

Article

# **Evaluation of Deep Neural Networks for Semantic Segmentation of Prostate in T2W MRI**

Zia Khan <sup>1</sup><sup>(b)</sup>, Norashikin Yahya <sup>1,\*</sup><sup>(b)</sup>, Khaled Alsaih <sup>1</sup><sup>(b)</sup>, Syed Saad Azhar Ali <sup>1</sup><sup>(b)</sup> and Fabrice Meriaudeau <sup>2</sup>

- <sup>1</sup> Centre for Intelligent Signal and Imaging Research (CISIR), Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Malaysia; zia.aseer@gmail.com (Z.K.); khaledalsaih@gmail.com (K.A.); saad.azhar@utp.edu.my (S.S.A.A.)
- <sup>2</sup> ImViA/ITFIM, University of Burgundy, 21078 Dijon, France; fabrice.meriaudeau@u-bourgogne.fr
- \* Correspondence: norashikin\_yahya@utp.edu.my

Received: 9 February 2020; Accepted: 12 April 2020; Published: 3 June 2020



Abstract: In this paper, we present an evaluation of four encoder–decoder CNNs in the segmentation of the prostate gland in T2W magnetic resonance imaging (MRI) image. The four selected CNNs are FCN, SegNet, U-Net, and DeepLabV3+, which was originally proposed for the segmentation of road scene, biomedical, and natural images. Segmentation of prostate in T2W MRI images is an important step in the automatic diagnosis of prostate cancer to enable better lesion detection and staging of prostate cancer. Therefore, many research efforts have been conducted to improve the segmentation of the prostate gland in MRI images. The main challenges of prostate gland segmentation are blurry prostate boundary and variability in prostate anatomical structure. In this work, we investigated the performance of encoder-decoder CNNs for segmentation of prostate gland in T2W MRI. Image pre-processing techniques including image resizing, center-cropping and intensity normalization are applied to address the issues of inter-patient and inter-scanner variability as well as the issue of dominating background pixels over prostate pixels. In addition, to enrich the network with more data, to increase data variation, and to improve its accuracy, patch extraction and data augmentation are applied prior to training the networks. Furthermore, class weight balancing is used to avoid having biased networks since the number of background pixels is much higher than the prostate pixels. The class imbalance problem is solved by utilizing weighted cross-entropy loss function during the training of the CNN model. The performance of the CNNs is evaluated in terms of the Dice similarity coefficient (DSC) and our experimental results show that patch-wise DeepLabV3+ gives the best performance with DSC equal to 92.8%. This value is the highest DSC score compared to the FCN, SegNet, and U-Net that also competed the recently published state-of-the-art method of prostate segmentation.

Keywords: encoder-decoder; CNNs; DNN; FCN; SegNet; U-Net; DeepLabV3+

# 1. Introduction

According to the report published by the GLOBOCAN 2018 [1], global statistics have indicated that prostate cancer is the second most leading diagnosed cancer in men, after lung cancer. In 2018 alone, more than 1.2 million new cases of prostate cancer were reported worldwide with higher cases in developed countries. Prostate cancer may be asymptomatic in the early stage and often has an indolent course, and symptoms may not appear until the cancer is in its more advanced stages [2]. Hence, early detection is highly critical in preventing further growth of cancerous cells in the prostate gland.

In the current clinical approach, the prostate-specific antigen (PSA) blood test and transrectal ultrasound (TRUS) biopsy are applied together for the diagnosis of prostate cancer [3]. However,



PSA testing has low specificity because men without cancer have also been identified with elevated PSA and many patients with cancer did not have elevated levels of PSA [4]. On the other hand, TRUS biopsy suffers from low sensitivity due to unclear optimal threshold for biopsy [5]. Due to the low specificity of PSA and low sensitivity of TRUS biopsy, these tests have restricted diagnostic value [6].

Machine learning approaches in image-based diagnosis, disease prognosis, and risk assessment are becoming increasingly successful. For the diagnosis of prostate cancer, magnetic resonance imaging (MRI) is a common modality for evaluation of prostate cancer, providing superior anatomic or pathologic visualization when compared with ultrasound and computed tomography (CT). In the past few years, the use of MRI was proposed by the prostate cancer research community as the most accurate non-invasive screening tool for prostate cancer diagnosis, staging [7], disease monitoring, and treatment management [8].

The prostate gland can be divided into three zones, peripheral zone (PZ), transitional zone (TZ), and the central zone (CZ). Having the advantages of good resolution and contrast in soft tissues, T2W MRI is used to differentiate the different zones of the prostate gland [9]. It was established that the ratio of developing prostate cancer in the prostate gland varies according to the zone. PZ has the highest rate of 70–80%, whereas the TZ rate is 10–20%, while, in CZ, the rate is only 5% [10]. In a typical examination of MRI images, the radiologist performs slice by slice visual inspection and prostate segmentation which can be time-consuming especially for large-scale samples. In addition, the result of manual segmentation varies depending on the skill of the expertise [11]. Hence, to improve the diagnosis of prostate cancer, there is a need for an automatic segmentation technique that is robust to variability in shape, size, image intensity around the gland and surrounding regions, anisotropic spatial resolution, and scanner type [12,13].

In this article, we evaluated and compared the performance of four encoder–decoder CNNs; FCN, SegNet, DeepLabV3+ and U-Net in segmenting prostate T2W MRI into whole gland (WG), or into three zones PZ, TZ, and CZ depending on availability of dataset label. The training and testing of the CNNs are conducted with selected image pre-processing techniques, data augmentation techniques and selected loss function for dealing with the issue of class imbalance.

The paper is outlined as follows. In Section 2, we reviewed papers related to image segmentation focusing on atlas-based segmentation, model-based segmentation, deep learning-based segmentation, and patch CNN-based segmentation. This section also includes a summary of our contribution. Next, Section 3 presents materials and methods, and Section 4 presents results and discussion. Lastly, Section 5 concludes the paper and suggests further work in this area.

#### 2. Related Work

Prostate segmentation, a necessary step in automatic image analysis, involves delineating the prostate boundaries. Segmentation will limit subsequent image processing only on the organ of interest. This section highlights some important works on prostate segmentation based on deformable models, and region-based and patch-based CNNs.

Before the breakthrough of deep learning, segmentation based on deformable models [14,15], hybrid methods [16–19], and atlas-based [20–23] were introduced. In principle, the deformable models were implemented as the physical bodies with certain elasticity and force properties that try to keep their form, while the image to be segmented was represented as a potential field with the force that tries to stretch the model and adapt to it. To achieve this, deformable-based segmentation incorporates knowledge about the shape, object orientation, continuity, elasticity, or smoothness of the region to be segmented.

In atlas-based segmentation, two major steps are atlas construction and atlas registration where the prior knowledge related to size, orientation, or shape needs to be incorporated [20–23]. The three basic building blocks for atlas registration are image transformation, similarity measure, and optimization algorithm that finds the maximum similarity measure between atlas and target region.

Deformable-based methods are sensitive to initialization because the contour is driven by a pre-defined energy function in a variational manner, giving local convergence. As a result, the shape of the object may not be well preserved, especially in the presence of image noise and in the non-homogeneous intensity region [14,15]. On the other hand, atlas-based methods lack the flexibility in local tuning the segmentation boundary.

Martin et al. [17] combined the deformable model and probabilistic atlas-based model for the segmentation of prostate MRI achieving the median Dice similarity coefficient (DSC) score of 0.86. Zhang et al. [18] proposed a method that uses a clustering method to separate the prostate gland from the surrounding tissues followed by a postprocessing phase through active contours, giving an overall accuracy with mean DSC of  $0.84 \pm 0.04$ .

Methods based on deep learning had shown remarkable results relative to the combination of deformable model and atlas-based segmentation. Hence, various methods were proposed to exploit the advantages of both methods while avoiding their respective drawbacks, including combining both deformable models and atlas-based models. In [16], Cheng et al. combined a supervised atlas-based active appearance model (AAM) and CNN was achieving DSC value equal to 0.925. The method utilized the volume of interest (VOI) generated by AAM and later inputs to CNN as learnable features, and this has increased the overall accuracy of prostate segmentation. Yanrong Guo et al. [19] presented a novel hybrid model by combining a deep learning model and deformable model where the features extracted by the deep learning model were matched to the atlas map by a sparse matching method. This method has resulted in DSC value of  $87.1\% \pm 4.2\%$ . The subsequent paragraphs highlight some semantic segmentation methods based on deep learning styles.

Long et al. [24] introduced a fully convolutional neural network (FCN) for scene segmentation tasks. The FCN is built from a classical CNN but only consists of locally connected layers including convolution, pooling, and upsampling, whereby the last fully connected layer is replaced by another broad receptive field convolution layer. Since the architecture does not include fully connected layers, the number of parameters and computation time is reduced. In addition, unlike the traditional CNN model, the network can work with any image size without requiring any fixed number of units at any stage, as all connections are local.

The application of CNN for the segmentation of prostate MRI was investigated in [25]. The network, known as PSNet is a fine-tuning of the FCN model proposed by Long et al. [24] trained on the PASCALVOC data set [26] and retrained for the prostate segmentation. Specifically, PSNet fine-tuned the last three layers of the FCN because these layers learn high-level specific features that are related to the prostate gland [27,28].

In [29], Ronneberger et al. proposed U-Net model, which is proposed for medical imaging segmentation tasks. The U-Net model concatenates the contracting side (encoder) and expanding side (decoder) to compensate the loss of spatial resolution due to the max pooling operations. In another work by Milletari et al. [30], the Dice loss function was used to solve the problem of class imbalance between background and foreground pixels and obtained accurate 3D segmentation of the prostate T2W MRI using U-Net architecture. In the work by Yu et al. [31], residual connections are incorporated into the convolutional and deconvolutional part of U-Net and the proposed architecture won the first place in 2017 Computer-Assisted Intervention Conference using the PROMISE12 challenge dataset.

In recent work by Rundo et al. [32], a new segmentation method based on U-Net was proposed. The CNN network known as Use-Net introduced Squeeze-and-Excitation (SE) blocks in both encoder and the decoder of U-Net architecture. The network was trained and tested on multi-dataset allowing for both intra and cross dataset generalization. Their findings revealed that the SE blocks enable adaptive feature recalibration, hence, providing excellent cross-dataset generalization.

Patch-based CNN segmentation methods were proven to result in better pixel-wise segmentation of biomedical images [33–39]. In [33], Zhang et al. applied patch-based CNN for segmentation of the multi-modality MRI brain dataset, T1, T2, and fractional anisotropy (FA) MRI. The brain region was

segmented into white matter, grey matter, and cerebrospinal fluid. For each case, more than 1K patches were extracted at four different sizes;  $9 \times 9$ ,  $13 \times 13$ ,  $17 \times 17$  and  $22 \times 22$  for training and testing.

In [34], Tajbakhsh et al. demonstrated that fine-tuning of a pre-trained CNN in a layer-wise manner leads to incremental performance improvement in classification, detection, and segmentation. The performance evaluation was made for three different medical imaging modalities; polyp detection in colonoscopy images, pulmonary embolism detection in CT pulmonary angiography, and intima-media boundary segmentation from the ultrasound of carotid intima-media thickness. The training set comprised of stratified image patches with an equal number of positive and negative classes. In another work by Kooi et al. [37], automatic detection of breast lesion in mammography was investigated using 39,872 patches of size 250 × 250 for training the CNN model.

In this article, we evaluated the performance of four CNNs; FCN, SegNet, DeepLabV3+, and U-Net for segmentation of prostate T2W MRI trained in a patch-wise framework. The four CNNs are selected as they have given excellent performance in other applications including road scene, biomedical image, and natural image segmentation. Our contribution in this work includes a head-to-head comparison of the four CNNs, trained in the best possible framework using selected image pre-processing techniques, methods for increasing the number of images and its variation, and a method to tackle the issue of class imbalance. In particular, image resizing, center-cropping, and intensity normalization are applied for addressing the issues of inter-patient and inter-scanner variability as well as the issue of dominating background pixel over a prostate pixel. Issues related to small training images are addressed by using patch extraction while data augmentation is used to increase the variation in images and has improved the network accuracy. In addition, the class imbalance issue was addressed using the weighted cross-entropy loss function during network training. The selected methodology is proven to improve the segmentation accuracy measured in terms of DSC values.

#### 3. Material and Methods

In this section, we cover the methodology for training and testing of the CNNs in segmenting the whole prostate gland, peripheral zone, and central gland from T2W MRI of two different datasets. Firstly, a brief description of the four CNNs architecture is discussed. Secondly, the details of the two datasets used in this work are described, covering the scanner types, image dimensions, and class labels. In the third section, methods for image pre-processing, data augmentation, and patch extraction are discussed. Next, the technique of class balancing is presented as a measure to prevent the network from being biased toward classes with a larger number of pixels. Then, the weighted cross-entropy loss function used in the optimization of the network parameters is presented. Lastly, the Dice similarity coefficient (DSC) is presented as the performance metric for evaluating the CNN-based segmentation methods. The general framework for evaluation of the CNNs is illustrated in Figure 1.

#### 3.1. Deep CNN Architectures for Prostate Segmentation

The following deep CNN models which follow encoder–decoder structure are selected for comparative performance evaluation of semantic segmentation of prostate T2W MRI; FCN [24], U-Net [29], SegNet [40], and DeepLabV3+ [41]. A brief description of each CNN models is presented in the subsequent paragraphs.

Fully convolutional neural network (FCNs) [24,42,43] is based on existing image classification network such as AlexNet [44], VGG16 [45], and GoogleNet [46] to serve as the encoder module (downsampling path) and appended with a decoder module (upsampling path) with transposed convolutional layers. The encoder module is built from local connected layers such as convolution, pooling, and upsampling without any dense layer in the architecture. Since all connections are local, the network can work with any image size. The downsampling path and upsampling path will respectively capture both the context and localization or spatial information. However, pooling or downsampling will result in loss of spatial information. Skip connections are used to recover the spatial information in downsampling layers, hence combining context information with spatial information.

Specifically, the skip connections concatenate the feature map from downsampling with upsampling layers at different levels of resolution.



**Figure 1.** General methodology for evaluation of encoder–decoder CNN architectures in semantic segmentation of prostate T2W MRI.

Originally developed for biomedical image segmentation, U-Net is a variant of FCN that was modified to work on a few training images giving better segmentation performance [29,47]. The network has a U-shaped architecture that consists of an encoder-decoder network and interconnected by skip connections. The encoder extracts the spatial feature map from the input image using the convolutional layer, while the decoder performs the upsampling process to get segmented output from the encoder feature map. In feature map construction, a  $3 \times 3$  convolution operations are performed, followed by stride of two  $2 \times 2$  max pooling. This process consists of four convolutional operations and, after each downsampling, the number of filters is doubled in the next convolutional layer. The output feature map of  $3 \times 3$  convolution operation is forwarded to the decoder part which is upsampled using  $2 \times 2$  transposed convolution operation, reducing the feature map by half. The upsampling stage and two convolution operations are repeated four times, while at each stage the numbers of filters are halved. The information of feature maps from the convolutional operation of an encoder is transferred to the decoder before the pooling operation. The skip connections between the encoder–decoder networks are used to recover the information lost from the pooling operation. The final segmented output is obtained by using a  $1 \times 1$  convolution operation on the upsampled feature map. All convolutional layers in U-Net used rectified linear unit (ReLU) as the activation function except for the final convolution layer which used sigmoid activation function.

SegNet is a pixel-wise semantic segmentation technique originally developed for outdoor and indoor scene understanding [40,48]. The architecture consists of encoders using only forward connections from VGG architecture, non-trainable layers, and a corresponding set of decoders followed by a pixel-wise classifier. With no fully connected layer, this had reduced the network parameters from 134 M to 14.7 M, which makes it suitable for computationally efficient embedded systems. If the classification network is directly adopted for pixel-wise segmentation, the segmentation result will

be poor. This is resulted from the max pooling and subsampling operations that reduce the feature map resolution and hence reduce the output resolution. To maintain the same output image resolution as the input image, the low-resolution feature maps are upsampled by SegNet using memorized pooling indices. This enables the network to store boundary information and reduce the number of

trainable parameters.

DeepLabV3+ is another CNN developed for semantic segmentation that is comprised of the encoder–decoder network [41,49,50]. The highlights of the architecture are the use of atrous or dilated convolution and atrous spatial pyramid pooling (ASPP) [51]. To maintain the spatial resolution of the output feature map, DeepLabV3+ used the Xception model as their backbone and made use of depth-wise separable convolution in place of max pooling layers. Instead of using standard convolution, depth-wise separable convolution is used which segregate the operation into depth-wise and point convolution. Feature maps are generated from each input channel by depth wise convolution which applies separable convolution with (r - 1) zeros being placed within the successive filters. The output of a depth wise convolution is accumulated by utilizing point wise convolution. In the decoder part, the encoder feature map is upsampled by a factor of 4 and then concatenated with low-level features. Here, the  $1 \times 1$  convolution is applied for achieving reduction in the number of channels. Lastly,  $3 \times 3$  convolution is applied on a feature map and subsequently being upsampling by factor of 4 to get the output having equal size to the input.

## 3.2. MRI Dataset

Two MRI datasets are used for evaluation of the CNN architectures, from (1) NCI-ISBI 2013 challenge and (2) Universiti Kebangsaan Malaysia Medical Centre (UKMMC). The summary of the two datasets is as follows:

- NCI-ISBI 2013 is a prostate 3T collection of 40 patients having 542 T2-weighted MR slices available at the Cancer Imaging Archive (TCIA) site [52]. The images were collected at Radboud University Medical Center, Netherlands, using Siemens TIM 3 Tesla scanners [53]. This dataset has three different image sizes, 384 × 384, 320 × 320 and 256 × 256, having a thickness of 4 mm. The image masks are labeled to 3-class, peripheral zone (PZ), Central Gland (CG), and background.
- 2. UKMMC dataset is a prostate 3T collection consists of 11 patients having 229 T2-weighted MR slices that are collected from UKMMC. The images were acquired using a 3 Tesla Siemens TIM MRI scanner with a surface coil. The image dimensions are 384 × 384 and 320 × 320, with a thickness of 3 mm. The image masks are labeled to 2-class, whole prostate gland (WG), and background.

## 3.3. Image Pre-Processing

Since the two datasets used in this work, NCI-ISBI 2013 and UKMMC, were obtained using two different scanners, there is variation in terms of image dimensions and resolutions. Therefore, image normalization is necessary in order to work with datasets having a common dimension while reducing both inter-patient and inter-scan variability. Details on the pre-processing steps are given in the subsequent paragraphs.

# 3.3.1. Image Resizing

The images of NCI-ISBI 2013 dataset come in three sizes,  $384 \times 384$ ,  $320 \times 320$  and  $256 \times 256$ , whereas, in the UKMMC dataset, the images come in two sizes,  $384 \times 384$  and  $320 \times 320$ . Therefore, in order to achieve common image size, image resizing is performed by scaling the images to a size of ( $320 \times 320$ )-pixel using the nearest-neighbor interpolation method [54].

#### 3.3.2. Center-Cropping

The size of the prostate gland in the MRI scan is relatively small compared to the background, and this will result in background pixels to dominate over the prostate pixel resulting in poor prostate segmentation. Hence, center-cropping is necessary for reducing the number of background pixels [55,56], and to reduce the computation time with redundant pixels. Since the output size from image resizing is ( $320 \times 320$ )-pixels, the dimensions of these images are center cropped to ( $256 \times 256$ )-pixel. A sample of the center cropped image and its corresponding mask is shown in Figure 2.



**Figure 2.** Center-cropping of MRI prostate image, original MRI (**top-left**), center-cropped image (**bottom-left**), and the corresponding masks (**top/bottom-right**).

# 3.3.3. Intensity Normalization

Intensity normalization is used to reduce variation in the intensity distribution of images in one dataset which occurs due to inter-patient variability [57]. This will also improve the comparability of the images across different subjects and give better segmentation results. Hence, intensity normalization is applied to both datasets.

## 3.4. Data Augmentation

During the training of deep neural networks, data augmentation is performed to improve the accuracy of the training networks. Data augmentation will essentially increase the spatial variation of the images and improve the network accuracy [58]. The two data augmentation techniques used are random reflection and translation by  $\pm$  10 pixels in the *x*-axis and *y*-axis directions.

# 3.5. Patch Extraction

Training DNNs for accurate prostate segmentation requires large datasets, hence we used a patch-wise framework. The four different encoder/decoder networks are evaluated in terms of DSC for different sizes of patches, including  $64 \times 64$ ,  $128 \times 128$ , and  $192 \times 192$ , similar to the work of Sekou et al. [59]. It is observed that the performance of the CNNs decreases in the smaller patch sizes because of the small prostate region compared to the background region. In smaller patch sizes

of  $64 \times 64$  and  $128 \times 128$ , the prediction of the prostate gland is lower than in higher patch size of  $192 \times 192$ . This explains why the CNN architectures will result in better segmentation in the case of  $(192 \times 192)$ -patch than the  $(64 \times 64)$ -patch and  $(128 \times 128)$ -patch.

In our case, from each MRI slices, 20 and 30 ( $192 \times 192$ )-pixel random patches were generated from NCI-ISBI 2013 and UKMMC dataset, respectively. For NCI-ISBI 2013, 460 images were selected for training from a total of 542 MRI slices, giving 9200 images, a 20-fold increase in the number of training images. On the other hand, since the UKMMC is a relatively small dataset, more patches are needed to be generated from this dataset. From 176 MRI slices, patch extraction has generated 5280 total training images, which is a 30-fold increase from the original 176 slices. A sample of patches generated from one of the datasets is shown in Figure 3.





**Figure 3.** Sample of three (192  $\times$  192)-pixel patches (**b**-**d**) generated from the (256  $\times$  256)-pixel image in (**a**).

#### 3.6. Class Weight Balancing

It is common in medical volume scan that the anatomy of interest occupies only a small region of the scan, like in the case of prostate MRI. As a result, the number of pixels for image background is higher than the prostate gland, resulting in imbalance class labels. Specifically, for the NCI-ISBI 2013 dataset, the ratio of pixels for background to central gland and peripheral zone is 95.5:1.9:2.6, whereas, in the UKMMC dataset, the ratio of background to whole gland is at 94.94:5.06. Class imbalance can be detrimental for the learning process as the network learns more on dominant classes and biases towards these dominant classes. To improve the learning process and avoid being trapped in local minima, the class weighting approach can be used as a systematic way to improve the low dominant class [60]. To counteract class imbalance present in the dataset, larger weights are assigned for labels with less total pixels and smaller weights for labels with more total pixels.

If the number of pixels in a particular class is denoted as  $N_c$ , where *c* corresponds to background, CG and PZ for 3-class dataset, and *c* corresponds to background and WG for 2-class dataset. Hence,  $N_c$  is  $(1 \times 3)$ -array and  $(1 \times 2)$ -array for 3-class dataset and 2-class dataset, respectively. Let *T* represent the total number of pixels in the images, then the  $(1 \times 2)$  or  $(1 \times 3)$ -array of image frequency,  $F_c$ , of a class is the ratio of  $N_c$  to *T* given by

$$F_c = \frac{N_c}{T}.$$
(1)

From here, the  $(1 \times 2)$  or  $(1 \times 3)$ -array of class weight,  $W_c$ , for a set of training data can be calculated by finding the ratio of median of  $F_c$  to  $F_c$ 

$$W_c = \frac{median(F_c)}{F_c}.$$
 (2)

Histograms of pixel distribution of 3-class NCI-ISBI 2013 dataset and 2-class UKMMC dataset, before and after applying class weighting approach, are shown respectively in Figures 4 and 5.



Figure 4. Pixel distribution of (a) original images and (b) after class weighting for the NCI-ISBI 2013 dataset.



Figure 5. Pixel distribution of (a) original images and (b) after class weighting for the UKMMC dataset.

# 3.7. Loss Function

In this work, the weighted cross-entropy (WCE) loss function is selected for the CNNs because it helps the network to better differentiate between background and prostate pixels, and is especially handy when there is class imbalance. The WCE loss function penalizes each class based on the median frequency of each class, which is formulated as follows [61]:

$$WCE = -\frac{1}{n} \sum_{i=1}^{n} W_{c,i} \left[ T_i \log P_i + (1 - T_i) \log(1 - P_i) \right],$$
(3)

where the sum is run over all training images, *n*. The variable  $P_i$  is the predicted segmentation class,  $T_i$  is the target or the ground truth segmentation label, and  $W_{c,i}$  is the class weight calculated from Equation (2).

#### 3.8. Evaluation Metric

The performance of the CNNs for segmentation of prostate T2W MRI is evaluated based on intersection between *target* and *prediction* sets using dice similarity coefficient (DSC) [62]. In essence, DSC measures the resemblance of elements between *prediction* and *target* sets given as

$$DSC = 2 \frac{|T \cap P|}{|T| + |P|} \times 100,$$
 (4)

where |T| and |P| represent the number of elements in target and prediction sets, respectively. The DSC values range between 0 to 1, with 1 indicating the best-case scenario.

## 4. Results and Discussion

#### 4.1. Experimental Set-up

The code for training and testing the CNNs model is developed on the MATLAB 2019b platform, using a PC equipped with 32 GB memory, Intel Core i7-8700k CPU, GeForce RTX 2080Ti graphic card with 32 GB memory and 10.1 Cuda edition. During the training, the network weight is updated by the Adam optimizer [63] with learning rate selected at 0.0001 and minibatch size equal to 16. Using patch extraction, total training images generated from NCI-ISBI 2013 and UKMMC dataset are 9200 and 5280, respectively. To ensure that the trained CNNs are well generalized, 5-fold cross validation strategy is utilized [64]. Here, the data are divided into 80% for training and 20% for testing the networks [64,65]. The performance of the CNNs is obtained from an average value of the 5-fold cross-validation.

#### 4.2. Quantitative and Qualitative Performance

The segmentation performance by the CNNs is quantitatively evaluated in terms of DSC values, averaged over all test images. The DSC values for datasets NCI-ISBI 2013 and UKMMC are shown in Tables 1 and 2, respectively. The results are presented for both without patch extraction and with patch extraction. Subsequent paragraphs provide an analysis of segmentation results for both datasets.

**Table 1.** DSC values (in percentage) of prostate peripheral zone (PZ) and central gland (CG) segmentation by FCN, SegNet, U-Net, and DeepLabV3+ for test images from the NCI-ISBI 2013 dataset.

	Withou	ıt Patch	Patch-Wise		
	PZ	CG	PZ	CG	
FCN	69.2±5.8	85.1±4.8	72.7±5.1	88.6±2.9	
SegNet	$75.3 {\pm} 4.8$	$88.9 \pm 3.8$	$76.0 \pm 3.9$	90.8±1.2	
U-Net	$76.3 \pm 3.5$	89.8±2.6	$76.8 \pm 3.3$	$91.6 {\pm} 1.4$	
DeepLabV3+	$76.4 \pm 3.3$	90.7±1.2	78.9±1.9	92.8±0.7	
L. Rundo [32]	-	-	$76.0{\pm}4.1$	91.5±3.2	

**Table 2.** DSC values (in percentage) of prostate whole gland (WG) segmentation by FCN, SegNet, U-Net, and DeepLabV3+ for test images from the UKMMC dataset.

	Without Patch	Patch-Wise
FCN	76.7±6.1	$86.6 {\pm} 4.8$
SegNet	$79.8 {\pm} 5.7$	$84.3 \pm 4.2$
U-Net	$85.0{\pm}3.6$	$88.4 \pm 3.7$
DeepLabV3+	$83.2 {\pm} 4.7$	$91.9{\pm}2.0$

The encoder–decoder CNNs are trained and tested in two scenarios: (1) without patch extraction using ( $256 \times 256$ )-pixel input images and (2) patch-wise extraction using ( $192 \times 192$ )-pixel patch size

input images. The DSC values for NCI-ISBI 2013 dataset for FCN, SegNet, U-Net, and DeepLabV3+ are presented in Table 1. The NCI-ISBI 2013 is a 3-class labeled data where the prostate regions are divided into two regions, peripheral zone (PZ) and central gland (CG). The DSC values in Table 1 indicate that the patch-wise method will result in better segmentation mainly because it has increased the amount of training data hence, enriching the deep neural network with more data. In particular, using patch-wise implementation, average DSC gain, calculated from the four CNNs is 2.5% and 2.1% for PZ and CG, respectively. In addition, the results clearly show that DeepLabV3+ gives the best prostate segmentation with the highest DSC values of 78.9% and 92.8% for PZ and CG, respectively. In general, the better performance for the segmentation of CG compared to PZ by the CNNs is due to the higher ratio of CG volume to PZ volume in the NCI-ISBI 2013 dataset.

As for qualitative analysis, we present a sample of segmented images as shown in Figure 6 for both without and with the patch-wise method for PZ and CG regions. Visual comparison between without patch and with patch-wise indicates that more accurate segmentation can be achieved by patch-wise method. This highlights the significance of patch extraction which has enriched the CNNs with more data and improves its accuracy. It is notable from Figure 6 that the mask generated by the best CNN, DeepLabV3+ is closely similar in shape and size to the ground truth mask. In addition, DeepLabV3+ can accurately segment the region around the prostate boundary. On the other hand, FCN, SegNet, and U-Net are also able to segment prostate into PZ and CG regions but with less accurate segmentation, especially around the pixels close to the prostate boundary.

Similar trends of performance are seen in the 2-class labeled UKMMC dataset, as given in Table 2. Here, the ground truth mask is only available for the whole prostate gland (WG). Again, the best DSC value of 91.9%, is generated by DeepLabV3+. Overall, the average improvement of the patch-wise method over without patch is by 8.7%. For qualitative analysis, visual inspection on the segmented image as shown in Figure 7 is performed. From Figure 7, it is clear that DeepLabV3+ is better than FCN, SegNet, and U-Net in segmenting the WG. Notably, FCN, SegNet, and U-Net are not able to accurately segment the WG, especially around the upper and lower sections of the WG.

To further confirm the better performance of patch-wise DeepLabV3+ over three other CNNs, a significance test is performed using the Wilcoxon signed-rank test generating *p*-values for the DeepLabV3+ versus the rest. The significance level is set at 0.05 and a null hypothesis is that the DeepLabV3+ versus three other CNNs will not improve the DSC values of the segmented images. The *p*-values tabulated in Table 3 for both NCI-ISBI 2013 and UKMMC datasets are all less than 0.05 which imply that we can reject the null hypothesis. In other words, this gives the indication that patch-wise DeepLabV3+ can better perform prostate segmentation than U-Net, SegNet, and FCN when measured in terms of DSC values.

	NCI-ISBI 2013	UKMMC
DeepLabV3+ vs. U-Net	0.025	0.041
DeepLabv3+ vs. SegNet	0.022	0.046
DeepLabV3+ vs. FCN	0.016	0.035

**Table 3.** *P*-value of DeepLabV3+ versus U-Net, SegNet and FCN for NCI-ISBI-2013 and UKMMC datasets, using patch-wise implementation.

The better performance of the DeepLabV3+ can be attributed to its architecture that uses the Xception model as the backbone and making use of atrous separable convolution for retaining the spatial resolution of the resultant feature maps. This allows the network to extract denser feature map by controlling the field view for precise localization and assimilation of context. In addition, utilizing atrous convolution enables the network to precisely define area with arbitrary scale by rebuilding them at various scales in atrous spatial pyramid pooling (ASPP). This has essentially resulted in a faster and efficient encoder–decoder network that covers better spatial information at multiple scales as compared to other encoder–decoder networks.

Performance comparison of DeepLabV3+ in terms of DSC values with some selected techniques including the state-of-the-art technique is summarized in Table 4. Direct comparison of DeepLabV3+ with other published techniques is difficult due to different in datasets, scanning protocols as well as numbers of images. Rough Comparison with other prostate segmentation techniques proposed in the past are listed in Table 4. If we consider the segmentation of WG by [19,20,30,31], the best DSC value is 89.4% by Yu et al. [31], which employed residual connection in U-Net. Our patch-wise DeepLabV3+ using a UKMMC dataset has resulted in a DSC value of 91.9%, which is 2.5% better than [31]. This may not be a completely fair comparison since it was not using the same dataset. However, this comparison may provide valuable insight on what is the best encoder-decoder segmentation method. Now, let's consider segmentation of PZ and CG proposed by Litjens et al. [23] and Rundo et al. [32]. Then, comparing the DSC values indicate that our deep-wise DeepLabV3+ gave better DSC by 4.8% (PZ) and 3.7% (CG) than the multi atlas-based segmentation [23]. For more fair comparison, we compared our deep-wise DeepLabV3+ with the recently published state-of-the-art method proposed by Rundo et al. [32] which uses squeeze-and-excitation blocks in encoders and decoders of U-Net architecture. The deep-wise DeepLabV3+ gave a higher DSC value than the U-Net method in [32], with an increase of 2.9% (PZ) and 1.3% (CG).



**Figure 6.** Segmentation of prostate T2W MRI, from NCI-ISBI 2013 dataset into PZ and CG regions. (a) Original image, (b) ground truth boundary, where predicted boundary (red), overlay on the ground truth (blue), based on the method without patch extraction using, (c) FCN, (d) SegNet, (e) U-Net and (f) DeepLabV3+, and based on patch extraction using (g) FCN, (h) SegNet (i) U-Net and (j) DeepLabV3+.

Work	Images	Dataset	Network	DSC (WC) or (P7	
	intages	Dutuset	THE WORK	CG)	
Klein et al. 2008 [20]	50 volumes	3T MRI	Atlas matching using localized information	84.4	
Guo et al. 2015 [19]	66 images	University of Chicago Hospital	Deformable model using SSAE	87.1	
Litjen et al. 2012 [23]	48 images	3T MRI	Multi atlas-based segmentation	74, 89	
Milletari et al. 2016 [30]	80 volumes	PROMISE12	V-Net	86.9	
Yu et al. 2015 [31]	80 volumes	PROMISE12	ConvNet with residual connections	89.4	
Rundo et al. 2019 [32]	40 images	NCI-ISBI 2013	U-Net with SE block	$76.0{\pm}4.1, 91.5{\pm}3.2$	
Our Work	40 images	NCI-ISBI 2013	Patch-wise DeepLabV3+	$78.9{\pm}1.9, 92.8{\pm}0.7$	
Our Work	11 images	UKMMC	Patch-wise DeepLabV3+	91.9±2.0	

Table 4.	Comparison	of DSC	values	(in	percentage)	of	atlas-based,	deformable	model-based,
CNN-based, and our patch-wise DeepLabV3+.									



(a)

(**d**)



(e)

(**f**)



Figure 7. Semantic segmentation of WG prostate T2W MRI, from UKMMC dataset. (a) Original image, (b) ground truth boundary where predicted boundary (red), overlay on the ground truth (blue), based on the method without patch extraction using, (c) FCN, (d) SegNet, (e) U-Net and (f) DeepLabV3+, and based on patch extraction using (g) FCN, (h) SegNet (i) U-Net and (j) DeepLabV3+.

# 5. Conclusions

(c)

In this paper, the evaluation of four encoder-decoder CNNs, including FCN, SegNet, U-Net, and DeepLabV3+ for segmentation of prostate T2W MRI is presented. The proposed methodology for the investigation includes methods for image pre-processing, patch extraction, data augmentation, and class weight balancing. These techniques have resulted in the normalized dataset, increased training images as well as its spatial variation, and reduced the effect of imbalance class. The selection of appropriate techniques prior to training the network is essential to achieve a well-trained network that gives good segmentation accuracy. Comparison with atlas-based, deformable model based, CNN-based, and state-of-the-art techniques has proven that patch-wise DeepLabV3+ gave better performance in segmentation of prostate WG, PZ, or CG. In summary, the best segmentation was achieved by a DeepLabV3+ model, which can be attributed to its atrous convolution operations and ASSP architecture. The accuracy can be further improved with more augmented data by injecting a generative adversarial network (GAN) to the deep models.

**Author Contributions:** N.Y. and F.M. devised the project, the main conceptual ideas and proof outline, as well as project planning and supervised the work. Z.K. contributed to data preprocessing, investigation, training, testing the algorithm and wrote the original manuscript. N.Y. secured the funding and worked on the manuscript to further improve the outline, technical details and result analysis. Z.K., F.M. and K.A. contributed to the design and implementation of the research, to the analysis of the results. S.S.A.A. reviewed the manuscript and provided valuable suggestions to further refine the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is supported by two research grants: (1) International Collaborative Research Fund between Universiti Teknologi PETRONAS (UTP) and Universitas Muhammadiyah Purwokerto (UMP) under grant number 015ME0-098, and (2) the Yayasan Universiti Teknologi PETRONAS under grant number YUTP-FRG 015LC0-292.

**Acknowledgments:** The authors would like to thank Hilwati Hashim from the UiTM Sungai Buloh Hospital and Syazarina Sharis Osman from the HUKM Medical Centre for helping with the manual prostate delineation.

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

## References

- Ferlay, J.; Colombet, M.; Soerjomataram, I.; Mathers, C.; Parkin, D.; Piñeros, M.; Znaor, A.; Bray, F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int. J. Cancer* 2019, 144, 1941–1953. [CrossRef] [PubMed]
- 2. Rawla, P. Epidemiology of Prostate Cancer. World J. Oncol. 2019, 10, 63. [CrossRef]
- 3. Gandellini, P.; Folini, M.; Zaffaroni, N. Emerging role of microRNAs in prostate cancer: Implications for personalized medicine. *Discov. Med.* **2010**, *9*, 212–218. [PubMed]
- 4. Loeb, S.; Catalona, W.J. What to do with an abnormal PSA test. *The Oncologist* **2008**, *13*, 299–305. [CrossRef] [PubMed]
- Ogden, J.M.; O'Keefe, S.J.; Marks, I.S. Development of a method for the purification of human trypsin by single step affinity chromatography suitable for human isotope incorporation studies. *Clin. Chim. Acta* 1992, 212, 141–147. [CrossRef]
- 6. Backer, H. Prostate cancer screening: Exploring the debate. *Permanente J* 1999, *3*, 330–340.
- 7. Turkbey, B.; Choyke, P.L. Multiparametric MRI and prostate cancer diagnosis and risk stratification. *Curr. Opin. Urol.* **2012**, *22*, 310. [CrossRef]
- 8. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]
- 9. Barentsz, J.O.; Richenberg, J.; Clements, R.; Choyke, P.; Verma, S.; Villeirs, G.; Rouviere, O.; Logager, V.; Fütterer, J.J. ESUR prostate MR guidelines 2012. *Eur. Radiol.* **2012**, *22*, 746–757. [CrossRef]
- 10. McNeal, J.E.; Redwine, E.A.; Freiha, F.S.; Stamey, T.A. Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread. *Am. J. Surg. Pathol.* **1988**, *12*, 897–906. [CrossRef]
- 11. Muller, B.G.; Shih, J.H.; Sankineni, S.; Marko, J.; Rais-Bahrami, S.; George, A.K.; de la Rosette, J.J.; Merino, M.J.; Wood, B.J.; Pinto, P.; et al. Prostate cancer: Interobserver agreement and accuracy with the revised prostate imaging reporting and data system at multiparametric MR imaging. *Radiology* **2015**, 277, 741–750. [CrossRef] [PubMed]

- Jia, H.; Xia, Y.; Song, Y.; Cai, W.; Fulham, M.; Feng, D.D. Atlas registration and ensemble deep convolutional neural network-based prostate segmentation using magnetic resonance imaging. *Neurocomputing* 2018, 275, 1358–1369. [CrossRef]
- Fasihi, M.S.; Mikhael, W.B. Overview of current biomedical image segmentation methods. In Proceedings of the 2016 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 15–17 December 2016; pp. 803–808.
- 14. Vincent, G.; Guillard, G.; Bowes, M. Fully automatic segmentation of the prostate using active appearance models. *MICCAI Grand Chall. Prostate MR Image Segmentation* **2012**, 2012, 2.
- 15. Kirschner, M.; Jung, F.; Wesarg, S. Automatic prostate segmentation in MR images with a probabilistic active shape model. In Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Nice, France, 1 October 2012; pp. 1–8.
- Cheng, R.; Roth, H.R.; Lu, L.; Wang, S.; Turkbey, B.; Gandler, W.; McCreedy, E.S.; Agarwal, H.K.; Choyke, P.; Summers, R.M.; et al. Active appearance model and deep learning for more accurate prostate segmentation on MRI. *Proc. SPIE* 2016, *9784*. [CrossRef]
- Martin, S.; Troccaz, J.; Daanen, V. Automated segmentation of the prostate in 3D MR images using a probabilistic atlas and a spatially constrained deformable model. *Med. Phys.* 2010, 37, 1579–1590. [CrossRef] [PubMed]
- Zhang, J.; Baig, S.; Wong, A.; Haider, M.A.; Khalvati, F. Segmentation of prostate in diffusion MR images via clustering. In Proceedings of the International Conference on Image Analysis and Recognition (ICIAR), Springer, At Montreal, QC, Canada, 2 June 2017; pp. 471–478.
- 19. Guo, Y.; Gao, Y.; Shen, D. Deformable MR prostate segmentation via deep feature learning and sparse patch matching. *IEEE Trans. Med. Imaging* **2015**, *35*, 1077–1089. [CrossRef] [PubMed]
- 20. Klein, S.; Van Der Heide, U.A.; Lips, I.M.; Van Vulpen, M.; Staring, M.; Pluim, J.P. Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Med. Phys.* **2008**, 35, 1407–1417. [CrossRef]
- 21. Langerak, T.R.; van der Heide, U.A.; Kotte, A.N.; Viergever, M.A.; Van Vulpen, M.; Pluim, J.P. Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE). *IEEE Trans. Med. Imaging* **2010**, *29*, 2000–2008. [CrossRef]
- 22. Dowling, J.A.; Fripp, J.; Chandra, S.; Pluim, J.P.W.; Lambert, J.; Parker, J.; Denham, J.; Greer, P.B.; Salvado, O. Fast automatic multi-atlas segmentation of the prostate from 3D MR images. In Proceedings of the International Workshop on Prostate Cancer Imaging, Toronto, ON, Canada, 22 September 2011; pp. 10–21.
- 23. Litjens, G.; Debats, O.; van de Ven, W.; Karssemeijer, N.; Huisman, H. A pattern recognition approach to zonal segmentation of the prostate on MRI. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Nice, France, 1–5 October 2012; pp. 413–420.
- 24. 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.
- 25. Tian, Z.; Liu, L.; Zhang, Z.; Fei, B. PSNet: Prostate segmentation on MRI based on a convolutional neural network. *J. Med. Imaging* **2018**, *5*, 021208. [CrossRef]
- 26. Everingham, M.; Van Gool, L.; Williams, C.K.; Winn, J.; Zisserman, A. The pascal visual object classes (voc) challenge. *Int. J. Comput. Vis.* **2010**, *88*, 303–338. [CrossRef]
- 27. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? *Adv. Neural Inf. Process. Syst.* **2014**, 3320–3328.
- 28. Zeiler, M.D.; Fergus, R. Visualizing and understanding convolutional networks. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 818–833.
- 29. 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; pp. 234–241.
- Milletari, F.; Navab, N.; Ahmadi, S.A. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016, pp. 565–571.

- Yu, L.; Yang, X.; Chen, H.; Qin, J.; Heng, P.A. Volumetric convnets with mixed residual connections for automated prostate segmentation from 3D MRI images. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
- 32. Rundo, L.; Han, C.; Nagano, Y.; Zhang, J.; Hataya, R.; Militello, C.; Tangherloni, A.; Nobile, M.S.; Ferretti, C.; Besozzi, D.; et al. USE-Net: Incorporating Squeeze-and-Excitation blocks into U-Net for prostate zonal segmentation of multi-institutional MRI datasets. *Neurocomputing* **2019**, *365*, 31–43. [CrossRef]
- Zhang, W.; Li, R.; Deng, H.; Wang, L.; Lin, W.; Ji, S.; Shen, D. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *NeuroImage* 2015, 108, 214–224. [CrossRef] [PubMed]
- Tajbakhsh, N.; Shin, J.Y.; Gurudu, S.R.; Hurst, R.T.; Kendall, C.B.; Gotway, M.B.; Liang, J. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Trans. Med. Imaging* 2016, 35, 1299–1312. [CrossRef] [PubMed]
- 35. Pereira, S.; Pinto, A.; Alves, V.; Silva, C.A. Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Trans. Med. Imaging* **2016**, *35*, 1240–1251. [CrossRef] [PubMed]
- Moeskops, P.; Viergever, M.A.; Mendrik, A.M.; de Vries, L.S.; Benders, M.J.; Išgum, I. Automatic segmentation of MR brain images with a convolutional neural network. *IEEE Trans. Med. Imaging* 2016, 35, 1252–1261. [CrossRef] [PubMed]
- Kooi, T.; Litjens, G.; Van Ginneken, B.; Gubern-Mérida, A.; Sánchez, C.I.; Mann, R.; den Heeten, A.; Karssemeijer, N. Large scale deep learning for computer aided detection of mammographic lesions. *Med. Image Anal.* 2017, *35*, 303–312. [CrossRef]
- Milletari, F.; Ahmadi, S.A.; Kroll, C.; Plate, A.; Rozanski, V.; Maiostre, J.; Levin, J.; Dietrich, O.; Ertl-Wagner, B.; Bötzel, K.; et al. Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound. *Comput. Vis. Image Underst.* 2017, 164, 92–102. [CrossRef]
- 39. Liu, Y.; Ren, Q.; Geng, J.; Ding, M.; Li, J. Efficient Patch-Wise Semantic Segmentation for Large-Scale Remote Sensing Images. *Sensors* **2018**, *18*, 3232. doi:10.3390/s18103232. [CrossRef]
- 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]
- 41. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818.
- 42. Chun, C.; Ryu, S.K. Road Surface Damage Detection Using Fully Convolutional Neural Networks and Semi-Supervised Learning. *Sensors* **2019**, *19*. doi:10.3390/s19245501. [CrossRef] [PubMed]
- Islam, M.M.M.; Kim, J.M. Vision-Based Autonomous Crack Detection of Concrete Structures Using a Fully Convolutional Encoder–Decoder Network. *Sensors* 2019, *19*, 4251. doi:10.3390/s19194251. [CrossRef] [PubMed]
- 44. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, 1097–1105. [CrossRef]
- 45. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
- Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
- 47. Zhao, X.; Yuan, Y.; Song, M.; Ding, Y.; Lin, F.; Liang, D.; Zhang, D. Use of Unmanned Aerial Vehicle Imagery and Deep Learning UNet to Extract Rice Lodging. *Sensors* **2019**, *19*, 3859, doi:10.3390/s19183859. [CrossRef]
- Zhu, Y.; Luo, K.; Ma, C.; Liu, Q.; Jin, B. Superpixel Segmentation Based Synthetic Classifications with Clear Boundary Information for a Legged Robot. *Sensors* 2018, *18*, 2808, doi:10.3390/s18092808. [CrossRef]
- Yao, X.; Yang, H.; Wu, Y.; Wu, P.; Wang, B.; Zhou, X.; Wang, S. Land Use Classification of the Deep Convolutional Neural Network Method Reducing the Loss of Spatial Features. *Sensors* 2019, 19, 2792, doi:10.3390/s19122792. [CrossRef]
- 50. Lobo Torres, D.; Queiroz Feitosa, R.; Nigri Happ, P.; Elena Cué La Rosa, L.; Marcato Junior, J.; Martins, J.; Olã Bressan, P.; Gonçalves, W.N.; Liesenberg, V. Applying Fully Convolutional Architectures for Semantic Segmentation of a Single Tree Species in Urban Environment on High Resolution UAV Optical Imagery. *Sensors* 2020, 20, 563, doi:10.3390/s20020563. [CrossRef]

- 51. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 834–848. [CrossRef]
- 52. Prior, F.; Smith, K.; Sharma, A.; Kirby, J.; Tarbox, L.; Clark, K.; Bennett, W.; Nolan, T.; Freymann, J. The public cancer radiology imaging collections of The Cancer Imaging Archive. *Sci. Data* **2017**, *4*, 170124. [CrossRef]
- 53. Bloch, N.; Madabhushi, A.; Huisman, H.; Freymann, J.; Kirby, J.; Grauer, M.; Enquobahrie, A.; Jaffe, C.; Clarke, L.; Farahani, K. NCI-ISBI 2013 challenge: Automated segmentation of prostate structures. *Cancer Imaging Arch.* 2015, 370.
- 54. Bovik, A.C. The Essential Guide to Image Processing; Elsevier: Amsterdam, The Netherlands, 2009.
- 55. Feng, X.; Qing, K.; Tustison, N.J.; Meyer, C.H.; Chen, Q. Deep convolutional neural network for segmentation of thoracic organs-at-risk using cropped 3D images. *Med. Phys.* **2019**, *46*, 2169–2180. [CrossRef] [PubMed]
- 56. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. *arXiv* **2014**, arXiv:1409.4842.
- 57. Zhou, X.Y.; Yang, G.Z. Normalization in training U-Net for 2-D biomedical semantic segmentation. *IEEE Robot. Autom. Lett.* **2019**, *4*, 1792–1799. [CrossRef]
- Zhu, Q.; Du, B.; Turkbey, B.; Choyke, P.L.; Yan, P. Deeply-supervised CNN for prostate segmentation. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), San Francisco, CA, USA, 4–9 February 2017; pp. 178–184.
- 59. Sekou, T.B.; Hidane, M.; Olivier, J.; Cardot, H. From patch to image segmentation using fully convolutional networks-application to retinal images. *arXiv* **2019**, arXiv:1904.03892.
- Dhivya, J.J.; Ramaswami, M. A Perusal Analysis on Hybrid Spectrum Handoff Schemes in Cognitive Radio Networks. In Proceedings of the International Conference on Intelligent Systems Design and Applications, Vellore, India, 6–8 December 2018; pp. 312–321.
- 61. Christ, P.F.; Elshaer, M.E.A.; Ettlinger, F.; Tatavarty, S.; Bickel, M.; Bilic, P.; Rempfler, M.; Armbruster, M.; Hofmann, F.; D Anastasi, M.; et al. Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 17–21 October 2016; pp. 415–423.
- 62. Litjens, G.; Toth, R.; van de Ven, W.; Hoeks, C.; Kerkstra, S.; van Ginneken, B.; Vincent, G.; Guillard, G.; Birbeck, N.; Zhang, J.; et al. Evaluation of prostate segmentation algorithms for MRI: The PROMISE12 challenge. *Med. Image Anal.* **2014**, *18*, 359–373. [CrossRef]
- 63. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. arxiv 2014.
- 64. Gandhi, T.; Panigrahi, B.K.; Bhatia, M.; Anand, S. Expert model for detection of epileptic activity in EEG signature. *Expert Syst. Appl.* **2010**, *37*, 3513–3520. [CrossRef]
- 65. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI), San Francisco, CA, USA, 1995; pp. 1137–1145.



© 2020 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 (http://creativecommons.org/licenses/by/4.0/).