International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 893-898 www.ijaem.net ISSN: 2395-5252

Advanced CNN Approach to Predict Water Body Segmentation over Satellite Images

B. Santhosh Kumar[1], Ch. Sripriya[2], B. Reethika Devi[3], Sk. Reshma[4], N. Chandra Lahari[5]

[1]*Assistant Professor, Dept. of CSE, Sai Spurthi Institute of Technology,Khammam,Telangana,India*
[2,3,4,5,]*B.Tech Student, Dept. of CSE, Sai Spurthi Institute of Technology, Khammam,Telangana,India*

Submitted: 01-06-2022 Revised: 10-06-2022 Accepted: 15-06-2022

ABSTRACT: Extricating water-bodies precisely is an extraordinary test from exceptionally high goal (VHR) remote detecting symbolism. The limits of a water body are usually difficult to recognize due to the complex unearthly combinations brought about by sea-going vegetation, particular lake/stream tones, sediments close the bank, shadows from the encompassing tall plants, etc. The variety and semantic data of elements should be expanded for a superior extraction of water-bodies from VHR remote detecting pictures. In this paper, we address these issues by planning a novel multi-include extraction furthermore mix module. This module comprises of three component extraction sub-modules dependent on spatial and divert connections in highlight maps at each scale, which separate the total objective data from the nearby space, bigger space, and between-channel relationship to accomplish a rich highlight portrayal. At the same time, to more readily anticipate the fine shapes of water-bodies, we take on a multi-scale forecast combination module. Plus, to settle the semantic irregularity of element combination between the encoding stage and the disentangling stage, we apply an encoder-decoder semantic highlight combination module to advance combination impacts. We complete broad analyses in VHR flying and satellite symbolism separately. The outcome shows that our technique accomplishes best in class division execution, outperforming the work of art and ongoing techniques. Besides, our proposed technique is vigorous in testing water-body extraction situations.

Keywords: water-body segmentation; multi-feature extraction and combination; aerial and satellite imagery; fully convolutional network.

I. INTRODUCTION:

Water-body extraction is of incredible importance in water assets checking, normal calamity appraisal and ecological security. These applications depend on the evaluation of the water-body change. Precisely getting water-body division from remote detecting pictures is a significant mission for checking water body changes. In this paper, our point is to precisely depict water-bodies in confounded and testing situations from exceptionally high goal (VHR) remote detecting symbolism. Instruments installed satellites and ethereal vehicles give remote detecting symbolism that covers enormous scope water surface on Earth. As the forms of water-body in VHR remote detecting pictures are regularly indistinct. Such corruptions are commonly brought about by oceanic vegetation impeding, sediments/boats close to the bank and shadows from the encompassing tall plants. The particular tones are generally brought about by imaging conditions, water quality what's more microorganisms. Consequently, it is an incredible test to separate the diagram of water-bodies precisely in complex scenes from VHR remote detecting imageryare.

Convolutional neural organization (CNN) has shown wonderful execution in picture grouping, target identification and semantic division, respectable to the solid include portrayal capacity of CNN. Long et al. first proposed the completely convolutional network (FCN), which replaces the last completely associated layers with convolutional ones to accomplish start to finish semantic division. In the future, FCNs in a start to finish way broadly applied and widely created, turning into a standard innovation in semantic division and edge discovery. Ronneberger et al. planned a contracting way and a symmetric growing way to blend distinctive semantic highlights for biomedical picture division. Lin et al. utilized the component data accessible in the down-examining process and utilized significant distance leftover associations with accomplish high-goal expectation. Yu et al. proposed a start to finish

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 893-898 www.ijaem.net ISSN: 2395-5252

profound semantic edge learning design for classification mindful semantic edge recognition. Bertasius et al.
introduced a multi-scale bifurcated profound organization, which took advantage of article related elements as undeniable level signals for shape discovery. Xie et al. fostered a novel convolutional neural-network-based edge discovery framework by consolidating multi-scale and staggered visual reactions.

As of late, profound taking in based water-body division from remote detecting symbolism has drawn in some consideration and advancements. Yu et al. pioneers at presenting a CNN-based strategy for water-body extraction from Landsat symbolism by considering both otherworldly and spatial data. Be that as it may, this CNN-based strategy cut a picture into little tiles for pixel-level expectations, which presented a ton of excess and is of low productivity. Miao et al. proposed a confined responsive field deconvolution organization to remove water bodies from high-goal remote detecting pictures. Li et al. embraced a common FCN model to remove water bodies from VHR pictures and altogether beat the standardized contrast water record (NDWI) based strategy, the help vector machine (SVM) based strategy, and the sparsity model (SM) based technique. In any case, these two methodologies didn't consider the multi-scale data from various decoder layers and the channel relationship of element maps in the encoder, which joined inadequate extraction of water bodies in complex scenes. Duan et al. proposed a book multi-scale refinement organization (MSR-Net) for water-body division, which made full utilization of the multi-scale highlights for more precise division. Be that as it may, the MSR-Net doesn't reuse significant level semantic data and the multi-scale module it has doesn't consider channel connections between highlight maps. Guo et al. took on a basic FCN-based strategy for water-body extraction and introduced a multi-scale include extractor, incorporating four widened convolutions with various rates, which was sent on top of the encoders. This FCN-based strategy basically utilized the multi-scale data of significant level semantic elements, however didn't extricate total highlights at different scales. It is clear that flow FCN-based water extraction concentrates on stressed component extraction also expectation enhancement, however the space for additional upgrades is extensive. Include combination in the FCN-based technique ideally consolidates high-semantic highlights and highlights with exact areas, which works with water-body distinguishing proof and the precise extraction of water-body edges. In this work, we plan our technique by thinking about three perspectives: include extraction, forecast advancement, and the element combination of shallow and profound layers.

II. METHODOLOGY:

In this part, we give the subtleties of our proposed multi-include extraction and mix organization (MECNet) for water-body division from separately ethereal furthermore satellite Imagery. From the get go, we present our proposed MECNet engineering. Then, at that point, we depict a multi-include extraction and blend (MEC) module to accomplish more extravagant and more different highlights and further developed semantic data. Consequently, to better anticipate the fine form of the water-body, we plan a multi-scale expectation combination (MPF) module to coordinate the forecast results at three unique scales. Finally, we present an encoder-decoder semantic element combination (DSFF) module to conquer the issue of semantic irregularity among encoder and decoder.

2.1. MECNet Architecture

The MECNet chiefly comprises of three modules. We initially plan a multi-highlight extraction and mix module to get more extravagant and more assorted highlights in the encoding stage. The proposed MEC module comprises of three diverse element extraction sub-modules to show the spatial and channel connections between highlight maps. These sub-modules are (1) a neighborhood highlight extraction sub-module, (2) a bigger open field featureextraction sub-module, and (3) a between-channel include extraction sub-module. To address the semantic irregularity of elements from the encoding stage and the translating stage, an encoder-decoder semantic component combination module is set up. A straightforward multi-scale expectation combination module utilizes the forecast results from three distinct scales as contribution to get very fine water-body division shapes.

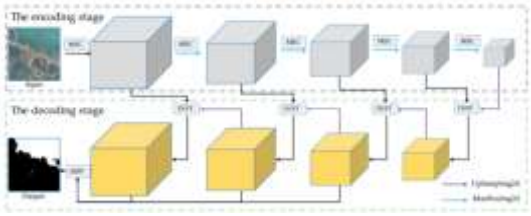

Fig 1. An overview of our proposed Multi-feature Extraction and Combination Network (MECNet)

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 893-898 www.ijaem.net ISSN: 2395-5252

2.2. Multi-Feature Extraction and Combination Module

The MEC module is made out of three sub-modules, to be specific a nearby component extraction sub-module (LFE), a more extended open field include extraction sub-module (LRFE) and a include extraction sub module for between-channel highlight improvement (CFE). The LFE what's more LRFE sub-modules depend on the spatial relations of element maps (i.e., from various responsive field scenes), and the CFE sub-module is intended to get additional rich component data by demonstrating the connections between channels of element maps.

The LFE sub-module, as displayed in Figure 3b, is intended to learn include maps recording nearby data. In particular, we play out a 3×3 convolution with a bunch standardization (BN) and a sigmoid capacity to gain proficiency with the weight guide of nearby highlights, and the weight map is increased by the information. And afterward, this outcome is added to the contribution as the last result of current layer.

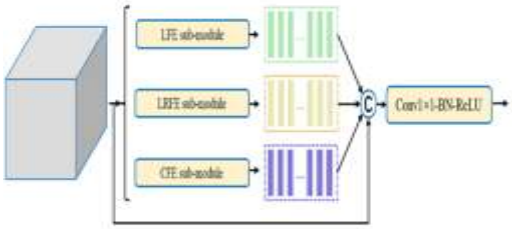

Fig 2 The MEC (Multi-feature Extraction and Combination) module

Utilizing pooling or walks can likewise get a bigger responsive field. Since the pooling activity will straightforwardly lose nearby data, we pick convolution with steps. Assume we have two continuous convolution layers, the principal convolution layer with channel piece K1 and step size S1, and the second convolutional layer with channel piece K2 and step size S2, the responsive field size is:

$$RFS_3 = K_1 + (K_2 - 1) \times S_1$$

2.3. Multi-Scale Prediction Fusion Module

Multi-scale expectation is demonstrated powerful in semantic division. In request to all the more likely foresee the fine shapes of water-bodies, we embrace a straightforward multi-scaleexpectation combination module. The MPF module advances the forecast aftereffects of three scales in disentangling stage. We first up-example the third-last and second-last encoder layers to the first picture measure and link them with the last forecast result. Then, at that point, we play out a 1 × 1 convolution with BN and ReLU to expand the quantity of channels, and individually utilize 3 × 3, 5×5, and 7 × 7 convolution with BN and a sigmoid capacity to learn multi-scale weight data. The loads contain significant signs from various open fields of the connected outcomes, which are duplicated separately by the loads. At last, we connect these outcomes and utilize a 3 × 3 convolution part Far off Sens. 2021
with BN and 1 × 1 convolution to acquire the last expectation result.

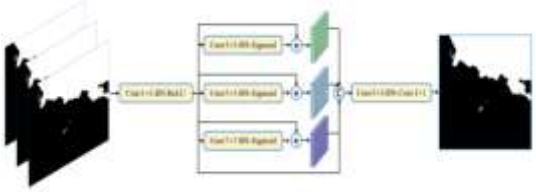

Fig 3. MPF: Multi-scale Prediction Fusion module

2.4. Encoder-Decoder Semantic Features Fusion Module

To take care of the issue of semantic irregularity in include combination at the interpreting stage, we apply the DSFF module), which expands the 3D divert consideration module proposed in our past work. The DSFF is intended for 2D tensors, first and foremost performs 1 × 1 convolution with BN and ReLU to diminish the channel number of the linked component maps at a similar scale from the encoding stage and the disentangling stage to half. Then, at that point, the worldwide setting is produced from the connected highlights by the worldwide pooling and is trailed by 1 × 1 convolution with BN and ReLU, and 1 × 1 convolution with a Sigmoid capacity. It is utilized as an aide for the combination of various semantic highlights, which consequently learns semantic associations between the channels of component maps. The worldwide setting data is increased with and added to the connected elements. At last, 3 × 3 convolutions with BN and a ReLU are applied to the got highlight maps. The DSFF module is conveyed on various scale highlights in the translating stage to accomplish proficient combination of various semantic highlights.

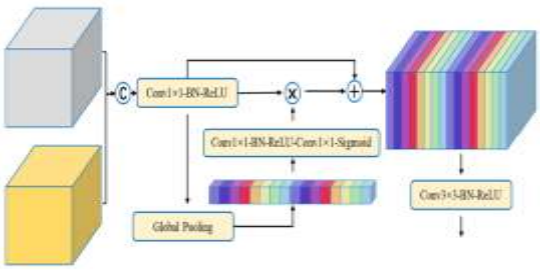

Fig 4 Different Semantic Feature Fusion module, DSFF

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 893-898 www.ijaem.net ISSN: 2395-5252

2.5. The total loss function

The preparation of profound neural organizations become more troublesome as the profundity of the organization increments. To prepare our proposed model all the more proficiently, we present a basic and successful result layer at each scale in the interpreting organize and force misfortune limitations between its outcome and ground truth. The result layer comprises of a 1 × 1 convolution layer and an up-examining layer, of which the quantity of result highlight map of the convolution layer is set to 1, and we utilize bilinear up-examining. The cross-entropy work L is utilized, and the complete misfortune work is as per the following:

$$L_{total} = \alpha L_{final} + \beta \sum_{i=1}^{5} L_i$$

2.6. Implementation Details

We carried out our strategy utilizing the Pytorch profound learning structure. Considering the restricted stockpiling of the GPU, we trimmed the first pictures into patches that measure 512 × 512 pixels with a cross-over proportion of 0.5 to kill the limit impacts. For a reasonable correlation among our and different strategies, we utilized the He introductionto introduce our model and different techniques in our work and train them without utilizing any pre-prepared loads. With the two bigger datasets presented in Section 3.1, we can extensively test model learning and speculation capacities. We applied arbitrary left-right also top-base flipping, Gaussian haze, and HSV change to contention the information. We set bunch size to 4 and embraced the Adam (versatile second assessment) streamlining agent also set the learning rate to 1e-4 and the quantity of ages to 32 in all trials.

III. RESULTS AND ANALYSIS

3.1. Water-Body Dataset

To assess our proposed strategy, we completed extensive examinations in aeronautical and GF2 (Gaofen2) faculties satellite symbolism. The ethereal pictures were caught in the Changxing space of Zhejiang Province, China, in 2018. What's more the dataset has an aggregate of 83 pictures, from which 63 and 20 are utilized for preparing and testing (Figure 1a). The size of every flying picture is 4994 × 4994 pixels, with the ground goal 0.2 m and three groups (red, green, and blue). From the aeronautical symbolism, it very well may be seen that there are weeds also residue on the two sides of some water-bodies, which makes the depiction of water-bodies seriously testing. Additionally, shadows are casted on some water-bodies at the vicinity of somewhat key position objects.

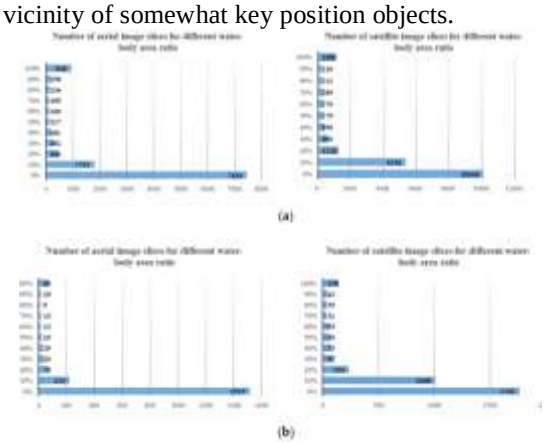

Fig 6. Distributions of the number of image slices over water-body area ratios, for (a) the two training sets and (b) test set

3.2. Evaluation Metrics

To look at results quantitatively, we utilized three assessment lists: Precision, Recall, what's more crossing point on association (IoU). They are characterized as:

$$IoU = \frac{TP}{TP + FP + FN}$$

3.3. Water-Body Segmentation Results

3.3.1. The Aerial Imagery

Our strategy accomplishes best in class precision. Our MECNet beated the second-best technique DANet 2.74% in IoU and 4.32% in review, which demonstrated that our technique enjoyed adequate benefits in distinguishing water-bodies and could assess more perplexing water tests. Be that as it may, our model acquired 1.02% lower in accuracy contrasted with the DANet. The explanation may be that our strategy perceived a few non-water pixels on the edge of water-bodies as water-body pixels. Among techniques with comparable constructions, our MECNet far outperformed U-Net and its variation RefineNet with 5.06% also 4.43% in IoU, which further exhibited the adequacy of the three modules we proposed.

Method	Backbone	Precision	Recall	IOU
U-Net	-	0.9076	0.9374	0.8558
RefineNet	resnet101	0.8741	0.9844	0.8623
DeeplabV3+	resnet101	0.9140	0.9417	0.8650
DANet	resnet101	**0.9259**	0.9456	0.8790
CascadePSP	DeeplabV3+&resnet50	0.9203	0.9409	0.8700
MECNet (ours)	-	0.9157	**0.9888**	**0.9064**

Table 1. The accuracy metrics of ours and other empirical networks using aerial imagery.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 893-898 www.ijaem.net ISSN: 2395-5252

3.3.2. The Satellite Imagery

A similar analysis is applied to the Gaofen2 satellite symbolism to additionally confirm the exhibition of our MECNet. Our proposed technique actually accomplishes the best exactness in IoU (Table 3). Our MECNet outperformed U-Net, RefineNet, DeeplabV3+, DANet and CascadePSP in accuracy, review and IoU, then again, actually DANet was somewhat better and U-Net was 1.26% higher as far as review. Figure 1 shows that the diagram of water-bodies in satellite symbolism is more clear and less difficult than that in airborne symbolism, which might be the motivation behind why U-Net, a lightweight and direct organization structure, could accomplish preferred outcomes over RefineNet and DANet. U-Net was 0.96% and 0.48% higher than the two techniques separately in IoU.

Method	Backbone	Precision	Recall	IOU
U-Net	-	0.9119	0.9756	0.8916
RefineNet	resnet101	0.9176	0.9578	0.8820
DeeplabV3+	resnet101	0.9379	0.9582	0.9010
DANet	resnet101	0.9156	0.9658	0.8868
CascadePSP	DeeplabV3+&resnet50	0.9378	0.9586	0.9013
MECNet (ours)	-	**0.9408**	**0.9630**	**0.9080**

Table: 2 Numerical comparisons with other empirical networks on the VHR satellite imagery. The bold format indicates the best results for each network in each evaluation metric

3.4. Ablation Studies

3.4.1. MECNet Components

In our proposed MECNet, the MEC module is intended to improve the element portrayal capacity at each scale. The MPF module is used at the last phase of decoder to completely incorporate the aftereffects of multi-scale forecast for the fine extraction of water-body shape. What's more the DSFF is embraced to settle the semantic irregularity of element combination between the encoder and the decoder. To confirm the presentation of our proposed modules, we directed broad trials with various settings. We dissected our strategies from quantitative and subjective viewpoints

Method	Parameter (M)	Flops (B)	IoU
FCN-8s	15.31	81.00	0.8399
FCN + MEC	26.11	105.59	0.8930
MEC + MPF	35.46	254.29	0.8974
MEC + MPF + DSFF (MECNet)	30.07	185.58	0.9064

.

Table 3. Our MECNet improve the performance of water-body segmentation on the VHR aerial imagery dataset. Parameters and FLOPs mean the parameters and floating-point operations per method. 'M': million, 'B': Billion.

3.4.2. LRFE Sub-Module

To plan and pick an appropriate LRFE sub-module, we led broad examinations with FCN as the pattern. The distinctions in execution of our planned LRFE sub-modules in VHR airborne symbolism. The three techniques, i.e., FCN+DCAClarge1, FCN+DCAC-large2 and FCN+DCAC-little (the various designs of LRFE have been recorded in Table 1), get comparable exactness, however FCN+DCAC-large2 has twice FLOPs contrasted with others since this model has the biggest number of boundaries. FCN with DCAC-little has a more modest open field than the others, yet its exactness is practically the same as that of the other two DCAC models, which uncovers that remote highlights has less consequences for the extraction of water data. FCN with JCC is near the precision of the DCAC model in IoU. Albeit the boundaries of its model were somewhat huge, it has just 55.29 GFLOPs. Considering the restricted equipment assets, we use JCC in the LRFE module.

Method	Parameters (M)	FLOPs(G)	IoU
FCN	15.31	81.10	0.8399
FCN+DCAC-large1	12.88	140.96	0.8801
FCN+DCAC-large2	13.39	248.30	0.8841
FCN+DCAC-small	12.70	106.57	0.8823
FCN+JCC	19.08	55.29	0.8816

Table 4. The performance of differently designed LRFE sub-modules in VHR aerial imagery.

3.4.3. MEC Module

In view of the perception in Section 2.2, we first independently dissected the exhibition of LFE, LRFE and CFE. Then, at that point, we considered the impacts of the in pairs blends of these three modules. At long last, we concentrated on the presentation of their blend and the effect of various mix ways. As displayed, we originally carried out FCN utilizing the LFE sub-module, and the IoU expanded from 83.99% to 84.78%. The IoU of FCN with the LRFE was 4.17% higher than the gauge. This shows that taking in highlights from bigger open field scenes is more worthwhile than gaining from nearby responsive fields, which might be a more significant part of spatial component extraction

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 893-898 www.ijaem.net ISSN: 2395-5252

Method	LFE	LRFE	CFE	IoU
FCN				0.8399
FCN	√			0.8478
FCN		√		0.8816
FCN			√	0.8835
FCN	√	√		0.8857
FCN	√		√	0.8851
FCN		√	√	0.8855
FCN (C)	√	√	√	0.8910
FCN (P)	√	√	√	0.8930

Table 6. Detailed performance of MEC module with different settings. '(C)' means the MEC module adopts a cascade way for the three sub-modules. '(P)' means the MEC module uses a parallel way for the three sub-modules.

IV. CONCLUSIONS

In this review, we develop dependent on the encoding–interpreting construction to work on fine water-body shape extraction from VHR remote detecting pictures, including aeronautical pictures also satellite pictures. Three modules are vital in our strategy: (1) a MEC module, for naturally removing more extravagant and more different elements in the encoding organize and acquire further developed semantic data for highlight combination in the interpreting stage; (2) a MPF module, which achieves the fine shape of the water-bodies; (3) a DSFF module, which takes care of the issue of semantic irregularity of element combination between the encoding stage also the interpreting stage. We completed tests on VHR airborne and satellite symbolism, individually, and the tests show that our technique accomplishes cutting edge exactness just as the best strength in testing situations. This original plan module for highlight extraction can be applied to other application situations, like semantic division also object identification.

REFERENCES

[1]. Mantzafleri, N.; Psilovikos, A.; Blanta, A. Water quality monitoring and modeling in Lake Kastoria, using GIS. Assessment and management of pollution sources. Water Resour. Manag. 2009, 23, 3221–3254. [CrossRef]

[2]. Frazier, P.S.; Page, K.J. Water body detection and delineation with Landsat TM data. Photogramm. Eng. Remote Sens. 2000, 66, 1461–1468.

[3]. Zhao, X.; Wang, P.; Chen, C.; Jiang, T.; Yu, Z.; Guo, B. Waterbody information extraction from remote-sensing images after disasters based on spectral information and characteristic knowledge. Int. J. Remote Sens. 2017, 38, 1404–1422. [CrossRef]

[4]. Yu, Z.; Feng, C.; Liu, M.-Y.; Ramalingam, S. Casenet: Deep category-aware semantic edge detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5964–5973.

[5]. Duan, L.; Hu, X. Multiscale Refinement Network for Water-Body Segmentation in High-Resolution Satellite Imagery. IEEE Geosci. Remote Sens. Lett. 2019, 17, 686–690. [CrossRef]

[6]. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.

[7]. Cheng, H.K.; Chung, J.; Tai, Y.-W.; Tang, C.-K. CascadePSP: Toward Class-Agnostic and Very High-Resolution Segmentation via Global and Local Refinement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Virtual Conference, Seattle, WA, USA, 14–19 June 2020; pp. 8890–8899.

[8]. Ji, S.; Zhang, Z.; Zhang, C.; Wei, S.; Lu, M.; Duan, Y. Learning discriminative spatiotemporal features for precise crop classification from multi-temporal satellite images. Int. J. Remote Sens. 2020, 41, 3162–3174. [CrossRef]

[9]. Gu, Y.; Lu, X.; Yang, L.; Zhang, B.; Yu, D.; Zhao, Y.; Gao, L.; Wu, L.; Zhou, T. Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs. Comput. Biol. Med. 2018, 103, 220–231. [CrossRef]

[10]. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L. Pytorch: An imperative style, high-performance deep learning library. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; pp. 8026–8037.

[11]. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. arXiv 2014, arXiv:1412.6980.