

Research Article

Segmentation of Breast Tubules in H&E Images Based on a DKS-DoubleU-Net Model

Yuli Chen ¹, Yao Zhou ¹, Guoping Chen ¹, Yuchuan Guo ², Yanquan Lv ²,
Miao Ma ¹, Zhao Pei ¹ and Zengguo Sun ¹

¹School of Computer Science, Shaanxi Normal University, Xi'an, China

²Shandong Junteng Medical Technology Co., LTD, China

Correspondence should be addressed to Zengguo Sun; sunzg@snnu.edu.cn

Received 5 August 2022; Accepted 10 September 2022; Published 30 September 2022

Academic Editor: Cheng Lu

Copyright © 2022 Yuli Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The formation of breast tubules plays an important role in the pathological grading of breast cancer. Breast tubules surrounded by a large number of epithelial cells are located in the subcutaneous tissue of the chest. The shapes of breast tubules are various, including tubular, round, and oval, which makes the process of breast tubule segmentation a difficult task. Deep learning technology, capable of learning complex data structures via efficient representation, could help pathologists accurately detect breast tubules in hematoxylin and eosin (H&E) stained images. In this paper, we propose a deep learning model named DKS-DoubleU-Net to accurately segment breast tubules with complex appearances in H&E images. The proposed DKS-DoubleU-Net model suggests using a DenseNet module as the encoder of the second subnetwork of DoubleU-Net, which utilizes dense features between layers and strengthens the propagation of features extracted in all previous layers, in order to better discover the intrinsic characteristics of breast tubules with complex structures and diverse shapes. Moreover, a feature fusing module called Kernel Selecting Module (KSM) is inserted before each output layer of the two U-Net branches of the DoubleU-Net, to implement a multiscale feature fusion via a self-adaptive kernel selecting for the sake of accurate segmentation of breast tubules in different sizes. The experiments on the public BRACS dataset and a private clinical dataset have shown that our model achieves better segmentation performance, compared to the state-of-art models of U-Net, DoubleU-Net, ResUnet++, HRNet, and DeepLabV3+. Specifically, on the public BRACS dataset, our method produced an F1-Score of 92.98%, which outperforms the F1-Score of U-Net, DoubleU-Net, and HRNet by 4.24%, 0.37%, and 1.68%, respectively, and is much better than performances of DeepLabV3+ and ResUnet++ by 7.83% and 23.84%, respectively. On the private clinic dataset, the proposed model achieved an F1-Score of 73.13%, which has shown an improvement of 10.31%, 1.89%, 4.88%, 15.47%, and 31.1% to the performances of the U-Net, DoubleU-Net, HRNet, DeepLabV3+, and ResUnet++, respectively. Superior performance could also be observed when comparing the proposed DKS-DoubleU-Net with the others using the metrics of Dice and mIoU.

1. Introduction

Breast cancer is one of the most common cancers for women in the world [1]. With the increase of breast cancer patients worldwide, the precise diagnosis and prognosis of breast cancer have become urgent tasks. The Nottingham grading system [2] has become the most adopted clinic routine criteria in the diagnosis and prognosis of breast cancer. As one of the three essential factors (tubule formation, nuclear pleomorphism, and mitotic counts) in the Nottingham grad-

ing system, the assessment of breast tubule formation plays an important role [2].

The arrangement and morphology of tumor cells within breast tubules are vital criteria for the diagnosis of breast ductal carcinoma. Breast ductal carcinoma will destroy the function of breast ducts, resulting in the disappearance of tubular structure of breast tubules. Therefore, it is particularly important to accurately identify breast tubules in pathological images for a precise grading and diagnosis of breast cancer patients.

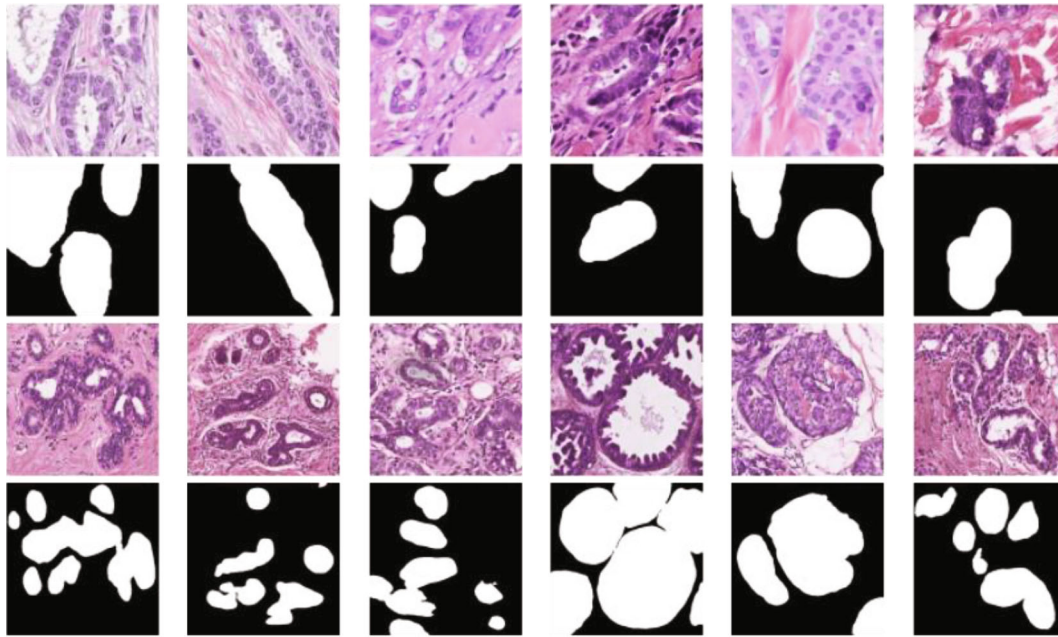


FIGURE 1: Demonstration of the complexity of breast tubules. Breast tubules are usually of diverse appearances, complex structures, and highly sophisticated morphology. The first and third row show several breast tubular images from public BRACS dataset and private clinical dataset. The second and fourth row represent their corresponding ground truths.

The identification of breast tubules is usually done by the manual observation of pathologists through microscopes. This process is not only time- and labor-consuming but also prone to errors due to interobserver and intraobserver discrepancies. Thus, it is necessary to develop a model to automatically recognize the breast tubule in histopathological images for a more objective and accurate computational pathology analysis.

Breast tubules are located in the subcutaneous tissue of the chest. Although breast tubules are characterized by clear lumina surrounded by epithelial cells, they are easily confused with clefts induced by shrinkage artifacts [2]. Moreover, breast tubules are typical of diverse appearances, complex structures, and highly sophisticated morphology, such as various shapes (tubular, round, or oval), colors, and textures, as shown in Figure 1. Therefore, it is a challenging task for traditional machine learning methods to identify breast tubules effectively. A promising way is to train end-to-end deep learning models [3–7], which could automatically discover the essential characteristics of breast tubules through the nonlinear representation of multiple layers of neurons.

Nowadays, with the rapid development of deep learning technology, increasing image segmentation models, such as U-Net [8], ResUNet++ [9], DeepLabV3+ [10], DoubleU-Net [11], and HRNet [12], have shown good performance in the field of medical image segmentation. For example, the U-Net model was applied to the segmentation of neuronal structures in electron microscopic recordings and the cell segmentation task in light microscopic images. Moreover, many U-Net-based Convolutional Neural Networks (CNNs) have also demonstrated superior performance in the task of medical image segmentation [13–17]. Particularly, among

the outstanding U-Net derived models, DoubleU-Net has shown promising results on various medical image datasets, including polyp detection dataset, lesion boundary segmentation dataset, and nuclei image dataset.

There have been some approaches dedicated to semantically segmenting breast tubules in hematoxylin and eosin (H&E) stained images. For example, Wang et al. [18] proposed a two-step clustering and random forest method for automatic recognition and segmentation of breast tubules. Basavanhally et al. [19] applied the O’Callaghan neighborhood in modeling and imposed spatial distance with directional constraints on object attributes for tubular identification on histopathological images stained with hematoxylin and eosin (H&E). Maqlin et al. [20] used K-means clustering to identify objects including glandular tubules and fatty regions. Zhang et al. [21] proposed a method for detecting tubules in testicular images based on boundary weighting and circular shortest path. Janowczyk and Madabhushi [22] detected breast tubules, lymphocytes, and epithelial cells by using the deep learning framework of the Caffe AlexNet network model.

However, the current deep learning models are still incapable of accurate segmentation of breast tubules, because breast tubules are usually of diverse appearance, complex structures, and different sizes for patients in different grading stages and various tissue transections.

Although deep-learning-based methods have been applied to the segmentation of breast tubules [8–12], most of them have not taken into account the reuse and fusion of multilevel features to better describe the tubule ROI regions of complex structures, diverse shapes, and different sizes for a more accurate segmentation performance.

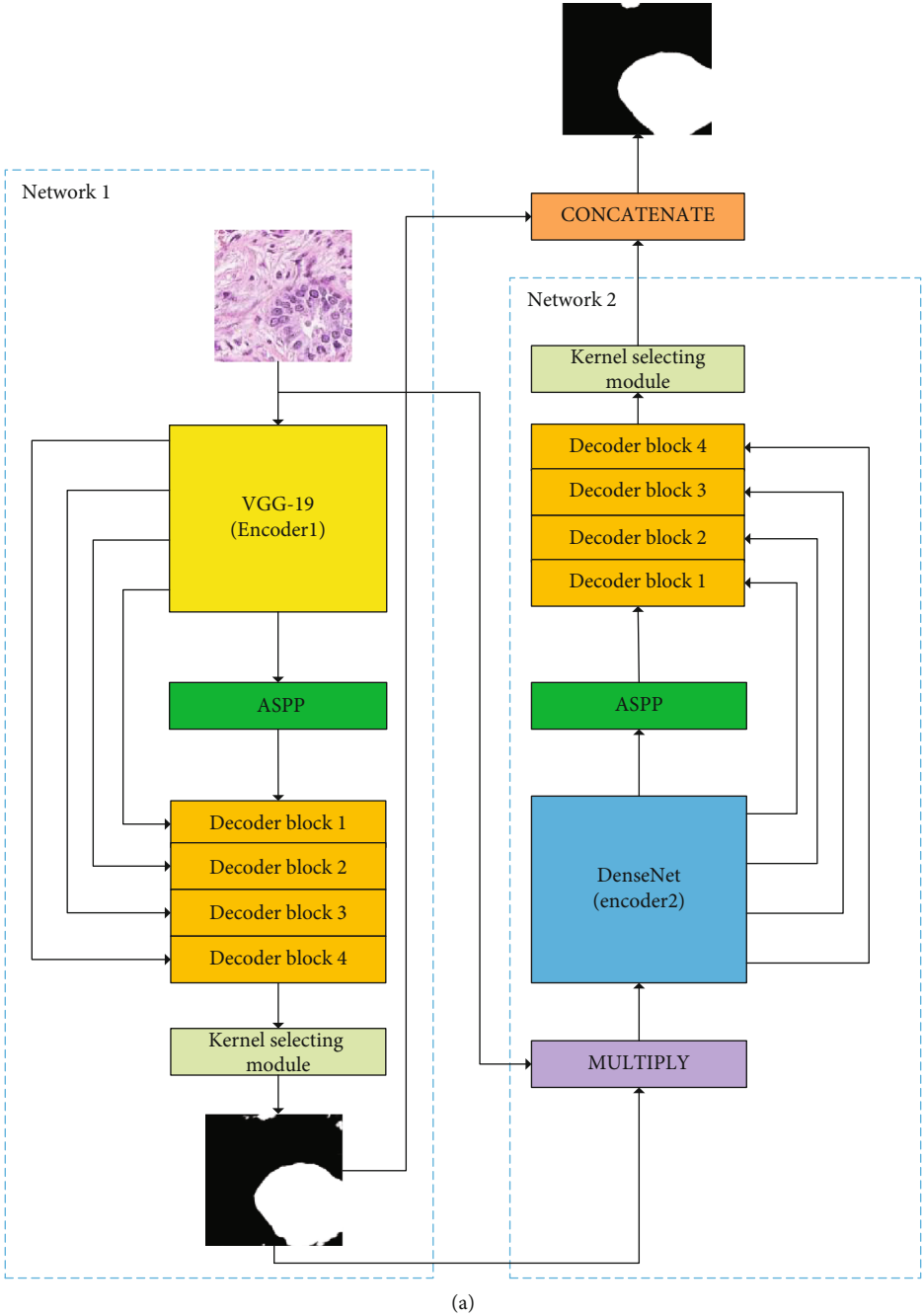


FIGURE 2: Continued.

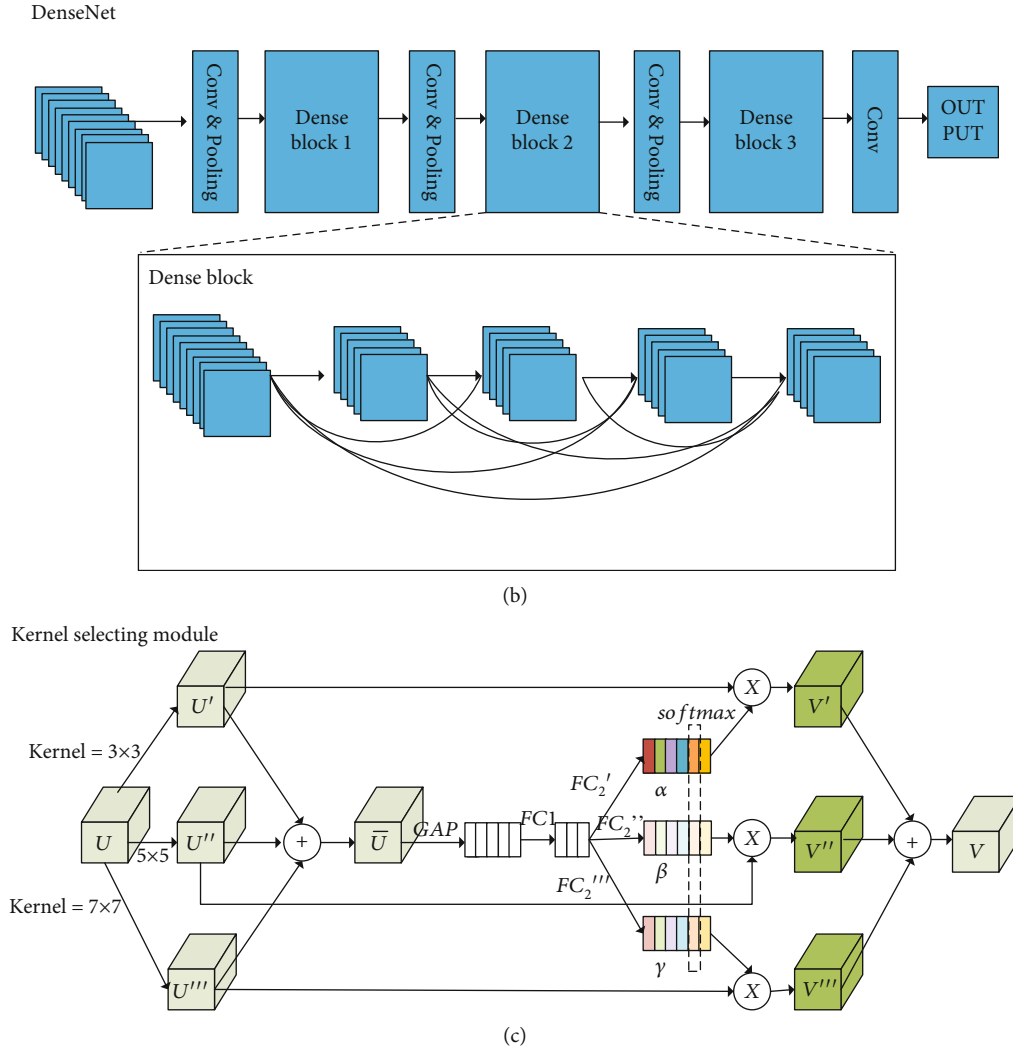


FIGURE 2: The overall architecture of the proposed DKS-DoubleU-Net model and the details of the introduced DenseNet and Kernel Selecting Module (KSM) for semantic breast tubule segmentation. (a) The overall architecture of the proposed DKS-DoubleU-Net model. (b) The detailed architecture of the DenseNet module [23]. (c) The detailed architecture of the Kernel Selecting Module (KSM) module [24].

TABLE 1: The quantitative comparison of the tubule segmentation results obtained by different models on the test set of the private clinical dataset.

Model	Dice	mIou	Recall	Precision	F1-score
DKS-DoubleU-Net	0.7002	0.5547	0.7636	0.7016	0.7313
DoubleU-Net+DenseNet	0.6897	0.5412	0.7442	0.7066	0.7249
DoubleU-Net+KSM	0.6820	0.5364	0.7165	0.7243	0.7204
DoubleU-Net	0.6798	0.5308	0.7450	0.6825	0.7124
U-Net	0.6115	0.4519	0.6568	0.6020	0.6282
ResNet++	0.3922	0.2672	0.6610	0.3081	0.4203
DeepLabV3+	0.5513	0.3955	0.5885	0.5651	0.5766
HRNet	0.6608	0.5085	0.6496	0.7188	0.6825

In this paper, we propose a novel semantic segmentation model named DKS-DoubleU-Net by integrating a DenseNet [23] module and a Kernel Selecting Module (KSM) [24] into the DoubleU-Net model for accurate segmentation of breast tubules in H&E images. The adopted DenseNet encoder utilizes

dense features between layers to discover intrinsic characteristics and strengthen the feature propagation of breast tubules. Besides, the KSM module conducts a self-adaptive kernel selecting for the sake of a multilevel feature fusion, facilitating the detection of breast tubules of different sizes. Experiments on

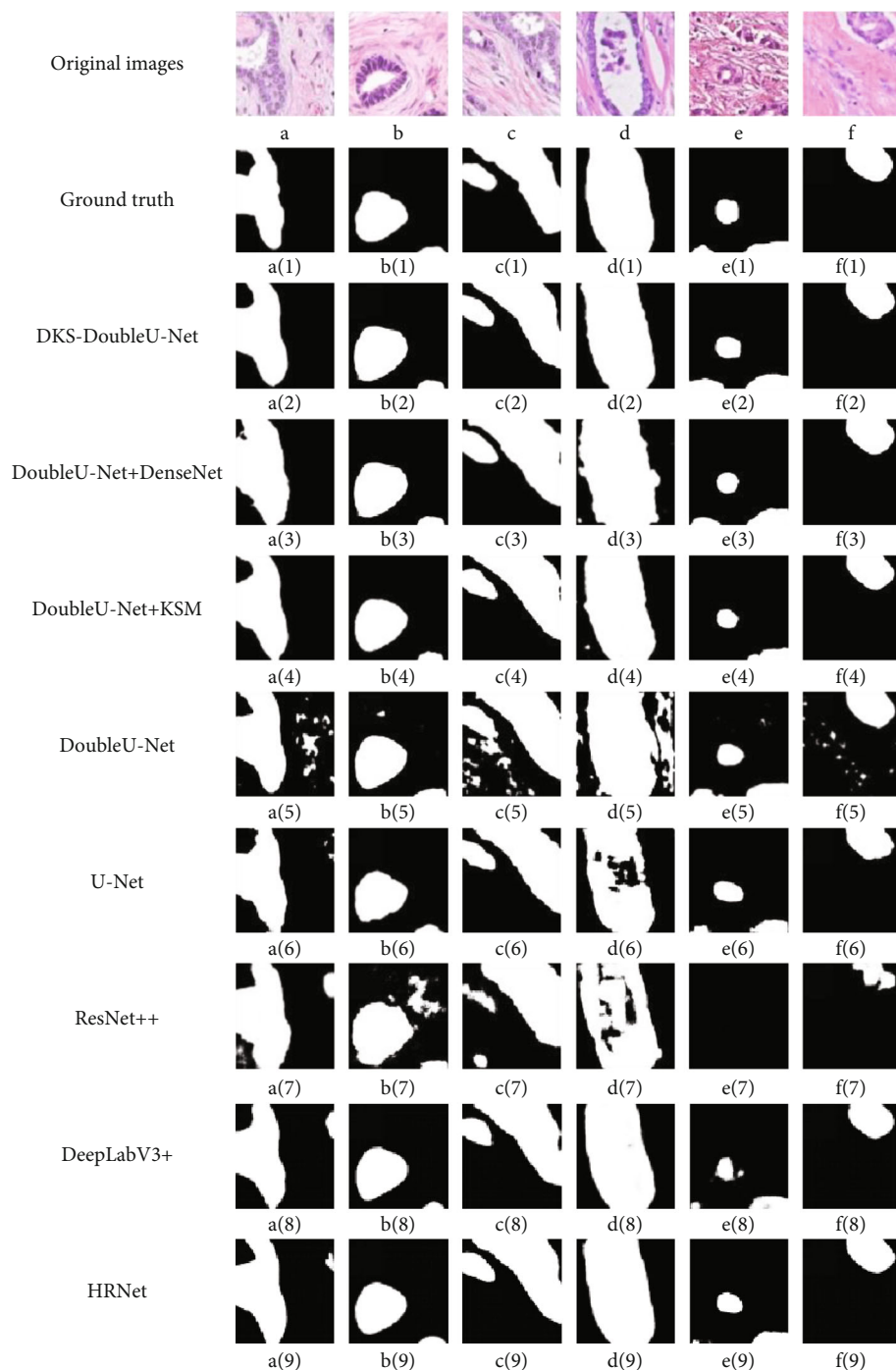


FIGURE 3: The illustrative tubule segmentation results of different models on the test set of the private clinical dataset. The first row lists the original images; images a(1)-f(1) in the second row are the corresponding tubule annotation masks; images a(2)-f(2) in the third row are the tubule segmentation results performed by DKS-DoubleU-Net; images a(3)-f(3) in the fourth row are the results output by the DoubleU-Net which added DenseNet; images a(4)-f(4) in the fifth row are the results obtained by the DoubleU-Net which added Kernel Selecting Module (KSM); images a(5)-f(5) in the sixth row are the results output by the DoubleU-Net; images a(6)-f(6) in the seventh row are the results output by U-Net; images a(7)-f(7) correspond to the results of ResUNet++; images a(8)-f(8) are the results obtained by DeepLabV3+, and images a(9)-f(9) show the tubule segmentation results of HRNet.

the public BRACS dataset [25] and a private clinical dataset have demonstrated the outperformance of the proposed DKS-DoubleU-Net in the semantic segmentation of the breast

tubules, compared to the original DoubleU-Net as well as the state-of-the-art segmentation models including U-Net [8], HRNet [12], DeepLabV3+ [10], and ResUNet++ [9].

2. Methods

2.1. The Proposed DKS-DoubleU-Net Architecture. For accurate semantic segmentation of breast tubules with complex structures and shapes, the proposed DKS-DoubleU-Net model is constructed through two main processes: (1) substituting the encoder of the second U-Net subnetwork branch of DoubleU-Net with a DenseNet [23] module; (2) inserting a Kernel Selecting Module (KSM) before each output layer of the two U-Net branches of the DoubleU-Net. The overall architecture of the proposed DKS-DoubleU-Net model is shown in Figure 2.

As shown in Figure 2, the proposed DKS-DoubleU-Net is composed of two U-Net subnetwork branches which are connected end to end, and each network branch is composed of an encoder and a decoder. Specifically, the VGG19 [26] is employed as the encoder module in the first network branch. The decoder modules of the two subnetwork branches are identical, both including four decoder blocks, and each of the decoder blocks consists of a squeeze and excitement block [27], 3×3 convolutional layers, and an upsampling layer. For each branch, the encoder and the decoder are connected with an Atrous Spatial Pyramid Pooling (ASPP) [10], which expands the receptive field of the convolutional layer. When the input image is fed to the first network, a preliminary binary segmentation mask will be generated to indicate the coarse detected regions of the tubules. Subsequently, the binary preliminary output tubule segmentation mask is multiplied by the original input image to remain only the coarse tubule ROI region in the original image. And the coarse tubule ROI image is further served as the input of the second network branch. The KSM module is inserted before each output layer of the two U-Net subnetwork branches. Finally, after the DenseNet encoder module and the decoder block of the second subnetwork branch, the final result of breast tubule segmentation on the input H&E image is output.

We will describe each component of the proposed DKS-DoubleU-Net model in detail.

2.2. Encoder. Feature extraction is the major function of the encoders in both branches [28]. The encoder of the first network branch adopts a VGG19 module, and all the convolutional layers of the module employ the same small size of 3×3 convolution kernels. Thus, VGG19 with a deep enough network structure could extract plenty of nonlinear features.

As the encoder of the second network branch, the DenseNet with a dense connection mechanism is introduced to extract the complex tubule structure features for a better tubule segmentation in H&E images. Specifically, each Dense Block (as shown in Figure 2(b)) of the introduced DenseNet module densely collects the output feature maps of all previous Dense Blocks as its input. Such dense connection of Dense Blocks makes the propagation and utilization of multilevel features more effective, resulting in a more efficient training process. That is, by establishing dense connections between the front and back layers, the reuse and propagation of features between different layers could alleviate the problem of gradient disappearance caused by the

deepening of the network to a certain extent. Therefore, the introduced DenseNet encoder has the potential to effectively extract, fuse, and propagate multilevel features of breast tubules on the ROI regions detected by the first U-Net branch.

2.3. Decoder. Decoders are used to restore the feature maps to the original resolution. The model has two decoders, each of which consists of four decoder blocks, as shown in Figure 2(a). For each layer of the decoder block, an upsampling operation is applied to enlarge the image through interpolation [29]. And then the upsampling feature map is concatenated with the feature map output from the corresponding encoder through a skip connection. Different from the skip connection from the encoder to the decoder in the first subnetwork branch, there are two skip connections from both encoder1 and encoder2 to the decoder in the second subnetwork branch. The purpose of this process is to reduce information loss in the process of feature extraction. After the upsampling, there are two 3×3 convolution operations, each of which is followed by the operations of normalization and ReLU [30]. The subsequent layers are composed of a squeeze and excitation block, a 1×1 convolution operation, and a sigmoid activation function [31]. The final output of the model is a binary image [32] having the same size as the input image to indicate the segmentation regions of breast tubules.

2.4. Kernel Selecting Module (KSM). The Kernel Selecting Module (KSM) fuses multiple convolutional branches with different kernel sizes to implement adaptive receptive field selection and multiscale feature fusion, as shown in Figure 2(c). Firstly, the input $U \in \mathfrak{R}^{H \times W \times C}$ performs 3×3 , 5×5 , and 7×7 convolution operations, respectively, to obtain $U' \in \mathfrak{R}^{H \times W \times C}$, $U'' \in \mathfrak{R}^{H \times W \times C}$, and $U''' \in \mathfrak{R}^{H \times W \times C}$. Secondly, U' , U'' , and U''' are added to obtain \bar{U} . The calculation formula of \bar{U} is

$$\bar{U} = U' + U'' + U''' \quad (1)$$

Then, the global average pooling operation (GAP) and the full connection operation (FC_1) are performed on \bar{U} , and after that, we perform the full connection operation (FC_2) on the obtained result again. The specific calculation formula is as follows:

$$\begin{aligned} \alpha' &= FC_2'(FC_1(GAP(\bar{U}))), \\ \beta' &= FC_2''(FC_1(GAP(\bar{U}))), \\ \gamma' &= FC_2'''(FC_1(GAP(\bar{U}))), \end{aligned} \quad (2)$$

where $\alpha' \in \mathfrak{R}^{1 \times 1 \times C}$, $\beta' \in \mathfrak{R}^{1 \times 1 \times C}$, and $\gamma' \in \mathfrak{R}^{1 \times 1 \times C}$. Three obtained outputs are calculated by softmax, respectively.

TABLE 2: The quantitative comparison of tubule segmentation results obtained by different models on the test set of the public BRACS dataset.

Model	Dice	mIou	Recall	Precision	F1-Score
DKS-DoubleU-Net	0.9272	0.8655	0.9292	0.9305	0.9298
DoubleU-Net+DenseNet	0.9247	0.8617	0.9288	0.9259	0.9273
DoubleU-Net+KSM	0.9269	0.8655	0.9328	0.9264	0.9296
DoubleU-Net	0.9219	0.8570	0.9204	0.9319	0.9261
U-Net	0.8793	0.7887	0.9239	0.8537	0.8874
ResNet++	0.6547	0.5020	0.7450	0.6450	0.6914
DeepLabV3+	0.8413	0.7301	0.8547	0.8483	0.8515
HRNet	0.9089	0.8348	0.9368	0.8903	0.9130

The specific calculation method is as follows:

$$\begin{aligned}\alpha_c &= \frac{e^{\alpha'_c}}{e^{\alpha'_c} + e^{\beta'_c} + e^{\gamma'_c}}, \\ \beta_c &= \frac{e^{\beta'_c}}{e^{\alpha'_c} + e^{\beta'_c} + e^{\gamma'_c}}, \\ \gamma_c &= \frac{e^{\gamma'_c}}{e^{\alpha'_c} + e^{\beta'_c} + e^{\gamma'_c}},\end{aligned}\quad (3)$$

where α , β , and γ denote the soft attention vectors, respectively. Note that α_c is the c -th element of α , and is the same for β_c and γ_c .

Finally, the feature maps obtained by different convolution kernels and the obtained attention vectors are multiplied and added to obtain the final output feature map V .

$$V_c = \alpha_c \cdot U' + \beta_c \cdot U'' + \gamma_c \cdot U''', \quad (4)$$

where α , β , and γ need to satisfy $\alpha_c + \beta_c + \gamma_c = 1$, and $V = [V_1, V_2, \dots, V_c]$, $V_c \in \mathcal{R}^{H \times W}$.

3. Experiments

3.1. Datasets. In the task of breast tubule segmentation, we conducted experiments on the public BRACS dataset [25] and a private clinical dataset. For each dataset, we randomly selected 10% of images for testing, 10% of images for verification, and the rest 80% of the images for model training.

The first experimental dataset is a private clinical dataset with a total of 398 images, each of which is 2000×2000 in size. The tubule masks on each of the images in this dataset were annotated by a pathologist, with white areas representing the breast tubules and black areas indicating irrelevant areas.

The second dataset is an openly accessible dataset named BRACS, which has a total of 4539 H&E images of various sizes. Since the breast tubules on the images of this dataset were not labeled, we did manual annotation of the breast tubules with the help of a pathologist before our experiments.

3.2. Implementation Details. Our experiments ran on the Ubuntu operation system with GPUs of GeForce GTX

1080 [33]. Our model was built using the Keras framework with TensorFlow [34, 35] as the back end. We used the Dice function [36] as the loss function of model training, the optimizer was Adam [37], the learning rate was set to $1e-5$, the batch size of the DKS-DoubleU-Net model and those of the other comparable models were set to 4, and the epoch was set to 200.

Our private dataset includes 398 original images with a size of 2000×2000 cropped from WSI images and their corresponding ground truths of tubule masks. Since the size of each of the private clinical pictures is 2000×2000 , to reduce the computational pressure of the graphics processing system [38] and improve the efficiency of code running, we cut each 2000×2000 image into 16 random overlapping 512×512 tiles, resulting in 6368 tiles in total. Since the cropped images do not change the structures of tubules, it will not affect the actual segmentation effects.

The public BRACS dataset is composed of 4539 H&E images of different image sizes, we resized the images to 512×512 for consistency with the 512×512 tile size of the private dataset. Since there are no tubule annotations in the public dataset, we generated the labels of the breast tubules beforehand. To speed up the annotation process, we adopted a coarse-to-fine annotation procedure on this dataset.

Specifically, as the first step, we randomly selected 200 H&E images and asked a pathologist to manually annotate the breast tubules using HistoView [39]. Subsequently, we trained the proposed model by using the 200 annotated images and obtained a preliminary model. The preliminary model in our experiment was a DoubleU-Net adopting a deep Residual module as the encoder of the second subnetwork. With the help of this preliminarily trained model, we tested the remaining unlabeled images and treat their outputs as coarse annotations of tubules. Then, the coarse annotations were denoised by applying the “bwareopen” function in MATLAB to remove small areas [40]. Finally, the refined breast tubule annotations were generated by manually eliminating the false positives and false negatives on the coarse annotation images with the help of the pathologist [41].

3.3. The Objective Function and Evaluation Metrics. To train the proposed DKS-DoubleU-Net model, Dice function is used as the loss function to emphasize the accurate semantic segmentation of the tubule regions. The Dice loss function is

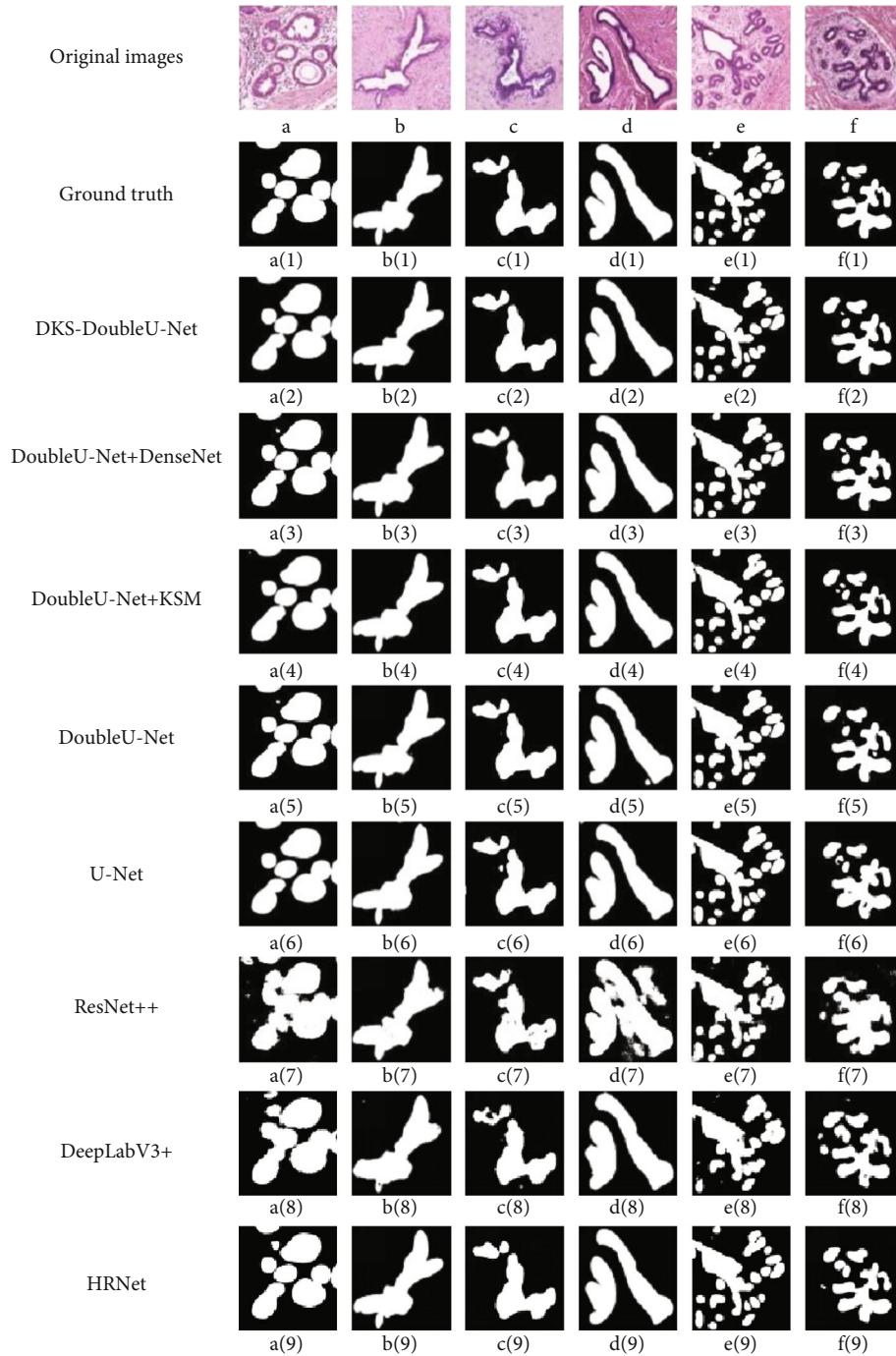


FIGURE 4: The illustrative tubule segmentation results of different models on the test set of the public BRACS dataset. The first row lists the original images; images a(1)-f(1) in the second row are the corresponding tubule annotation masks; images a(2)-f(2) in the third row are the tubule segmentation results performed by DKS-DoubleU-Net; images a(3)-f(3) in the fourth row are the results output by the DoubleU-Net which added DenseNet; images a(4)-f(4) in the fifth row are the results output by the DoubleU-Net which added Kernel Selecting Module (KSM); images a(5)-f(5) in the sixth row are the results output by the DoubleU-Net; images a(6)-f(6) in the seventh row are the results output by U-Net; images a(7)-f(7) correspond to the results of ResUnet++; images a(8)-f(8) are the results obtained by DeepLabV3+, and images a(9)-f(9) show the tubule segmentation results of HRNet.

TABLE 3: Comparison of the computational complexity of the models.

Model	Trainable params	FLOPs
DKS-DoubleU-Net	127.38 M	254.71 M
DoubleU-Net+DenseNet	127.24 M	254.42 M
DoubleU-Net+KSM	111.88 M	223.69 M
DoubleU-Net	111.73 M	223.40 M
U-Net	98.68 M	197.31 M
ResNet++	15.50 M	30.97 M
DeepLabV3+	155.88 M	311.76 M
HRNet	108.93 M	217.85 M

widely used in medical image segmentation tasks. The definition of Dice Loss is as follows:

$$\text{DiceLoss} = 1 - \frac{2|X \cap Y|}{|X| + |Y|}, \quad (5)$$

where $X \cap Y$ denotes the intersection between X and Y , $|X|$ and $|Y|$ represent the number of elements, respectively, and the numerator is multiplied by 2 to ensure that the value range is between 0 and 1.

To measure the superiority of the proposed model more quantitatively and comprehensively, Dice, mIoU, Precision, Recall, and F1-Score, which are the most important and commonly used metrics in semantic segmentation tasks [42–45], served as evaluation metrics in our experiments:

$$\begin{aligned} \text{Dice} &= \frac{2|X \cap Y|}{|X| + |Y|}, \\ \text{mIoU} &= \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}, \\ \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{F1-Score} &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \end{aligned} \quad (6)$$

where TP represents true positive, FP denotes false positive, and FN means false negative.

4. Results and Analysis

To evaluate the performance of the proposed DKS-DoubleU-Net model, we conducted experiments on both the private clinical dataset and the public BRACS dataset independently. The experiment results of the proposed DKS-DoubleU-Net model are compared with those of the state-of-the-art segmentation models, including U-Net [8], DoubleU-Net [11], ResUnet++ [9], DeepLabV3+ [10], and HRNet [12] models. To assess the performances of different models objectively, we employ the most commonly used evaluation metrics in semantic segmentation tasks, such as Dice, mIoU, Precision, Recall, and F1-Score in our experiments.

4.1. Experiments on the Private Clinical Dataset. The quantitative comparison of the tubule segmentation results obtained by different models on the private clinical dataset is shown in Table 1. The proposed DKS-DoubleU-Net model achieves a Dice of 70.02%, a mIoU of 55.47%, a Recall of 76.36%, a Precision of 70.16%, and an F1-Score of 73.13%, which is superior to the original DoubleU-Net model by increasing 2.04%, 2.39%, 1.86%, 1.91%, and 1.89%, respectively. The individual combination of the DenseNet module or KSM module with the original DoubleU-Net also exhibited an improvement in performance as shown in Table 1. This may be attributed to the combined action of the powerful feature extraction ability of DenseNet and the multiscale feature fusion mechanism of KSM. The former could accurately identify the various shapes of tubules and the latter improves the detection accuracy of tubules in different sizes. Moreover, compared to the U-Net model, the proposed DKS-DoubleU-Net has improved the scores of Dice, mIoU, Recall, Precision, and F1-Score obtained by 8.87%, 10.28%, 10.68%, 9.96%, and 10.31%, respectively. In summary, the proposed DKS-DoubleU-Net model has achieved the highest scores in Dice, mIoU, and F1-Score, except for precision, among all listed state-of-the-art models. Although the model of “DoubleU-Net+KSM” has the highest precision score, the decrease in its recall score leads to a decrease in its F1-Score (i.e., the harmonic mean of Recall and Precision).

To visualize the performances of different models, we demonstrate several tubule semantic segmentation results on the test set of the private clinical dataset, as shown in Figure 3. As can be seen from Figure 3, DKS-DoubleU-Net has the most excellent performance even in segmenting the challenging targets. Specifically, its segmentation results are the closest to the ground truth images and with the fewest noise spots.

4.2. Experiments on the Public BRACS Dataset. To further analyze the performance of breast tubule segmentation by the proposed DKS-DoubleU-Net, we conducted experiments on the public BRACS dataset [25]. The comparative evaluation results of different models on the public BRACS dataset are shown in Table 2.

Table 2 shows that the proposed DKS-DoubleU-Net achieves scores of 92.72% in Dice, 86.55% in mIoU, 92.92% in Recall, 93.05% in Precision, and 92.98% in F1-Score. Specifically, after introducing the modules of DenseNet and KSM into the original DoubleU-Net model, the segmentation results have increased by 0.53%, 0.85%, 0.88%, and 0.37% in Dice, mIoU, Recall, and F1-Score, respectively. This may be put down to the flexible ability of the introduced DenseNet and KSM modules in segmenting the breast tubules in different sizes and diverse shapes. Compared to the U-Net, the proposed DKS-DoubleU-Net has gained the biggest increase in Dice, mIoU, Recall, Precision, and F1-Score by 4.79%, 7.68%, 0.53%, 7.68%, and 4.24%, respectively. On the public BRACS data set, our model fails to reach the highest recall and precision scores, but its F1-Score, resulting from the harmonic mean of recall and precision, is the most superior among all state-of-the-art methods as shown in Table 2.

The illustrative tubule segmentation results obtained by different models on the test set of the public BRACS dataset are shown in Figure 4. It can be seen from Figure 4 that, compared to the state-of-the-art networks, the tubule segmentation results performed by the proposed DKS-DoubleU-Net are more effective in suppressing noisy spots and more accurate in detecting the tubules with various sizes and diverse shapes.

In summary, the proposed DKS-DoubleU-Net model has shown outstanding performance in detecting breast tubules of diverse shapes and different sizes on both the private clinical dataset and the public BRACS dataset.

Specifically, from the quantitative comparison results in Table 1 and Table 2, it can be seen that, among the state-of-the-art models, the proposed DKS-DoubleU-Net model achieves the highest evaluation scores of Dice, mIoU, and F1-Score in the tubule segmentation on both datasets. Besides, the compared illustrative tubule segmentation results, as shown in Figures 3 and 4, also demonstrate that the proposed DKS-DoubleU-Net performs best in the accurate breast tubule segmentation, with the results closest to the ground truth among all state-of-the-art models including DoubleU-Net, U-Net, ResUnet++, DeepLabV3+, and HRNet. Therefore, the introduced modules of DenseNet and Kernel Selecting Module (KSM) in the proposed DKS-DoubleU-Net model could effectively improve the model's ability in complex feature extraction and multiscale feature fusion, which are important in the performance promotion of breast tubule segmentation.

4.3. Computational Complexity Analysis. In order to compare the computational complexity of the models, we measure the computational complexity from the amount of parameters (Paras) and the amount of calculation (FLOPs) as shown in Table 3. Each parameter is a float, that is, a parameter is 4 bytes. And the units of these two metrics are MB. The amount of parameters corresponds to the space complexity of the model, that is, the size of the model. The amount of computation corresponds to the time complexity of the model, that is, the length of the network execution time. It can be seen from Table 3 that the amounts of parameters and calculation of the DKS-DoubleU-Net network are a little bit larger than those of the original DoubleU-Net model, due to the addition of the DenseNet and KSM modules. Although the space complexity and time complexity of the DKS-DoubleU-Net network is a little bit higher, it results in a high performance.

5. Discussion

Breast tubule formation is one of the three essential factors in the clinical routine Nottingham grading system. The assessment of breast tubule formation is critical to the accurate diagnosis and prognosis of breast cancer [2]. In practice, it is a challenging task to accurately segment the breast tubules in H&E images due to their diverse appearance, complex structures, and highly sophisticated morphology. Although deep-learning-based methods have been applied to the segmentation of breast tubules, most of them have

not taken into account the reuse of dense features among layers and the fusion of multilevel features for a more accurate semantic segmentation performance of breast tubules. To solve this problem, the proposed DKS-DoubleU-Net model introduced a DenseNet [23] as the encoder module of the second network branch in DoubleU-Net. Besides, the Kernel Selecting Module (KSM) was inserted in front of the outputs of the two U-Net branches in the proposed DKS-DoubleU-Net.

The potential roles of the two modules played in the task of tubule segmentation could be inferred as follows: (1) the DenseNet module, utilizing dense connections between layers through Dense Blocks, would help better extract features from the tubules with complex structures and diverse shapes and (2) the KSM module, enabling a self-adaptive kernel selecting for feature fusion, would facilitate the discovery of tubules of different sizes. Therefore, both qualitative and quantitative experiments have shown that our model achieves competitive performance on both the private clinical dataset and public BRACS dataset. Specifically, in view of quantitative assessment, the DKS-DoubleU-Net produced a Dice of 70.02% and an F1-Score of 73.13% on the private data, which has increased by 2.04% in Dice and 1.89% in F1-Score compared to the original DoubleU-Net. On the public BRACS dataset, our proposed model achieved a Dice of 92.72% and an F1-Score of 92.98%, which outperformed the DoubleU-Net by 0.53% and 0.37%, respectively. Both in terms of quantitative metrics and qualitative visual perception of the tubule segmentation results, our model showed the best performance among the state-of-the-art models, including DoubleU-Net [11], U-Net [8], ResUnet++ [9], DeepLabV3+ [10], and HRNet [12].

The limitation of the DKS-DoubleU-Net is that after the introduction of DenseNet and Kernel Selecting Module (KSM), the network structure becomes more complex and has more parameters to learn, therefore, in our future work, we will focus on optimizing the network structure for more efficient and accurate segmentation performance on diverse tubule regions. As another limitation, our network cannot handle images of random size. In our experiment, we cropped or resize the input image to the size of 512×512 . Therefore, to develop an algorithm dealing with input images of random sizes will be one of our further works.

6. Conclusions

In this paper, we proposed a semantic segmentation model termed DKS-DoubleU-Net for accurate semantic segmentation of breast tubules in H&E images. The proposed DKS-DoubleU-Net adopted a DenseNet module as the second encoder of the DoubleU-Net and inserted a Kernel Selecting Module (KSM) before the output layer in each of the two U-Net branches. Our purpose is to discover, reuse, fuse, and propagate the dense features and multilevel features extracted from the coarse breast tubular regions suggested by the first U-Net subnetwork branch of DoubleU-Net. Moreover, in the experiment section, we applied the proposed DKS-DoubleU-Net to two datasets. In terms of both quantitative evaluation and visual perception quality, the

proposed DKS-DoubleU-Net achieved competitive performance in the semantic segmentation of breast tubules in the H&E images on both datasets, as compared with the state-of-the-art models. Therefore, DKS-DoubleU-Net has the potential of being a baseline for breast tubule segmentation in H&E images.

Data Availability

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.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 61501287, No. 61971273, No. U2001205, and No. 61877038), the Key Research and Development Program in Shaanxi Province of China (No. 2021GY-032), the State Key Laboratory of Geo-Information Engineering (SKLGIE2019-M-3-5), the Fundamental Research Funds for the Central Universities (No. GK202003077), the Research and Development Project funded by Central University/Public Health and Health Research Project (GK202007033). The authors would like to thank the Xi'an Future Artificial Intelligence Computing Center for providing the artificial intelligence computing power and technical support.

References

- [1] M. Al-Hajj, M. S. Wicha, A. Benito-Hernandez, S. J. Morrison, and M. F. Clarke, "Prospective identification of tumorigenic breast cancer cells," *Proceedings of the National Academy of Sciences*, vol. 100, no. 7, pp. 3983–3988, 2003.
- [2] C. W. Elston and I. O. Ellis, "Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up," *Histopathology*, vol. 19, no. 5, pp. 403–410, 1991.
- [3] I. H. Sarker, "Deep cybersecurity: a comprehensive overview from neural network and deep learning perspective," *SN Computer Science*, vol. 2, no. 3, pp. 1–16, 2021.
- [4] J. Liu, Y. Wu, Y. Wu et al., "DAM: discrepancy alignment metric for face recognition," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3814–3823, Montreal, QC, Canada, 2021.
- [5] A. Choudhary, L. Tong, Y. Zhu, and M. D. Wang, "Advancing medical imaging informatics by deep learning-based domain adaptation," *Yearbook of Medical Informatics*, vol. 29, no. 1, pp. 129–138, 2020.
- [6] M. H. Laves, S. Ihler, T. Ortmaier, and L. A. Kahrs, "Quantifying the uncertainty of deep learning-based computer-aided diagnosis for patient safety," *Current Directions in Biomedical Engineering*, vol. 5, no. 1, pp. 223–226, 2019.
- [7] L. Deng, Y. Liu, Y. Shi, and H. Liu, "A deep neural network approach using distributed representations of RNA sequence and structure for identifying binding site of RNA-binding proteins," in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 12–17, San Diego, CA, USA, 2019.
- [8] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Springer, Cham, 2015.
- [9] D. Jha, P. H. Smedsrud, M. A. Riegler et al., "Resunet++: An advanced architecture for medical image segmentation," in *2019 IEEE International Symposium on Multimedia (ISM)*, pp. 225–2255, San Diego, CA, USA, 2019.
- [10] L. C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017, <http://arxiv.org/abs/1706.05587>.
- [11] D. Jha, M. A. Riegler, D. Johansen, P. Halvorsen, and H. D. Johansen, "Doubleu-net: A deep convolutional neural network for medical image segmentation," in *2020 IEEE 33rd international symposium on computer-based medical systems (CBMS)*, pp. 558–564, Rochester, MN, USA, 2020.
- [12] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep high-resolution representation learning for human pose estimation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5693–5703, Long Beach, CA, USA, 2019.
- [13] X. Li, H. Chen, X. Qi, Q. Dou, C. W. Fu, and P. A. Heng, "HDenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes," *IEEE Transactions on Medical Imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.
- [14] L. Kong and Y. Zhang, "Multi-modal brain tumor Segmentation using cascaded 3D U-Net," in *2021 6th International Conference on Image, Vision and Computing (ICIVC)*, pp. 129–133, Qingdao, China, 2021.
- [15] H. Khodajou-Chokami, S. A. Hosseini, and M. R. Ay, "PARS-NET: a novel deep learning framework using parallel residual conventional neural networks for sparse-view CT reconstruction," *Journal of Instrumentation*, vol. 17, no. 2, p. P02011, 2022.
- [16] V. Iglovikov and A. Shvets, "Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation," 2018, <http://arxiv.org/abs/1801.05746>.
- [17] X. Guan, G. Yang, J. Ye et al., "3D AGSE-VNet: an automatic brain tumor MRI data segmentation framework," *BMC Medical Imaging*, vol. 22, no. 1, pp. 1–18, 2022.
- [18] S. Wang, J. Liu, B. I. Yao-Yao, Z. Chen, Q. H. Zheng, and H. F. Duan, "Automatic recognition of breast gland based on two-step clustering and random Forest," *Computer Science*, vol. 45, pp. 247–252, 2018.
- [19] A. Basavanhally, E. Yu, J. Xu et al., "Incorporating domain knowledge for tubule detection in breast histopathology using O'Callaghan neighborhoods," *Medical Imaging 2011: Computer-Aided Diagnosis*, vol. 7963, pp. 305–319, 2011.
- [20] P. Maqlin, R. Thamburaj, J. J. Mammen, and A. K. Nagar, "Automatic detection of tubules in breast histopathological images," in *Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012)*, pp. 311–321, Springer, India, 2013.
- [21] C. Zhang, C. Sun, R. Davey et al., "Tubule detection in testis images using boundary weighting and circular shortest path," in *2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pp. 3319–3322, Osaka, Japan, 2013.

- [22] A. Janowczyk and A. Madabhushi, "Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases," *Journal of Pathology Informatics*, vol. 7, no. 1, p. 29, 2016.
- [23] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708, Honolulu, HI, USA, 2017.
- [24] Y. Zhao, Z. Jiang, A. Men, and G. Ju, "Pyramid real image denoising network," in *2019 IEEE visual communications and image processing (VCIP)*, pp. 1–4, Sydney, NSW, Australia, 2019.
- [25] N. Brancati, A. M. Anniciello, P. Pati et al., "Bracs: A dataset for breast carcinoma subtyping in h&e histology images," 2021, <http://arxiv.org/abs/2111.04740>.
- [26] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, <http://arxiv.org/abs/1409.1556>.
- [27] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141, Salt Lake City, UT, USA, 2018.
- [28] M. Nixon and A. Aguado, *Feature Extraction and Image Processing for Computer Vision*, Academic press, 2020.
- [29] F. Handoyo, *Resizing or Scaling Image from Small Image to Enlarge Image Using Bilinear Interpolation Algorithm, Median Filter Algorithm, and Unsharp Masking Algorithm*, Fakultas Ilmu Komputer, 2011.
- [30] J. Ge, X. Cui, K. Xiao, C. Zou, Y. Chen, and R. Wei, "BNReLU: combine batch normalization and rectified linear unit to reduce hardware overhead," in *2019 IEEE 13th International Conference on ASIC (ASICON)*, pp. 1–4, Chongqing, China, 2019.
- [31] S. Narayan, "The generalized sigmoid activation function: competitive supervised learning," *Information Sciences*, vol. 99, no. 1-2, pp. 69–82, 1997.
- [32] M. Wu and B. Liu, "Data hiding in binary image for authentication and annotation," *IEEE Transactions on Multimedia*, vol. 6, no. 4, pp. 528–538, 2004.
- [33] D. G. Peng, H. Zhang, L. Yang, and H. Li, "Design and realization of modbus protocol based on embedded linux system," in *2008 International Conference on Embedded Software and Systems Symposia*, pp. 275–280, Chengdu, China, 2008.
- [34] K. Ramasubramanian, A. Singh, K. Ramasubramanian, and A. Singh, *Deep Learning Using Keras and Tensorflow*, Machine Learning Using R. Apress, Berkeley, CA, 2019.
- [35] K. Jakhar and N. Hooda, "Big data deep learning framework using keras: a case study of pneumonia prediction," in *2018 4th International Conference on computing communication and automation (ICCCA)*, pp. 1–5, Greater Noida, India, 2018.
- [36] Q. Huang, J. Sun, H. Ding, X. Wang, and G. Wang, "Robust liver vessel extraction using 3D U-net with variant dice loss function," *Computers in Biology and Medicine*, vol. 101, pp. 153–162, 2018.
- [37] D. P. Kingma and B. J. Adam, "A method for stochastic optimization," 2014, <http://arxiv.org/abs/1412.6980>.
- [38] J. Yuan, G. Xu, Y. Yu et al., "Real-time photoacoustic and ultrasound dual-modality imaging system facilitated with graphics processing unit and code parallel optimization," *Journal of Biomedical Optics*, vol. 18, no. 8, article 086001, 2013.
- [39] <https://engineering.case.edu/centers/ccipd/content/software>.
- [40] W. Wu and P. Cai, "Simulation of wavelet denoising based on MATLAB," *Information and electronic engineering*, vol. 6, no. 3, pp. 220–222, 2008.
- [41] H. Yu, L. T. Yang, Q. Zhang, D. Armstrong, and M. J. Deen, "Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives," *Neurocomputing*, vol. 444, pp. 92–110, 2021.
- [42] P. Anilkumar and P. Venugopal, "A Survey on Semantic Segmentation of Aerial Images using Deep Learning Techniques," in *2021 Innovations in Power and Advanced computing technologies (i-PACT)*, pp. 1–7, Kuala Lumpur, Malaysia, 2021.
- [43] H. H. Jebamikyous and R. Kashef, "Deep Learning-Based Semantic Segmentation in Autonomous Driving," in *2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on dependability in sensor, Cloud & big Data Systems & application (HPCC/DSS/Smart-City/DependSys)*, pp. 1367–1373, Haikou, Hainan, China, 2021.
- [44] L. Yu, Z. Hu, F. Zhang, and K. Yang, "Unmanned aerial vehicle image biological soil crust recognition based on UNet++," *International Journal of Remote Sensing*, vol. 43, no. 7, pp. 2660–2676, 2022.
- [45] Y. Cai, H. He, K. Yang et al., "A comparative study of deep learning approaches to rooftop detection in aerial images," *Canadian Journal of Remote Sensing*, vol. 47, no. 3, pp. 413–431, 2021.