2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Induced Pluripotent Stem Cells Detection via Ensemble Yolo Network

Xinglie Wang, Jinqi Liao, Guanghui Yue, Liangge He, Tianfu Wang, Guangqian Zhou, Beiying Lei

Abstract— **Induced pluripotent stem cells (iPSCs) have huge potential in regenerative medicine research and industrial applications. However, building automatic method without using cell staining technique for iPSCs identification is an important challenge. To improve the efficiency of producing iPSCs, we build an accurate and noninvasive iPSCs colonies detection method via ensemble Yolo network based on the self-collected bright-field microscopy images. Meanwhile, test-time augmentation (TTA) is leveraged to further improve the detection result of our iPSCs colonies detection method. Extensive experimental results on our dataset demonstrate that our method obtains quite favorable detection performance with the highest F1 score of 0.867 and the highest mean average precision score of 0.898, which outperforms most mainstream methods.**

I. Introduction

As a type of pluripotent stem cell derived from somatic cells through co-expression of defined pluripotency-associated factors, induced pluripotent stem cells (iPSCs) hold great promise in the field of regenerative medicine [1, 2]. However, iPSCs derivation is a slow and inefficient process, which takes 1-2 weeks for mouse cells and 3-4 weeks for human cells [1, 3]. Researchers have to spend a lot of time and energy on iPSCs microscopy images analysis as shown in Fig. 1, which indicates the urgent need to develop an automatic and fast method with reliable performance to detect iPSCs in microscopy images for further research.

Since Tokunaga *et al*. [4] first introduced machine learning into the quality evaluation of iPSCs in 2014, the automated analysis of iPSCs microscopy images via machine learning has been widely developed [5], but the images and features of iPSCs still need to be selected manually for training these researches, which requires numerous time and efforts from researchers. More recently, deep learning has been widely used in the field of iPSCs microscopy images analysis. For example, Kimmel *et al*. [6] developed a vector-based convolutional neural network (CNN) for distinguishing characteristics of iPSC colonies. Waisman *et al*. [7] used CNN for accurate cellular morphology recognition based on cell staining technique.

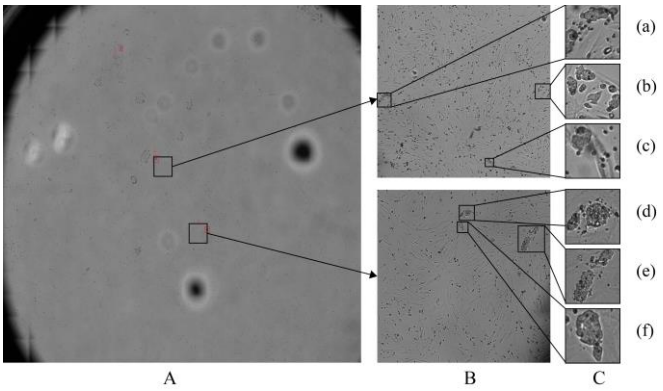

Fig. 1. Flow chart of iPSCs clone analysis under microscope manually. A is the view of iPSCs culture medium under microscope. B is the magnified view of specific area in A. (a) to (f) in C is the iPSCs colonies. Researchers have to find specific areas (B) containing iPSCs colonies under the whole view of A to make further analysis on iPSCs colonies (C).

However, searching for iPSCs colonies under extremely wide microscopic view to prepare datasets for the above methods still brings overwhelming workload for researchers. Also, the cell staining technique leads to massive death of iPSCs. To solve these problems, we propose a noninvasive iPSCs colonies detection method via ensemble Yolo network from a self-collected bright-field microscopy dataset without using cell staining technique.

Yolo network is famous for its excellent trade-off between detection accuracy and speed, especially the latest version Yolo v5 [8]. Enhanced by cross stage pyramid (CSP) structure, the backbone network Darknet53 named CSPDarknet53 used in Yolo v5 can not only ensure the network's strong feature learning ability, but also accelerate the detection speed of iPSCs colonies [8-10]. Feature pyramid network (FPN) [11], path aggregation network (PAN) [12] and spatial pyramid pooling (SPP) [13] module used in Yolo v5 can refine multi-scale feature learning from CSPDarknet53 and enhance the receptive field of Yolo v5, which assists the detection of iPSCs colonies in various scales.

This work was supported partly by National Natural Science Foundation of China (Nos.61871274, U1909209, and No. 62001302), Key Laboratory of Medical Image Processing of Guangdong Province (No. K217300003), Guangdong Pearl River Talents Plan (2016ZT06S220), Guangdong Basic and Applied Basic Research Foundation (Nos. 2021A1515011348, 2019A1515111205), Shenzhen Peacock Plan (Nos. KQTD2016 053112051497 and KQTD2015033016104926), and Shenzhen Key Basic Research Project (JCYJ20170818094109846), Natural Science Foundation of Shenzhen (Nos. JCYJ20190808145011259, RCBS20200714114920379).

T. Wang, B. Lei, Yue. G, X. Wang are with National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, China

G. Zhou, L. He, J. Liao are with Department of Medical Cell Biology and Genetics, Guangdong Key Laboratory of Genomic Stability and Disease Prevention, Shenzhen Key Laboratory of Anti-Aging and Regenerative Medicine, and Shenzhen Engineering Laboratory of Regenerative Technologies for Orthopaedic Diseases, Health Science Center, Shenzhen University, Shenzhen, 518060, China

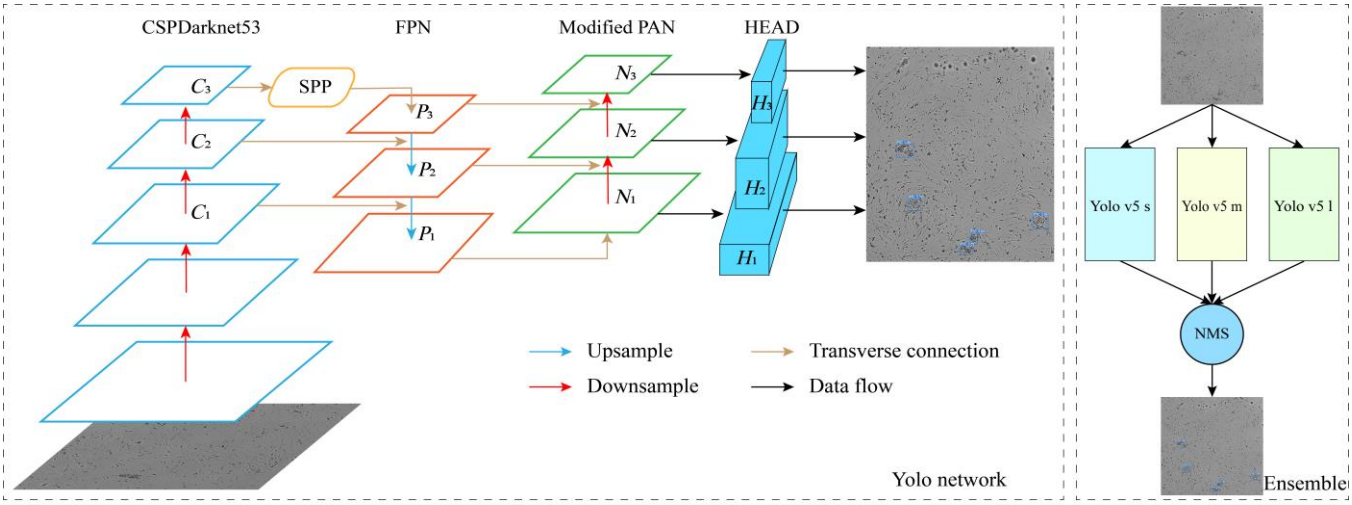

Fig. 2. (a) The detailed architecture of the basic Yolo network. C_1 to C_3 are feature maps at different levels generated by CSPDarknet53, P_1 to P_3 are coarse feature maps at different levels generated by FPN, N_1 to N_3 are feature maps at different levels generated by PAN. H_1 to H_3 are the multi-scale output of Yolo v5 head. (b) The proposed ensemble Yolo network for iPSCs colonies detection. Yolo v5 s, Yolo v5 m, Yolo v5 l indicate small, medium and large version of Yolo v5. NMS means Non-max Suppression algorithm.

To get a more accurate iPSCs colonies detection result, we utilize model ensemble technique to reduce the generalization error of the iPSCs colonies detection [14]. Meanwhile, test-time augmentation (TTA) techniques are used to further improve the performance of proposed method. Overall, the contribution of our works are as follows:

1) We first introduce a noninvasive way on iPSCs colonies detection with the self-collected bright-field microscopy images;
2) We combine ensemble learning with the state-of-the-art objection detection algorithm Yolo v5 to achieve better iPSCs colonies detection result;
3) TTA techniques are used in our proposed iPSCs detection method to further improve the performance.

II. METHODS

Fig. 2 shows the detail architecture of the proposed iPSCs colonies detection method. We adopt different versions of Yolo network as the basic iPSCs colonies detection network. Fig.2 (b) indicates the proposed model ensemble technique. Meanwhile, TTA techniques are used during detection time for more accurate detection results.

A. Basic Object Detection Algorithm

As shown in Fig. 2, we choose Yolov5 as the detection algorithm to locate iPSCs colonies in bright-field microscopy images. Yolov5 consists of 3 parts: backbone, neck and head. Compared to the original version of Darknet53 in Yolo network, CSPDarknet53 combines the CSP structure to reduce the repeated gradient information and improve learning ability of network. As illustrated in Fig. 3 (a), the basic block of CSPDarknet53 divides feature map of base layer C_i into two parts $C_{(i,1)}$ and $C_{(i,2)}$ along the channel dimension, $i = \{1, 2, 3\}$ represent the feature level of FPN or modified PAN. Part 2 $C_{(i,2)}$ will go through a Darknet53 block D. The output of D will undergo the first transition layer T_1 and concatenate with feature map $C_{(i,1)}$ coming from part 1 to undergo another transition layer T_2 .A hierarchical feature fusion mechanism consisting of 1×1 convolutional layer, batch normalization layer and Mish activation layer [15]. The continuous and differentiable activation function in Mish layer is used in CSPDarknet53 to increase the accuracy of detector. Mish activation function $f(x)$ mathematically is defined as:

$$f(x) = xtanh(\ln(1 + e^x)), \quad (1)$$

where x is the output vector of the normalization layer before activation layer, ln is natural logarithm, $tanh$ is hyperbolic tangent.

The neck of Yolo v5 mainly consists of three parts: top-down path FPN structure, modified bottom-up PAN structure by replacing shortcut connection to concatenate original PAN, additional SPP module. As shown in Fig.3 (c), FPN up-samples the high level feature P_{i+1} at a factor of 2 concatenates the coarse feature map C_{i+1} from CSPDarknet53 to generate new feature map P_i. Similar to FPN but in a different direction, the modified PAN block takes low level feature N_i and a coarse feature map P_{i+1} from lateral connection and generates the new feature map N_{i+1}. Fig. 3 (b) indicates the SPP module used in Yolo v5. The entire input feature C_i will go through 3 max-pooling layers M_1, M_2, M_3, respectively. Then the outputs of those layers are concatenated together with C_i to enhance the receptive field of the input feature map C_i and generate the output feature map C', which could enhance the model by fusing features at multi-scales.

As illustrated in Fig. 2 (a), the head of Yolo v5 consists of three parts H_1, H_2, H_3 at three different resolutions, which is the same as Yolo v3. The outputs for those three head will concatenate together and go through the NMS algorithm to generate final detection results.

B. Model Ensemble

Ensemble learning combines multiple learning algorithms for generating better predictive performance strategically. In order to take advantage of Yolo v5 network, we employ model ensemble technique for iPSCs colonies detection during inference time as illustrated in Fig.2 (b).

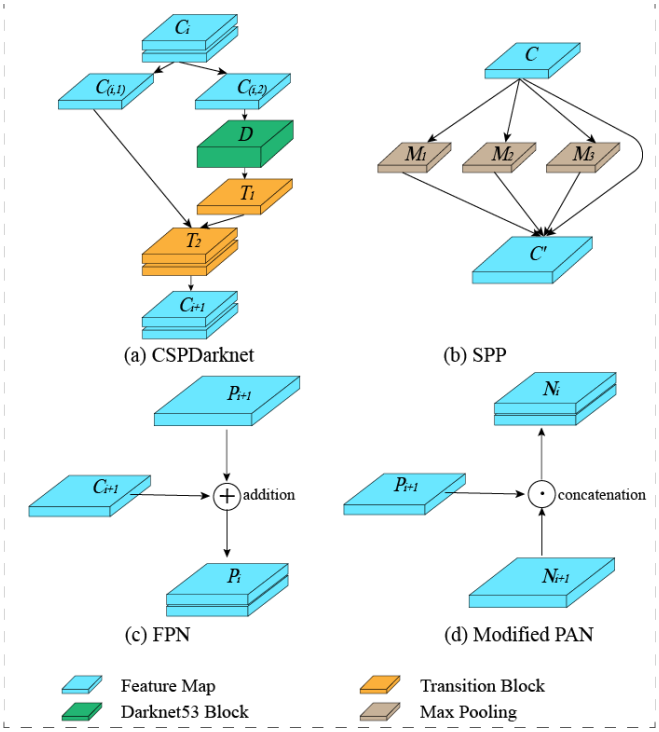

Fig. 3. Some basic block in Yolo v5 detection algorithm. (a) Basic block of CSPDarknet53; (b) SSP module; (c) FPN module; (d) Modified PAN block.

First, we train small, medium and large versions of Yolo v5 network separately. Then we concatenate the outputs including the bounding boxes coordinates and confidence scores of these 3 models, which can minimize the missed detections of iPSCs colonies and retain higher confidence scores. Finally, the ensemble output will go through the final Non-max Suppression algorithm to generate better iPSCs colonies detection results by discarding objects with high overlap ratios and low confidence scores.

C. Test-Time Augmentation

Similar to data augmentation for the training set, TTA is an application of random transformations to the test images. In order to generate more precise detection results, we perform TTA techniques for our iPSCs test set.

TTA techniques used in our experiments including left-right flipping and scaling of images. First, the input images are being left-right flipped and scaling to 3 different sizes with the ratios of 1, 0.83, 0.67, respectively. Those augmented images will go through the Yolo network and generate united predictions. Then we de-scale and de-flip the outputs and concatenate them as the inputs of the final NMS algorithm to generate the final detection result of iPSCs colonies.

III. Experiments

A. Experimental Setting-up

In order to evaluate the performance of our proposed iPSCs colonies detection method, we collect bright-field microscopy images for about 25 days since we carry out the reprogrammed procedure. During the image acquisition process, the microscope magnification is set to 4 times, and the imaging mode is set as bright-field. We collect images at the same 10 wells and select 527 available images with size 2048×2048 pixels per day. We choose image data on the 18^{th} day that could be labeled by our researchers for the following analysis, which is a much earlier time compared to the manual recognition time of researchers. Finally, the number of images we collected is 525, and we randomly divide them into training, validation and testing set at a ratio of 6:2:2. The summary information of collected iPSCs bright-field datasets of our experiments is given in Table 1.

Table 1. Summary of the iPSCs colonies bright-field microscopy images.

Data	Number	Object
Train	315	1034
Validation	105	310
Test	105	322

To solve imbalance problem in contrast and brightness of iPSCs bright-field microscopy images, we apply some data preprocessing methods to enhance the data as follows: 1) HSV color mode augmentation to adjust the contrast and brightness of images. 2) Random affine transformations including flipping, scaling and translation. Then, we resize all images to 640×640 with bilinear interpolation.

To evaluate the performance of all deep learning detection algorithm in a fair way, all hyper-parameters and settings are consistent in our experiments. The chosen optimizer during model training is "Adam" with the same learning rate 10^{-4}. The number of epoch for training is set to 200, while the batch-size is 8. All models in this study are trained on single NVIDIA GeForce GTX 2080 Ti 24G graphic cards.

B. Evaluation Metrics

Several evaluation metrics for object detection are adopted, including precision (Pre), recall (Rec), F1 score (F1), and mean average precision at intersection over union threshold 0.5 (mAP50). Pre, Rec and F1 score can illustrate the detection performance of foreground iPSCs colonies, while mAP50 has a more comprehensive ability for evaluating the foreground objects and background regions classification performance of the model.

C. Experiment Results

The iPSCs colonies detection results of ensemble Yolo v5 network compared to the basic Yolo v5 network are present in Table 2. Model ensemble technique might lead to the dropping of precision, but obtains higher recall, F1 score and mAP50.

Test-time Augmentation results are shown in Table 3. Compared to the experiment results in Table 2, TTA brings obvious improvement of precision, F1 score and mAP50, while the recall drops slightly.

We also compare our proposed method with state-of-the-art object detection methods including Faster R-CNN enhanced by ResNeXt101 and FPN, Cascade R-CNN enhanced by ResNeXt101 and FPN, Libra R-CNN enhanced by ResNet101 and FPN, RetinaNet enhanced by ResNeXt101 and FPN, EfficientDet-D1, EfficientDet-D2 [16-21]. The results in Table 4 show that our proposed ensemble Yolo network with TTA has better comprehensive performance.

IV. Conclusion

In this paper, we propose an automatic iPSCs colonies detection method via ensemble Yolo v5 network and test-time

augmentation techniques. The experimental results based on the self-collected datasets show that model ensemble achieves better comprehensive performance of Yolo network. Moreover, test-time augmentation further improves the distinguishing ability of our proposed method. In the future work, we expect to explore a better ensemble model for iPSCs colonies detection and improve our test-time augmentation algorithm to get better performance.

V. Compliance With Ethical Standards

We wish to confirm that there are no known conflicts of interest associated with this publication. This research study is approved by the ethical review board of the institute.

Table 2. Model ensemble results of Yolo network. + means ensemble detection. Alphabet s, m, l means small, medium and large version of Yolo v5, respectively.

Yolo v5	Precision	Recall	F1	mAP50
s	**0.889**	0.823	0.854	0.864
m	0.879	0.832	0.854	0.855
l	0.882	0.842	0.861	0.862
s + m	0.878	0.848	**0.863**	0.871
s + m + l	0.858	**0.863**	0.860	**0.873**

Table 3. Test-time Augmentation results of Yolo network. * means test-time augmentation.

Yolo v5	Precision	Recall	F1	mAP50
s *	0.892	0.822	0.856	0.877
m *	0.875	**0.847**	0.861	0.879
l *	0.891	0.838	0.864	0.868
s + m + l *	**0.894**	0.841	**0.867**	**0.898**

Table 4. Our proposed method compared to other state-of-the-art object detection methods.

Method	Precision	Recall	F1	mAP50
Faster R-CNN	0.864	0.851	0.858	0.836
Cascade R-CNN	0.845	0.845	0.845	0.829
RetinaNet	0.883	0.823	0.852	0.805
Libra R-CNN	0.840	**0.863**	0.851	0.831
EfficientDet-D1	0.852	0.820	0.835	0.891
EfficientDet-D2	0.869	0.848	0.860	0.819
ours	**0.894**	0.841	**0.867**	**0.898**

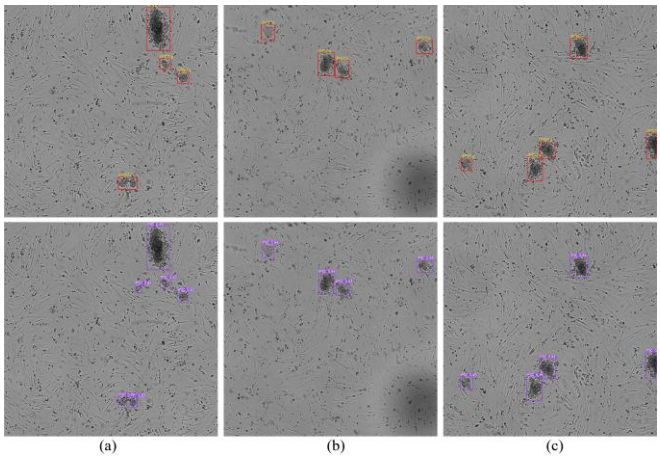

Fig. 4. Inference results visualization of the proposed ensemble Yolo network with test-time augmentation. The first raw is the images labeled by researchers. The second row is the corresponding detection results with confidence scores of our proposed method.

References

[1] K. Takahashi and S. Yamanaka, "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors," *cell,* vol. 126, no. 4, pp. 663-676, 2006.

[2] R. S. Mahla, "Stem cells applications in regenerative medicine and disease therapeutics," *International journal of cell biology,* vol. 2016, 2016.

[3] T. Zhou, C. Benda, S. Dunzinger, Y. Huang, J. C. Ho, J. Yang, Y. Wang, Y. Zhang, Q. Zhuang, and Y. Li, "Generation of human induced pluripotent stem cells from urine samples," *Nature protocols,* vol. 7, no. 12, p. 2080, 2012.

[4] K. Tokunaga, N. Saitoh, I. G. Goldberg, C. Sakamoto, Y. Yasuda, Y. Yoshida, S. Yamanaka, and M. Nakao, "Computational image analysis of colony and nuclear morphology to evaluate human induced pluripotent stem cells," *Scientific reports,* vol. 4, p. 6996, 2014.

[5] H. Joutsijoki, M. Haponen, J. Rasku, K. Aalto-Setälä, and M. Juhola, "Machine learning approach to automated quality identification of human induced pluripotent stem cell colony images," *Computational and mathematical methods in medicine,* vol. 2016, 2016.

[6] M. S. Kavitha, T. Kurita, S.-Y. Park, S.-I. Chien, J.-S. Bae, and B.-C. Ahn, "Deep vector-based convolutional neural network approach for automatic recognition of colonies of induced pluripotent stem cells," *PloS one,* vol. 12, no. 12, 2017.

[7] A. Waisman, A. La Greca, A. M. Möbbs, M. A. Scarafía, N. L. S. Velazque, G. Neiman, L. N. Moro, C. Luzzani, G. E. Sevlever, and A. S. Guberman, "Deep learning neural networks highly predict very early onset of pluripotent stem cell differentiation," *Stem cell reports,* vol. 12, no. 4, pp. 845-859, 2019.

[8] G. Jocher. (2020/04/29). *Yolo Network from Ultralytics LLC*. Available: https://github.com/ultralytics/yolov5

[9] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767,* 2018.

[10] C.-Y. Wang, H.-Y. Mark Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, "CSPNet: A new backbone that can enhance learning capability of cnn," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR)*, 2020, pp. 390-391.

[11] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2017, pp. 2117-2125.

[12] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2018, pp. 8759-8768.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE transactions on pattern analysis and machine intelligence,* vol. 37, no. 9, pp. 1904-1916, 2015.

[14] T. G. Dietterich, "Ensemble learning," *The handbook of brain theory and neural networks,* vol. 2, pp. 110-125, 2002.

[15] D. Misra, "Mish: A self regularized non-monotonic neural activation function," *arXiv preprint arXiv:1908.08681,* 2019.

[16] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91-99, 2015.

[17] M. Tan, R. Pang, and Q. V. Le, "Efficientdet: Scalable and efficient object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 10781-10790.

[18] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2017, pp. 1492-1500.

[19] Z. Cai and N. Vasconcelos, "Cascade R-CNN: high quality object detection and instance segmentation," *IEEE transactions on pattern analysis and machine intelligence,* 2019.

[20] J. Pang, K. Chen, J. Shi, H. Feng, W. Ouyang, and D. Lin, "Libra r-cnn: Towards balanced learning for object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 821-830.

[21] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition* 2017, pp. 2980-2988, 2017.