WEU-Net: A Weight Excitation U-Net for Lung Nodule Segmentation

Syeda Furruka BANU [a,1], Md. Mostafa Kamal SARKER [b],
Mohamed ABDEL-NASSER [a], Hatem A. RASHWAN [a], and Domenec PUIG [a]

[a] *Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili,
43007 Tarragona, Spain*
[b] *Precision Medicine Centre of Excellence, The Patrick G Johnston Centre for Cancer
Research, Queen's University Belfast, BT9 7BL, United Kingdom*

Abstract. Lung cancer is a dangerous non-communicable disease attacking both women and men and every year it causes thousands of deaths worldwide. Accurate lung nodule segmentation in computed tomography (CT) images can help detect lung cancer early. Since there are different locations and indistinguishable shapes of lung nodules in CT images, the accuracy of the existing automated lung nodule segmentation methods still needs further enhancements. In an attempt towards overcoming the above-mentioned challenges, this paper presents WEU-Net; an end-to-end encoder-decoder deep learning approach to accurately segment lung nodules in CT images. Specifically, we use a U-Net network as a baseline and propose a weight excitation (WE) mechanism to encourage the deep learning network to learn lung nodule-relevant contextual features during the training stage. WEU-Net was trained and validated on a publicly available CT images dataset called LIDC-IDRI. The experimental results demonstrated that WEU-Net achieved a Dice score of 82.83% and a Jaccard similarity coefficient of 70.55%.

Keywords. Lung Cancer, Computed Tomography (CT), Lung Nodule Segmentation, Computer-Aided Diagnosis (CAD), Deep Learning.

1. Introduction

Lung cancer causes around 1.3 million deaths each year over the world according to the World Health Organization (WHO). The death rate in European Union (EU) is about 44%, in turn in Japan is 35%. Since lung cancer detection is a challenging task in early stages; most cases are frequently diagnosed in the late-stage, and thus patients cannot survive because of improper prediction of disease and treatment at the late-stage of the disease at the hospital.The accurate detection and analysis of the lung nodules in the lung tissue in an early phase highly increase the patient chance of survival and facilitates effective treatment [1]. Computed tomography (CT) scans are a popularly utilized and profoundly accurate format for analyzing the lung nodules.The multi-detector row CT scanners are utilized to collect these lung scans.The precise segmentation of the lung nodules has a great impact on the diagnosis and prognosis of lung cancer [2]. Therefore,

---

[1]Corresponding Author: E-mail: syedafurruka.banu@estudiants.urv.cat

a radiologist must go through all individual patient CT scans that contain so many slices (about 150-500), which is a very difficult and tedious task [3]. However, it is difficult to distinguish the internal lung structure and the nodules because of the complex tissue environment, particularly when the nodule is located on the lung surface or attached to the border of a vessel in the lung tissue. To an accurate segmentation of the lung nodules by using normal image processing technique (e.g., threshold and morphological-based methods) is very challenging due to the large variety in size and types of lung nodules [4]. Another, the segmentation of nodules with small diameter and intensity variations with the surrounding noise is also very difficult for segmenting the nodule accurately. Thus, it is necessary to develop a robust segmentation system that can adapt to all the limitations and achieves efficient performance. Recently, convolutional neural networks (CNN) have become dominant in the area of computer vision (e.g., semantic segmentation, image classification, object detection, etc.). In the biomedical image analysis domain, an encoder-decoder like CNN architecture called U-Net [5], shown exceptional results on the task of segmentation. Several modified version of the U-Net yields state-of-the-art results in this domain. Nonetheless, the development of CNN architectures for the segmentation of lung nodules is yet immature. Hence, it is necessary to develop advanced architectures that can deal with the weaknesses of the previous architectures. In this paper, a weight excitation (WE) mechanism through the weight reparameterization-based backpropagation corrections is used from [6] and implemented with the U-Net encoder and decoder architecture to cope with the heterogeneity of lung nodule feature extraction efficiently, which is suitable for the segmentation of various forms of lung nodules.The main contributions of this research can be summarized as follows:

- We propose WEU-Net, which is a robust and efficient lung nodule segmentation model.
- A weight excitation (WE) mechanism through the weight reparameterization-based backpropagation corrections is adapted with the U-Net.
- The WEU-Net model achieved an overall precise segmentation performance on different challenging cases of nodule segmentation.

The preparation of this article is as follows. Section II discusses recent lung nodules segmentation methods based on traditional methods and deep learning techniques. The details of the proposed models architecture and the experimental setup and results are described in Section III and IV, respectively. Finally, Section V concludes and proposes some future works of this research.

2. Related Works

Several traditional (e.g., morphological, region-growing, and energy-based techniques, etc.), machine learning and deep learning-based lung nodule segmentation approaches that have been introduced in the last few years are described in this section. In traditional approaches, a morphological based-operation to remove the nodules associated with the vessels, and isolating the lung nodules by choosing the connected region is introduced in [7]. Another improvement of the separation of nodules from the lung surface is presented in [8] by using the combination of morphological based-method with the shape condition. However, the segmentation of lung nodules by morphological operations is

extremely challenging [9]. [4] presented that other well-known region-growing methods are not able to segment the nodules perfectly among the different types. In [10] noticed these challenges and proposed a region-growing technique based on the intensity information, distance, fuzzy connectivity, and peripheral variation of the nodules. The challenge of these tasks is the convergence condition of the methods. Furthermore, it is also hard to segment the irregular and fuzzy-shaped nodules using the region-growing techniques because of the nodule shape condition. Moreover, the energy based-optimization methods are also introduced to segment the lung nodules by [11], [12], [13], [14] using a level set function to minimize the energy function for achieving the segmented contour which reaches the boundary of the lung nodule. In [15]; [16]; [17] presented a maximum flow concept similar to the region-growing methods for segmenting the low contrast nodules.

Over the last decade, researchers are introduced machine-learning approaches for the classification, segmentation, and detection of lung nodules. In [18] presented extraction of rich feature maps with the invariance of translation and rotation. Other textures and shape-dependent features are utilized for the classification of voxels of lung nodules with conditional random field (CRF) model is presented in [19]. The Hessian matrix-based vascular feature extraction procedure is introduced in [20]. and used lung blood vessel segmentation and classification. Another Hessian-based 3D large-scale nodule segmentation method is utilized by [21]. Currently, deep learning approaches outperform the traditional and machine learning-based methods in most of the domains. Several CNN-based models were proposed for the task of the classification and segmentation of voxels in a supervised manner. For example, the multi-view convolutional neural network (MVCNN) is proposed by [22]. for nodule segmentation, which consists of three branches of CNN modules that are related to the three positions of the sagittal, axial, and coronal plane. Familiarly, fully convolutional networks (FCN), 2D U-Net, and 3D U-Net architectures are introduced by [23], [5], and [24], respectively are commonly used deep learning models for the segmentation of the biomedical imaging. Similarly, the central focused convolutional neural network(CF-CNN) is based on the shape-aware method proposed by [25], for the lung nodules segmentation task. Recently, a dual-branch residual network (DBResNet) is proposed by [26] using the combination of intensity features into the CNN model to achieve better segmentation results on lung nodules.

3. Proposed Model

The proposed WEU-Net is a encoder-decoder model shown in Figure 1 that modified version of the popular U-Net [5] model for the biomedical image segmentation task. The encoder part of the WEU-Net is composed by weight excitation based CNN (WE-CNN) [6], ReLU activation and MaxPooling layers. The input size of the model is $256x256x1$, because of the single channel CT lung nodule image. Initially, the WE-CNN and ReLU layers are apply to capture the contextual features from the input image. The encoder is a collection of WE-CNN and max pooling layers. The size of the feature maps are gradually decreases while the depth successively increases in the encoder from $256x256x64$ to $16x16x1024$. The decoder is the stack of transposed WE-CNN layers to up-sample the feature maps to the same size of the model input. The size of the decoder features are gradually increases and the depth gradually decreases from $32x32x512$ to $256x256x2$.
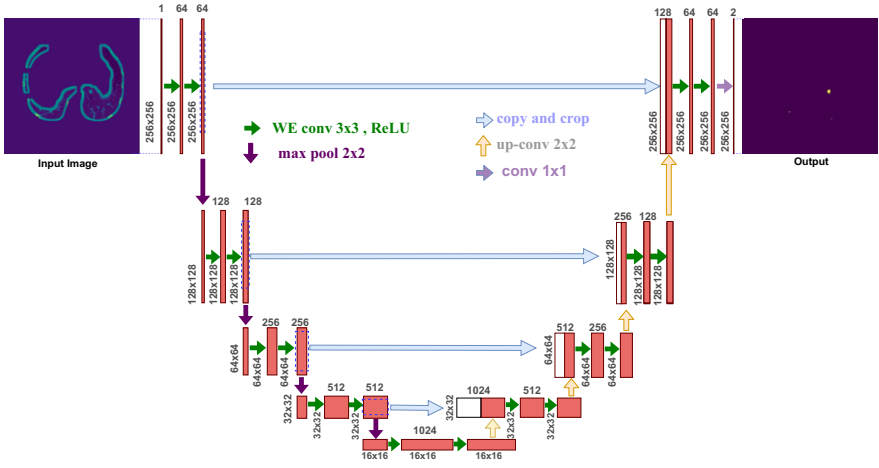
Figure 1. The architecture of the proposed WEU-Net.

After every decoder block, the feature maps are up-sampled and get the same size as the corresponding encoder block output to maintain harmony and concatenated it. This mechanism helps to keep the features that are learned from the encoder phase and use them for the reconstruction method. A 1×1 WE-CNN is used at the final layer of the network to map the final 64 feature vector to the targeted number of the segmentation classes which is two (background and lung nodule). A total of 23 WE-CNN layers are used in the U-Net model.

4. Experimental Setup and Results

4.0.1. Dataset

We used a publicly available dataset from the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) [27] to train and evaluate the proposed WEU-Net. The dataset consists of 888 paitents CT scans with annotations that were annotated by four experienced radiologists by using two-phase annotation process. The dataset contains annotations of 1186 nodules with the annotation files indicating different features of the nodules. The dataset remains with a total of 1166 CT images with corresponding ground truth masks after the pre-processing method is used for cleaning the dataset. We then split the dataset into two subsets training and test with 922 and 244 CT images, respectively.

4.0.2. Evaluation Metrics

We evaluate the proposed WEU-Net by using five metrics, accuracy (ACC), intersection over union (IoU), Dice similarity coefficient (DIC), precision (PRE), and recall (REC). Let the annotated ground-truth is *y*, the true positive (TP) and the false positive (FP) rates are correctly and incorrectly classified pixels as *y*, whereas the true negative (TN) and the false negative (FN) rates are correctly and incorrectly classified pixels as not *y*. The mathematical definitions of the five metrics: ACC, IoU DIC, PRE, and REC are presented following.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$IoU = \frac{TP}{TP + FP + FN} \tag{2}$$

$$DIC = \frac{2.TP}{2.TP + FP + FN} \tag{3}$$

$$PRE = \frac{TP}{TP + FP} \tag{4}$$

$$REC = \frac{TP}{TP + FN} \tag{5}$$

4.0.3. Data Augmentation and Implementation

We used online data augmentation methods during the training phase of the proposed model to increase the amount of training data by flipping the images horizontally and vertically, random cropping, and rotating. We implemented the proposed model on the PyTorch framework [28] and used NVIDIA 1080Ti GPU with 8GB memory. The Adam optimizer [29] is used with the learning rate and the batch size are set to 0.0002 and 8, respectively. The binary cross-entropy with dice loss is used as a loss function to train the proposed model.

4.0.4. Results

We carried a comprehensive investigation to evaluate how our proposed WEU-Net model improves the segmentation performance compared to the four state-of-the-art semantic segmentation models, PSPNet [30], Linknet [31], FPN [32], and vanila U-Net [5]. A performance comparison between the proposed model and four segmentation methods on the LIDC-IDRI dataset presents in Table 1.

Table 1. A comparison between the proposed model and four segmentation methods on the LIDC-IDRI dataset

Model Name	Metrices				
	ACC	IoU	DIC	PRE	REC
PSPNet [30]	99.87	62.12	77.34	73.00	83.78
Linknet [31]	99.89	63.30	77.43	75.00	83.58
FPN [32]	99.91	64.77	78.56	78.62	80.81
U-Net [5]	99.95	67.70	80.49	81.21	82.28
WEU-Net	**99.99**	**70.55**	**82.83**	**81.66**	**86.53**

The proposed model achieved the highest performance of 99.99%, 70.55%, 82.83%, 81.66%, and 86.53% in terms of ACC, IoU, DIC, PRE, and REC, respectively. Comparing to vanila U-Net, it improves 0.04%, 2.85%, 2.34%, 0.45%, and 4.25% of ACC, IoU, DIC, PRE, and REC, respectively. Therefore, it clearly shows that the effect of the weight excitation-based CNN is significantly improving the performance of U-Net. On

the other hand, it also improves the IoU of 8.43%, 7.25%, 5.78%, and 2.85% and the DIC of 5.49%, 5.40%, 4.27%, and 2.34% compared to the PSPNet, Linknet, FPN, and vanila U-Net, respectively.
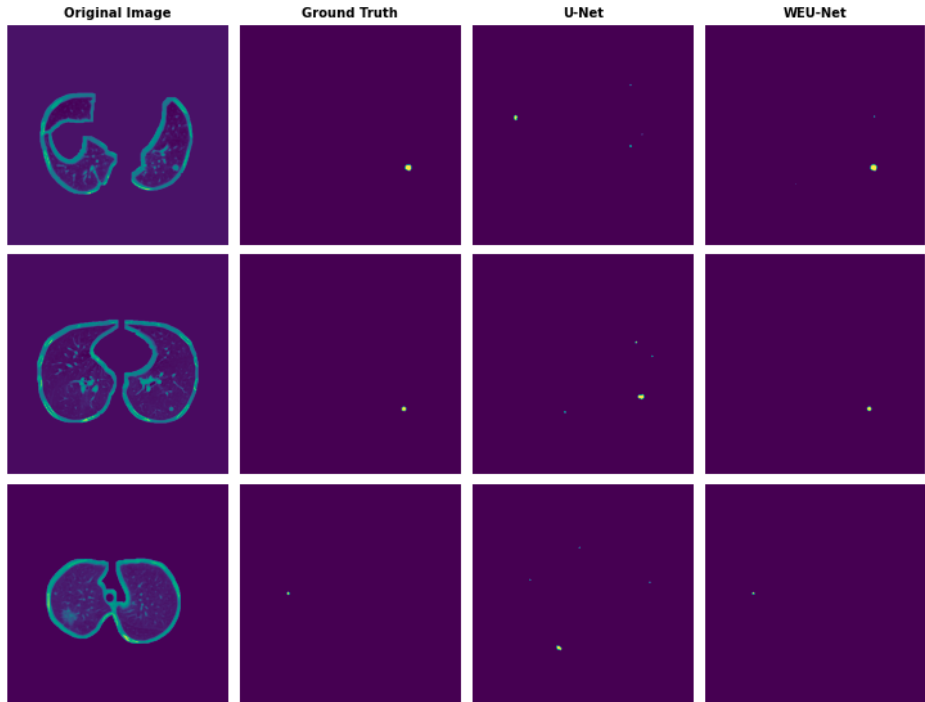

Figure 2. Examples of segmentation results (from left to right column represents; original image, ground truth, segmented by U-Net and proposed WEU-Net).

Figure 2 illustrates the qualitative segmentation results of WEU-Net from some examples of the LIDC-IDRI dataset. As can be seen in the predictions of U-Net model contains wrong segmentation and consists of false positive, whereas the proposed model segment the nodule region accurately. The proposed model can also segment the tiny nodule region preciously when the vanila U-Net failed to segment it. Therefore, the visualization demonstrates that the proposed model improved the U-Net model performance significantly for the lung nodule segmentation task.

5. Conclusion

This paper proposes a weight excitation U-Net for lung nodule segmentation. The weight excitation mechanism extracts the features in a weighted manner which leads to improving the performance of the vanila U-Net. The experimental evaluation shows that the proposed model demonstrated promising improvement compared with the four state-of-the-art semantic segmentation models including vanila U-Net. The proposed model achieved 99.99%, 70.55%, 82.83%, 81.66%, and 86.53% of ACC, IoU, DIC, PRE, and REC, respectively on the LIDC-IDRI dataset. The future work focuses on developing a

3D WEU-Net model based on the 3D weight excitation-based CNN (3D WE-CNN) for fully automated segmentation and classification of lung cancer.

Acknowledgement

References

[1]   A. El-Baz and J.S. Suri, *Lung imaging and computer aided diagnosis*, CRC Press, 2011.

[2]   H. MacMahon, J.H. Austin, G. Gamsu, C.J. Herold, J.R. Jett, D.P. Naidich, E.F. Patz Jr and S.J. Swensen, Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner Society, *Radiology* **237**(2) (2005), 395–400.

[3]   T. Way, H.-P. Chan, L. Hadjiiski, B. Sahiner, A. Chughtai, T.K. Song, C. Poopat, J. Stojanovska, L. Frank, A. Attili et al., Computer-Aided Diagnosis of Lung Nodules on CT Scans:: ROC Study of Its Effect on Radiologists' Performance, *Academic radiology* **17**(3) (2010), 323–332.

[4]   T. Kubota, A.K. Jerebko, M. Dewan, M. Salganicoff and A. Krishnan, Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models, *Medical Image Analysis* **15**(1) (2011), 133–154.

[5]   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*, Springer, 2015, pp. 234–241.

[6]   N. Quader, M.M.I. Bhuiyan, J. Lu, P. Dai and W. Li, Weight Excitation: Built-in Attention Mechanisms in Convolutional Neural Networks, in: *European Conference on Computer Vision*, Springer, 2020, pp. 87–103.

[7]   W.J. Kostis, A.P. Reeves, D.F. Yankelevitz and C.I. Henschke, Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images, *IEEE transactions on medical imaging* **22**(10) (2003), 1259–1274.

[8]   D. Sargent and S.Y. Park, Semi-automatic 3D lung nodule segmentation in CT using dynamic programming, in: *Medical Imaging 2017: Image Processing*, Vol. 10133, International Society for Optics and Photonics, 2017, p. 101332.

[9]   S. Diciotti, S. Lombardo, M. Falchini, G. Picozzi and M. Mascalchi, Automated segmentation refinement of small lung nodules in CT scans by local shape analysis, *IEEE Transactions on Biomedical Engineering* **58**(12) (2011), 3418–3428.

[10] J. Dehmeshki, H. Amin, M. Valdivieso and X. Ye, Segmentation of pulmonary nodules in thoracic CT scans: a region growing approach, *IEEE transactions on medical imaging* **27**(4) (2008), 467–480.

[11] T.F. Chan and L.A. Vese, Active contours without edges, *IEEE Transactions on image processing* **10**(2) (2001), 266–277.

[12] P.P. Rebouças Filho, A.C. da Silva Barros, J.S. Almeida, J. Rodrigues and V.H.C. de Albuquerque, A new effective and powerful medical image segmentation algorithm based on optimum path snakes, *Applied Soft Computing* **76** (2019), 649–670.

[13] E.E. Nithila and S. Kumar, Segmentation of lung nodule in CT data using active contour model and Fuzzy C-mean clustering, *Alexandria Engineering Journal* **55**(3) (2016), 2583–2588.

[14] A.A. Farag, H.E. Abd El Munim, J.H. Graham and A.A. Farag, A novel approach for lung nodules segmentation in chest CT using level sets, *IEEE Transactions on Image Processing* **22**(12) (2013), 5202–5213.

[15] X. Ye, G. Beddoe and G. Slabaugh, Automatic graph cut segmentation of lesions in CT using mean shift superpixels, *International journal of biomedical imaging* **2010** (2010).

[16] S. Mukherjee, X. Huang and R.R. Bhagalia, Lung nodule segmentation using deep learned prior based graph cut, in: *2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017)*, IEEE, 2017, pp. 1205–1208.

[17] Y. Boykov and V. Kolmogorov, An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, *IEEE transactions on pattern analysis and machine intelligence* **26**(9) (2004), 1124–1137.

[18] L. Lu, P. Devarakota, S. Vikal, D. Wu, Y. Zheng and M. Wolf, Computer aided diagnosis using multilevel image features on large-scale evaluation, in: *International MICCAI Workshop on Medical Computer Vision*, Springer, 2013, pp. 161–174.

[19] D. Wu, L. Lu, J. Bi, Y. Shinagawa, K. Boyer, A. Krishnan and M. Salganicoff, Stratified learning of local anatomical context for lung nodules in CT images, in: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, 2010, pp. 2791–2798.

[20] Y. Hu and P.G. Menon, A neural network approach to lung nodule segmentation, in: *Medical Imaging 2016: Image Processing*, Vol. 9784, International Society for Optics and Photonics, 2016, p. 97842.

[21] L. Gonçalves, J. Novo and A. Campilho, Hessian based approaches for 3D lung nodule segmentation, *Expert Systems with Applications* **61** (2016), 1–15.

[22] J. Wang and H. Guo, Automatic approach for lung segmentation with juxta-pleural nodules from thoracic CT based on contour tracing and correction, *Computational and mathematical methods in medicine* **2016** (2016).

[23] J. Long, E. Shelhamer and T. Darrell, Fully convolutional networks for semantic segmentation, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[24] Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, T. Brox and O. Ronneberger, 3D U-Net: learning dense volumetric segmentation from sparse annotation, in: *International conference on medical image computing and computer-assisted intervention*, Springer, 2016, pp. 424–432.

[25] S. Wang, M. Zhou, O. Gevaert, Z. Tang, D. Dong, Z. Liu and T. Jie, A multi-view deep convolutional neural networks for lung nodule segmentation, in: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2017, pp. 1752–1755.

[26] H. Cao, H. Liu, E. Song, C.-C. Hung, G. Ma, X. Xu, R. Jin and J. Lu, Dual-branch residual network for lung nodule segmentation, *Applied Soft Computing* **86** (2020), 105934.

[27] S.G. Armato III, G. McLennan, L. Bidaut, M.F. McNitt-Gray, C.R. Meyer, A.P. Reeves, B. Zhao, D.R. Aberle, C.I. Henschke, E.A. Hoffman et al., The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans, *Medical physics* **38**(2) (2011), 915–931.

[28] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai and S. Chintala, PyTorch: An Imperative Style, High-Performance Deep Learning Library, in: *Advances in Neural Information Processing Systems 32*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox and R. Garnett, eds, Curran Associates, Inc., 2019, pp. 8024–8035. `http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf`.

[29] D.P. Kingma and J. Ba, Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980* (2014).

[30] H. Zhao, J. Shi, X. Qi, X. Wang and J. Jia, Pyramid scene parsing network, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2881–2890.

[31] A. Chaurasia and E. Culurciello, Linknet: Exploiting encoder representations for efficient semantic segmentation, in: *2017 IEEE Visual Communications and Image Processing (VCIP)*, IEEE, 2017, pp. 1–4.

[32] A. Kirillov, K. He, R. Girshick and P. Dollár, A unified architecture for instance and semantic segmentation, 2017.