote the classification score for healthy, inflammatory, and malignant/dysplastic epithelium classes, respectively. The results show that the proposed network outperforms all the others in terms of  $F_c^1$ ,  $F_c^2$ , and  $F_c^3$  scores. In Figure 8, we illustrate the results of nuclei segmentation and classification on the sample images and compare them with those of [13–15,30].



**Figure 8.** Comparative results for nuclei classification and segmentation. The normal epithelium, fibroblast, muscle, endothelial, and miscellaneous are combined into a single type corresponding to the normal region, and the malignant/dysplastic epithelial and inflammatory are considered as two separate types.

As can be seen from this figure, our lightweight method is successful in segmenting overlapping and clustered nuclei. It is also excellent to complete the task of nuclei classification at the same time. Overall, our proposed model achieves state-of-the-art accuracy on nuclei segmentation and classification tasks while maintaining low computation cost. Our idea can be directly deployed in the cell pathology diagnosis system to reduce the workload of pathologists.

**Table 5.** Comparative results for nuclei segmentation. The Dice coefficient, AJI, and PQ are used to evaluate the instance segmentation performance of ten networks on the CoNSeP, Kumar, and CPM-17 datasets.

Method	CoNSeP			Kumar			CPM-17		
	Dice	AJI	PQ	Dice	AJI	PQ	Dice	AJI	PQ
HoVer-Net [13]	0.838	0.525	0.494	0.826	0.618	0.597	0.869	0.705	0.697
SegNet [48]	0.796	0.194	0.270	0.811	0.377	0.407	0.857	0.491	0.531
FCN8 [49]	0.756	0.123	0.163	0.797	0.281	0.312	0.840	0.397	0.435
U-Net [28]	0.724	0.482	0.328	0.758	0.556	0.478	0.813	0.643	0.578
DIST [30]	0.798	0.495	0.386	0.789	0.559	0.443	0.826	0.616	0.504
DCAN [47]	0.733	0.289	0.256	0.792	0.525	0.492	0.828	0.561	0.545
Micro-Net [14]	0.784	0.518	0.421	0.797	0.560	0.519	0.857	0.668	0.661
Mask-RCNN [15]	0.740	0.474	0.460	0.760	0.546	0.509	0.850	0.684	0.674
CIA-Net [31]	-	-	-	0.818	0.620	0.577	-	-	-
Our proposed	0.861	0.602	0.566	0.879	0.635	0.644	0.899	0.701	0.683

**Table 6.** Comparative results for 3-class nuclei classification on the CoNSeP dataset.  $F_d$  denotes the  $F_1$  score for nuclei detection.  $F_c^1$ ,  $F_c^2$ , and  $F_c^3$  denote the classification score for healthy, inflammatory, and malignant/dysplastic epithelium classes, respectively.

Method	$F_d$	$F_c^1$	$F_c^2$	$F_c^3$
HoVer-Net [13]	0.784	0.488	0.525	0.517
Micro-Net [14]	0.812	0.487	0.549	0.546
DIST [30]	0.782	0.489	0.569	0.526
Mask-RCNN [15]	0.701	0.413	0.568	0.514
Our proposed	0.820	0.514	0.572	0.519

## 5. Conclusions

In this paper, we designed a lightweight, multi-task deep learning framework for nuclei segmentation and classification. Our model follows an encoder–decoder architecture, and the decoder consists of three branches, each outputting a prediction for a sub-task. To sufficiently use the correlation among the three branches, we employ NSS and HV branches to complete the nuclei instance segmentation and use NC branch to predict the classes of each nucleus in a learning process. Two newly designed blocks, Residual-Ghost-SN and Dense-Ghost-SN, are employed in the encoder and decoder parts, respectively, to reduce the computational cost and improve the network stability under small batch sizes. Extensive experiments have been carried out on the CoNSeP, Kumar, and CPM-17 datasets, and the results demonstrate that our model offers a state-of-the-art trade-off between computational efficiency and both segmentation and classification accuracy.

Ultimately, our idea is generic, and can be easily deployed to other histopathology images analysis works. Moreover, the blocks proposed in this paper, including Residual-Ghost-SN and Dense-Ghost-SN, are also generic and can be flexibly embedded into other deep CNNs for histopathology image diagnostic tasks. However, regarding their application in the field of natural images, we have not conducted experiments, and the effects cannot be guaranteed. Thus, we pose this as an open problem and expect to provide a theoretical analysis with complete proof in further research.

**Author Contributions:** Conceptualization, T.Z. and C.F.; methodology, T.Z. and C.F.; validation, T.Z. and Y.T.; formal analysis, T.Z., W.S. and C.-W.S.; investigation, T.Z. and C.F.; resources, T.Z. and W.S.; data curation, T.Z. and C.F.; writing—original draft preparation, T.Z. and Y.T.; writing—review and editing, T.Z., C.F., Y.T., W.S. and C.-W.S.; visualization, T.Z. and Y.T.; supervision, T.Z.; project administration, T.Z.; funding acquisition, C.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is supported by the National Natural Science Foundation of China (No. 62032013), the Fundamental Research Funds for the Central Universities (Nos. N2224001-7 and N2116020), and the Natural Science Foundation of Liaoning Province (No. 2021-YGJC-24).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Pantanowitz, L. Digital images and the future of digital pathology. *J. Pathol. Inform.* **2010**, *1*, 15. [\[CrossRef\]](#)
2. Alsubaie, N.; Sirinukunwattana, K.; Raza, S.E.A.; Snead, D.; Rajpoot, N. A bottom-up approach for tumour differentiation in whole slide images of lung adenocarcinoma. In *Medical Imaging 2018: Digital Pathology*; Tomaszewski, J.E., Gurcan, M.N., Eds.; International Society for Optics and Photonics, SPIE: Houston, TX, USA, 2018; Volume 10581, p. 105810E. [\[CrossRef\]](#)

3. Lu, C.; Romo-Bucheli, D.; Wang, X.; Janowczyk, A.; Ganesan, S.; Gilmore, H.; Rimm, D.; Madabhushi, A. Nuclear shape and orientation features from H&E images predict survival in early-stage estrogen receptor-positive breast cancers. *Lab. Investig.* **2018**, *98*, 1438–1448. [[CrossRef](#)]
4. Javed, S.; Fraz, M.M.; Epstein, D.; Snead, D.; Rajpoot, N.M. Cellular community detection for tissue phenotyping in histology images. In *Computational Pathology and Ophthalmic Medical Image Analysis*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 120–129. [[CrossRef](#)]
5. Sirinukunwattana, K.; Snead, D.; Epstein, D.; Aftab, Z.; Mujeeb, I.; Tsang, Y.W.; Cree, I.; Rajpoot, N. Novel digital signatures of tissue phenotypes for predicting distant metastasis in colorectal cancer. *Sci. Rep.* **2018**, *8*, 13692. [[CrossRef](#)] [[PubMed](#)]
6. Corredor, G.; Wang, X.; Zhou, Y.; Lu, C.; Fu, P.; Syrigos, K.; Rimm, D.L.; Yang, M.; Romero, E.; Schalper, K.A.; et al. Spatial architecture and arrangement of tumor-infiltrating lymphocytes for predicting likelihood of recurrence in early-stage non-small cell lung cancer. *Clin. Cancer Res.* **2019**, *25*, 1526–1534. [[CrossRef](#)]
7. Sharma, H.; Zerbe, N.; Heim, D.; Wienert, S.; Behrens, H.M.; Hellwich, O.; Hufnagl, P. A multi-resolution approach for combining visual information using nuclei segmentation and classification in histopathological images. In Proceedings of the VISAPP, Berlin, Germany, 11–14 March 2015; Volume 3, pp. 37–46.
8. Wang, P.; Hu, X.; Li, Y.; Liu, Q.; Zhu, X. Automatic cell nuclei segmentation and classification of breast cancer histopathology images. *Signal Process.* **2016**, *122*, 1–13. [[CrossRef](#)]
9. Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448. [[CrossRef](#)]
10. Kendall, A.; Gal, Y.; Cipolla, R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7482–7491. [[CrossRef](#)]
11. Zhou, Y.; Chen, H.; Li, Y.; Liu, Q.; Xu, X.; Wang, S.; Yap, P.T.; Shen, D. Multi-task learning for segmentation and classification of tumors in 3D automated breast ultrasound images. *Med. Image Anal.* **2021**, *70*, 101918. [[CrossRef](#)] [[PubMed](#)]
12. Shan, P.; Wang, Y.; Fu, C.; Song, W.; Chen, J. Automatic skin lesion segmentation based on FC-DPN. *Comput. Biol. Med.* **2020**, *123*, 103762. [[CrossRef](#)]
13. Graham, S.; Vu, Q.D.; Raza, S.E.A.; Azam, A.; Tsang, Y.W.; Kwak, J.T.; Rajpoot, N. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Med. Image Anal.* **2019**, *58*, 101563. [[CrossRef](#)]
14. Raza, S.E.A.; Cheung, L.; Shaban, M.; Graham, S.; Epstein, D.; Pelengaris, S.; Khan, M.; Rajpoot, N.M. Micro-Net: A unified model for segmentation of various objects in microscopy images. *Med. Image Anal.* **2019**, *52*, 160–173. [[CrossRef](#)]
15. Abdulla, W. Mask R-CNN for Object Detection and Instance Segmentation on Keras and TensorFlow. 2017. Available online: [https://github.com/matterport/Mask\\_RCNN](https://github.com/matterport/Mask_RCNN) (accessed on 31 January 2023).
16. Song, W.; Zheng, Y.; Fu, C.; Shan, P. A novel batch image encryption algorithm using parallel computing. *Inf. Sci.* **2020**, *518*, 211–224. [[CrossRef](#)]
17. Song, W.; Fu, C.; Zheng, Y.; Cao, L.; Tie, M. A practical medical image cryptosystem with parallel acceleration. *J. Ambient. Intell. Humaniz. Comput.* **2022**, 1–15.
18. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
19. Song, W.; Fu, C.; Zheng, Y.; Cao, L.; Tie, M.; Sham, C.W. Protection of image ROI using chaos-based encryption and DCNN-based object detection. *Neural Comput. Appl.* **2022**, *34*, 5743–5756.
20. Song, W.; Fu, C.; Zheng, Y.; Tie, M.; Liu, J.; Chen, J. A parallel image encryption algorithm using intra bitplane scrambling. *Math. Comput. Simul.* **2023**, *204*, 71–88. [[CrossRef](#)]
21. Vincent, L.; Soille, P. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Comput. Archit. Lett.* **1991**, *13*, 583–598. [[CrossRef](#)]
22. Yang, X.; Li, H.; Zhou, X. Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and Kalman filter in time-lapse microscopy. *IEEE Trans. Circuits Syst. Regul. Pap.* **2006**, *53*, 2405–2414. [[CrossRef](#)]
23. Mohammed, J.G.; Boudier, T. Classified region growing for 3D segmentation of packed nuclei. In Proceedings of the 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), Beijing, China, 29 April–2 May 2014; pp. 842–845.
24. Lu, Z.; Carneiro, G.; Bradley, A.P. An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells. *IEEE Trans. Image Process.* **2015**, *24*, 1261–1272. [[CrossRef](#)]
25. Nguyen, K.; Sarkar, A.; Jain, A.K. Prostate cancer grading: Use of graph cut and spatial arrangement of nuclei. *IEEE Trans. Med. Imaging* **2014**, *33*, 2254–2270. [[CrossRef](#)]
26. Mitra, S.; Dey, S.; Das, N.; Chakrabarty, S.; Nasipuri, M.; Naskar, M.K. Identification of Benign and Malignant Cells from cytological images using superpixel based segmentation approach. In Proceedings of the Annual Convention of the Computer Society of India, Kolkata, India, 19–21 January 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 257–269. [[CrossRef](#)]
27. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
28. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241. [[CrossRef](#)]

29. Graham, S.; Rajpoot, N.M. SAMS-NET: Stain-aware multi-scale network for instance-based nuclei segmentation in histology images. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 590–594. [[CrossRef](#)]
30. Naylor, P.; Laé, M.; Reyal, F.; Walter, T. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Trans. Med. Imaging* **2018**, *38*, 448–459. [[CrossRef](#)]
31. Zhou, Y.; Onder, O.F.; Dou, Q.; Tsougenis, E.; Chen, H.; Heng, P.A. Cia-net: Robust nuclei instance segmentation with contour-aware information aggregation. In Proceedings of the International Conference on Information Processing in Medical Imaging, Hong Kong, China, 2–7 June 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 682–693. [[CrossRef](#)]
32. Jones, T.R.; Carpenter, A.E.; Lamprecht, M.R.; Moffat, J.; Silver, S.J.; Grenier, J.K.; Castoreno, A.B.; Eggert, U.S.; Root, D.E.; Golland, P.; et al. Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 1826–1831. [[CrossRef](#)]
33. Yuan, Y.; Failmezger, H.; Rueda, O.M.; Ali, H.R.; Gräf, S.; Chin, S.F.; Schwarz, R.F.; Curtis, C.; Dunning, M.J.; Bardwell, H.; et al. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Sci. Transl. Med.* **2012**, *4*, 157ra143. [[CrossRef](#)] [[PubMed](#)]
34. Sirinukunwattana, K.; Raza, S.E.A.; Tsang, Y.W.; Snead, D.R.; Cree, I.A.; Rajpoot, N.M. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Trans. Med. Imaging* **2016**, *35*, 1196–1206. [[CrossRef](#)]
35. Luo, P.; Ren, J.; Peng, Z.; Zhang, R.; Li, J. Differentiable Learning-to-Normalize via Switchable Normalization. In Proceedings of the International Conference on Learning Representation (ICLR), Vancouver, BC, Canada, 30 April–3 May 2019.
36. Han, K.; Wang, Y.; Tian, Q.; Guo, J.; Xu, C.; Xu, C. Ghostnet: More features from cheap operations. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 1580–1589. [[CrossRef](#)]
37. Huang, X.; Belongie, S. Arbitrary style transfer in real-time with adaptive instance normalization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 1501–1510. [[CrossRef](#)]
38. Ba, J.L.; Kiros, J.R.; Hinton, G.E. Layer normalization. *arXiv* **2016**, arXiv:1607.06450.
39. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 7–9 July 2015; pp. 448–456.
40. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [[CrossRef](#)]
41. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708. [[CrossRef](#)]
42. Kumar, N.; Verma, R.; Sharma, S.; Bhargava, S.; Vahadane, A.; Sethi, A. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Trans. Med. Imaging* **2017**, *36*, 1550–1560. [[CrossRef](#)]
43. Vu, Q.D.; Graham, S.; Kurc, T.; To, M.N.N.; Shaban, M.; Qaiser, T.; Koohbanani, N.A.; Khurram, S.A.; Kalpathy-Cramer, J.; Zhao, T.; et al. Methods for segmentation and classification of digital microscopy tissue images. *Front. Bioeng. Biotechnol.* **2019**, *7*, 53. [[CrossRef](#)] [[PubMed](#)]
44. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1026–1034. [[CrossRef](#)]
45. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
46. Kirillov, A.; He, K.; Girshick, R.; Rother, C.; Dollár, P. Panoptic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 9404–9413. [[CrossRef](#)]
47. Chen, H.; Qi, X.; Yu, L.; Heng, P.A. DCAN: Deep contour-aware networks for accurate gland segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2487–2496. [[CrossRef](#)]
48. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [[CrossRef](#)]
49. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440. [[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.