

## Multiclass Segmentation of Pulmonary Diseases using Convolutional Neural Network

Muhammad Arnaldo<sup>1</sup>, Siti Nurmaini<sup>2\*</sup>, Hadipurnawan Satria<sup>1</sup>, Muhammad Naufal Rachmatullah<sup>2</sup>

<sup>1</sup>Department of Informatics, Faculty of Computer Science, Universitas Sriwijaya, Indonesia

<sup>2</sup>Intelligent System Research Group, Universitas Sriwijaya, Indonesia

\*muhammad.arnaldo@gmail.com

### ABSTRACT

Pulmonary disease has affected tens of millions of people in the world. This disease has also become the cause of death of millions of its sufferers every year. In addition, lung disease has also become the cause of other respiratory complications, which also causes the death of the sufferer. The diagnosis of pulmonary diseases through medical imaging is a significant challenge in computer vision and medical image processing. The difficulty is due to the wide variety in infected areas' shape, dimension, and location. Another challenge is to differentiate one lung disease from the other. Discriminating pulmonary diseases is a notable concern in the diagnosis of pulmonary disease. We have adopted the deep learning convolutional neural network in this study to address these challenges. Seven models were constructed using the Mask Region-based Convolutional Neural Network (Mask-RCNN) architecture to detect and segment infected areas within the lung region from CT scan imagery. The evaluation results show that the best model obtained scores of 91.98%, 85.25%, and 93.75% for DSC, MIoU, and mAP, respectively. The segmentation results are then visualized.

**Keywords:** Mask-RCNN, Medical Images, Pulmonary Disease, Segmentation.

### 1. INTRODUCTION

Various kinds of pulmonary diseases have been the cause of death for more than 3 million people every year. In addition, the pulmonary disease also causes other complications related to the respiratory system, which also causes millions of deaths. [1]. Among the dangerous lung diseases are lung cancer [2], tuberculosis (TB) [3], and Covid-19 which became a pandemic beginning in 2019 [4].

Computed tomography (CT) scans and chest X-rays (CXR) are the two primary imageries to diagnose those pulmonary diseases, and the CT scans are the gold standard between the two [5][6]. The advantages of CT scans are wide availability, short scan time, and low cost. It also generates a very detailed view of the lungs [7]. Academia and industry have started using CT scans to test for potential lung disease in high-risk individuals. However, this becomes a heavy burden for the radiologist because of the large number of CT scans that must be analyzed. Thus, automation using computer algorithms is becoming an increasingly popular option [8].

One Artificial Intelligence (AI) approach that has become a promising technique in the detection and segmentation of CT imaging is Deep Learning (DL), specifically the Convolutional Neural Networks (CNNs) [9]. The utilization of DL on CT image data has been successfully carried out in several studies of various

pulmonary diseases [7][10][11]. However, the segmentation of lung disease lesions on CT images has been a challenge in computer-aided chest radiograph analysis. The challenge is due to the different shapes, locations, and sizes of the infection areas [8]. Moreover, if the type of disease to be detected varies. Therefore, a method that can produce high accuracy in segmentation is critical in building an automatic identification system. Thus, this study proposes a multiclass pulmonary diseases segmentation applied to CT images using the Mask R-CNN [12] method.

## **2. MATERIAL AND METHODS**

This study is addressed to segment various pulmonary diseases by employing a CNN method, utilizing the Mask-RCNN architecture [12]. Afterward, the performances will be evaluated to see if the models produce good results.

### **2.1 DATASET**

The datasets used in this study were obtained from various sources, from private dan public data. CT image data for lungs infected with COVID-19 was obtained from a hospital in Indonesia. The data is in video format, which is then processed into 725 images. Furthermore, data for healthy or normal lungs were obtained from MosMedData in NIfTI format, which used as many as 658 images. TB data was obtained from the TB Portals ([tbportals.niaid.nih.gov](http://tbportals.niaid.nih.gov)) in DICOM format, which is an open-access TB data resource supported by the National Institute of Allergy and Infectious Diseases (NIAID) Office of Cyber Infrastructure and Computational Biology (OCICB) in Bethesda, MD [13]. These data were collected and submitted by members of the TB Portals Consortium (<https://tbportals.niaid.nih.gov/Partners>). Investigators and other data contributors that originally submitted the data to the TB Portals did not participate in the design or analysis of this study. Whereas the lung cancer dataset was obtained from The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI) [14], which were also in DICOM format. 640 and 663 CT images were used for TB disease and lung cancer, respectively. Moreover, the expert has marked the areas that have the infection as ground truth.

The data that has become an image is then annotated based on the existing ground truth. Figure 1 shows the raw data (a), the ground-truth (b), and the annotated data (c).

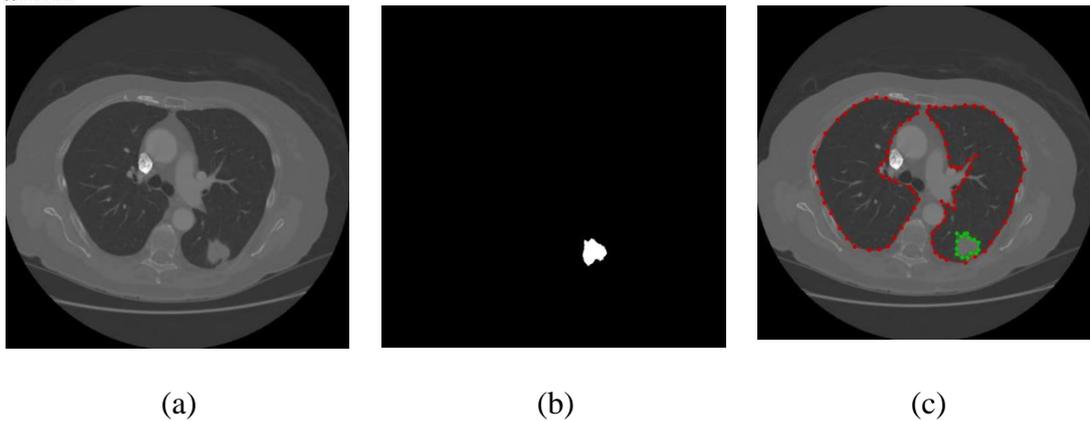


FIGURE 1. Raw data, ground-truth and annotated data

## 2.2 MASK-RCNN ARCHITECTURE

Further development of the object detection method Faster-RCNN [15] resulted in a new method called Mask-RCNN. It was developed by He et al. [12] to perform an instance segmentation task. Instance segmentation is a task capable of detecting all objects in an image correctly and segmenting each instance simultaneously. Initially, faster-RCNN was not made for pixel alignment. So then this lack of faster-RCNN is enhanced by He et al. [12] on mask-RCNN. Mask-RCNN has proven to be good at performing segmentation tasks in medical images of lung diseases [16] and other diseases [17]. Therefore, it was chosen as the method used in this study.

Similar to Faster-RCNN, Mask-RCNN also works in two stages [12][18]. The first stage will generate region proposals from the feature maps of an input image. These proposals are areas within the image that likely contain objects. Afterward, in the second stage, it will classify the region proposals detected previously and generate bounding boxes along with the masks [19]. Figure 2 illustrates the architecture of Mask-RCNN.

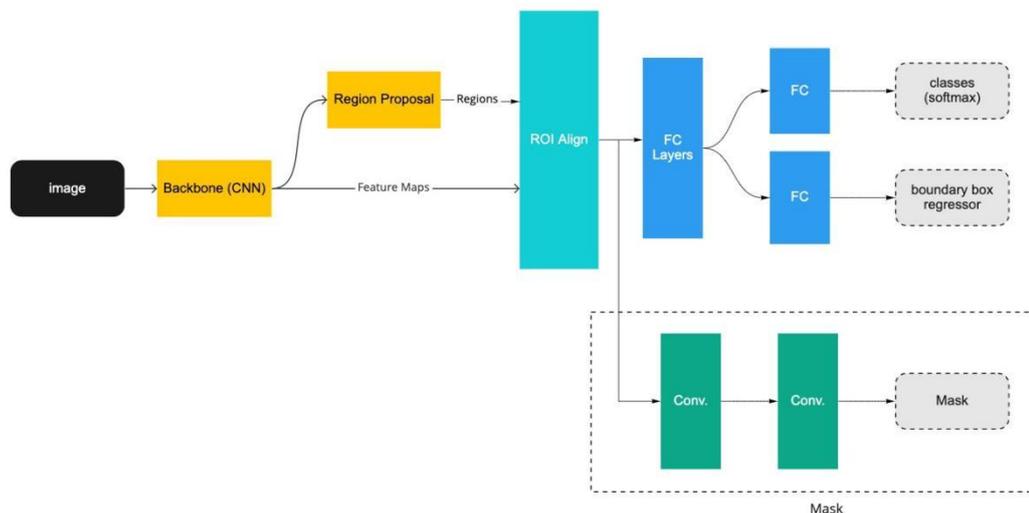


FIGURE 2. The architecture of Mask-RCNN

At a high level, Mask-RCNN consists of several components: Backbone, Region Proposal Network (RPN), ROI Classifier and Bounding Box Regressor, and Segmentation Masks [20]. The backbone is a Feature Pyramid Network (FPN) [21][22] style feature extractor. It will generate feature maps for the following process. The RPN is a lightweight neural network that performs the scanning process through a set of anchors [23]. The anchors are boxes distributed throughout the image. They differ in size and aspect ratios and overlap each other. These anchors work for binding features to the original image. The RPN then will select the anchors that most likely contain the object we want to detect, known as regions of interest (ROI). Furthermore, the ROI classifier and bounding box regressor will assign ROIs to specific classes and refine its bounding boxes, respectively, while the mask branch will generate the segmentation mask.

### 2.3 PERFORMANCE AND EVALUATION

Evaluation of the performance of each model is carried out with several measures, namely, dice score similarity (DSC), mean intersection over union (MIoU), and mean average precision (mAP). The DSC will be used to assess how closely predicted segmentation matched the annotated ground truth segmentation. The DSC is presented in (1) and defined as follows:

$$DSC(P, G) = 2 \frac{\sum_{i=0}^{n-1} \sum_{j=0}^{m-1} P_{ij} G_{ij}}{\sum_{i=0}^{n-1} \sum_{j=0}^{m-1} P_{ij} + \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} G_{ij}} \quad (1)$$

where P represents the predicted segmentation and G represents the ground truth segmentation. N is the height of the image with i as its indices. While M is the width of the image with j as its indices.

The MIoU is another popular performance evaluation tool for image segmentation. As the name suggests, it quantifies the intersection percentage between the ground-truth mask and the predicted mask [24]. The IoU for each class will be calculated, and the result will be averaged. The MIoU is calculated as shown in (2):

$$MIoU = \frac{1}{N_{cls}} \sum_{x=1}^{N_{cls}} \frac{N_{xx}}{\sum_{y=1}^{N_{cls}} N_{xy} + \sum_{y=1}^{N_{cls}} N_{yx} - N_{xx}} \quad (2)$$

The mAP metric will be used to quantify both segmentation and classification for all object classes. It can comprehensively measure precision, recall, and threshold. Therefore, this metric is convincing when evaluating the model. The mAP is calculated as in (3):

$$mAP = \sum_{i=1}^N \frac{AP_i}{N} \quad (3)$$

where  $i$  in  $AP_i$  is the class number, and  $n$  is the total number of all classes being assessed. The mAP is used as an overlap criterion defined as an intersection-over-union at a certain value. If a predicted box satisfies this criterion with respect to a ground-truth box, it is considered a detection.

### 3. RESULT AND DISCUSSION

In this study of multiclass segmentation of pulmonary diseases, there are in total four classes to be detected and segmented. They are lung region, lung cancer, TB, and Covid-19. Six models were constructed using different backbones and different hyperparameter configurations. The training processes also involved the transfer learning approach from a pre-trained model of the Microsoft Common Objects in Context (MS COCO) dataset. The backbones employed are Resnet50 and Resnet101 [25]. While the fine-tuned hyperparameter is the learning rate and the epoch. Details from each model and its configurations, along with the performances, are summarized in Table 1.

TABLE 1.  
Result of Mask-RCNN's model

Models	DSC	MIoU	mAP
Resnet50_LR 0.01_epoch 50	91.87	85.27	93.83
Resnet50_LR 0.01_epoch 100	91.41	84.96	94.12
Resnet50_LR 0.001_epoch 50	91.85	85.19	93.51
Resnet50_LR 0.001_epoch 100	91.98	85.25	93.75
Resnet50_LR 0.0001_epoch 50	89.58	82.81	92.33
Resnet101_LR 0.001_epoch 50	91.76	85.13	93.90
Resnet101_LR 0.001_epoch 100	91.59	84.96	93.75

Each model constructed has successfully segmented the images in the test set with scores over 80%. It has scored 91.98% for DSC, 85.25% for MIoU, and 93.75% for mAP. The model acquired the lowest score with the backbone Resnet50 utilizing a learning rate value of 0.0001. That learning rate was the smallest learning rate configuration of all models. Whereas the top score was also acquired by the model using backbone Resnet50, the learning rate was fine-tuned to 0.001, and the epoch increased to 100. Subsequently, the segmentation was visualized from the top-scoring model for a better review. The visualization of the segmentation result is

**Muhammad Arnaldo, Siti Nurmaini,  
Hadipurnawan Satria, Muhammad Naufal Rachmatullah**  
**Multiclass Segmentation of Pulmonary Diseases using Convolutional Neural Network**

presented in Figure 3. The colors for the segmentation mask in (a) and (b) are arbitrary.

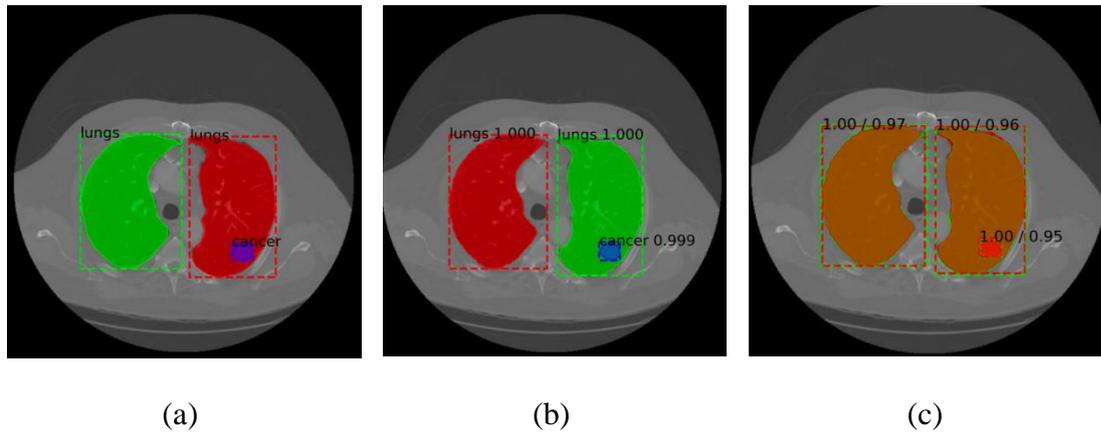


FIGURE 3. Visualization of the segmentation result

The visualization is displayed with three kinds of images. The first image shown in (a) is the ground truth and the second image shown in (b) is the prediction. Each instance has its bounding box, mask, and its instance name. The image in (b) shows that the model has successfully detected and segmented all the instances. Furthermore, to better compare (a) and (b), the third image is presented in (c). The image in (c) is the overlapping mask between the ground truth and the prediction mask. The confidence score and the IoU score are presented in (c) respectively to support the visual comparison. Two lung instances and a single instance of an infected area are shown in Figure 3. The following image shows more instances of infected areas.

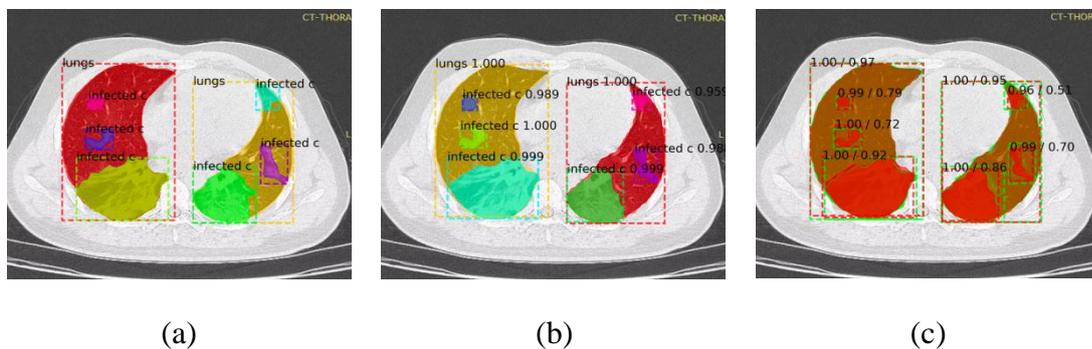


FIGURE 4. Segmentation result with multiple infected areas

There are in total six instances of infected areas in Figure 4. The image in (b) shows that the model can detect and segment multiple instances of infected areas and the lung area within the CT imagery.

#### 4. CONCLUSION

The demand for rapid, accurate, and automatic diagnosis of pulmonary diseases requires eminently reliable models. The model developed in this study shows that they can successfully segment the infected areas from various pulmonary diseases and lung regions. The top segmentation results in this study were shown by Mask-RCNN using backbone Resnet50, 100 epochs, and a learning rate configuration of 0.001. The values obtained are 91.98%, 85.25%, and 93.75% for DSC, MIoU, and mAP, respectively.

#### ACKNOWLEDGEMENTS

We want to express our gratitude toward Intelligent System Research Group (ISYSRG) Universitas Sriwijaya for supporting infrastructures.

#### REFERENCES

- [1] Xu, Y., Souza, L. F. F., Silva, I. C. L., Marques, A. G., Silva, F. H. S., Nunes, V. X., Han, T., Jia, C., de Albuquerque, V. H. C., & Filho, P. P. R. (2021). A soft computing automatic based in deep learning with use of fine-tuning for pulmonary segmentation in computed tomography images. *Applied Soft Computing*, 112, 107810. <https://doi.org/10.1016/J.ASOC.2021.107810>
- [2] Wang, D., Zhang, T., Li, M., Bueno, R., & Jayender, J. (2021). 3D deep learning based classification of pulmonary ground glass opacity nodules with automatic segmentation. *Computerized Medical Imaging and Graphics*, 88. <https://doi.org/10.1016/j.compmedimag.2020.101814>
- [3] Verma, D., Bose, C., Tufchi, N., Pant, K., Tripathi, V., & Thapliyal, A. (2020). An efficient framework for identification of Tuberculosis and Pneumonia in chest X-ray images using Neural Network. *Procedia Computer Science*, 171, 217–224. <https://doi.org/10.1016/J.PROCS.2020.04.023>
- [4] Prakash, N. B., Murugappan, M., Hemalakshmi, G. R., Jayalakshmi, M., & Mahmud, M. (2021). Deep transfer learning for COVID-19 detection and infection localization with superpixel based segmentation. *Sustainable Cities and Society*, 75. <https://doi.org/10.1016/j.scs.2021.103252>
- [5] Mamalakis, M., Swift, A. J., Vorselaars, B., Ray, S., Weeks, S., Ding, W., Clayton, R. H., Mackenzie, L. S., & Banerjee, A. (2021). DenResCov-19: A deep transfer learning network for robust automatic classification of COVID-19, pneumonia, and tuberculosis from X-rays. *Computerized Medical Imaging and Graphics*, 102008. <https://doi.org/10.1016/j.compmedimag.2021.102008>
- [6] Park, B., Park, H., Lee, S. M., Seo, J. B., & Kim, N. (2019). Lung Segmentation on HRCT and Volumetric CT for Diffuse Interstitial Lung Disease Using Deep Convolutional Neural Networks. *Journal of digital imaging*, 32(6), 1019–1026. <https://doi.org/10.1007/s10278-019-00254-8>
- [7] Oulefki, A., Aghaian, S., Trongtirakul, T., & Kassah Laouar, A. (2021). Automatic COVID-19 lung infected region segmentation and measurement

**Muhammad Arnaldo, Siti Nurmaini,  
Hadipurnawan Satria, Muhammad Naufal Rachmatullah**  
**Multiclass Segmentation of Pulmonary Diseases using Convolutional Neural Network**

- using CT-scans images. *Pattern Recognition*, 114. <https://doi.org/10.1016/j.patcog.2020.107747>
- [8] M. Liu, J. Dong, X. Dong, H. Yu and L. Qi, "Segmentation of Lung Nodule in CT Images Based on Mask R-CNN," 2018 9th International Conference on Awareness Science and Technology (iCAST), 2018, pp. 1-6, doi: 10.1109/ICAwST.2018.8517248.
- [9] Siti Nurmaini, Alexander Edo Tondas, Radiyati Umi Partan, Muhammad Naufal Rachmatullah, Annisa Darmawahyuni, Firdaus Firdaus, Bambang Tutuko, Rachmat Hidayat, and Ade Iriani Sapitri, "Automated Detection of COVID-19 Infected Lesion on Computed Tomography Images Using Faster-RCNNs," *Engineering Letters*, vol. 28, no.4, pp1295-1301, 2020
- [10] Kopelowitz, E., & Engelhard, G. (2019). Lung Nodules Detection and Segmentation Using 3D Mask-RCNN. <http://arxiv.org/abs/1907.07676>
- [11] Gao, X. W., James-Reynolds, C., & Currie, E. (2020). Analysis of tuberculosis severity levels from CT pulmonary images based on enhanced residual deep learning architecture. *Neurocomputing*, 392, 233–244. <https://doi.org/10.1016/j.neucom.2018.12.086>
- [12] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. <http://arxiv.org/abs/1703.06870>
- [13] Rosenthal A, Gabrielian A, Engle E, Hurt DE, Alexandru S, Crudu V, et al. The TB Portals: an Open-Access, Web-Based Platform for Global Drug-Resistant-Tuberculosis Data Sharing and Analysis. *J Clin Microbiol*. 2017;55(11):3267-82
- [14] Armato, S. G., Mclellan, G., Bidaut, L., Mcnitt-Gray, M. F., Meyer, C. R., Reeves, A. P., Zhao, B., Aberle, D. R., Henschke, C. I., Hoffman, E. A., Kazerooni, E. A., Macmahon, H., van Beek, E. J. R., Yankelevitz, D., Biancardi, A. M., Bland, P. H., & Brown, M. S. (2011). The Lung Image Database Consortium "LIDC... and Image Database Resource Initiative "IDRI...: A Completed Reference Database of Lung Nodules on CT Scans.
- [15] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. <http://arxiv.org/abs/1506.01497>.
- [16] Jaiswal, A. K., Tiwari, P., Kumar, S., Gupta, D., Khanna, A., & Rodrigues, J. J. P. C. (2019). Identifying pneumonia in chest X-rays: A deep learning approach. *Measurement: Journal of the International Measurement Confederation*, 145, 511–518. <https://doi.org/10.1016/j.measurement.2019.05>
- [17] Nurmaini, S., Rachmatullah, M. N., Sapitri, A. I., Darmawahyuni, A., Jovandy, A., Firdaus, F., Tutuko, B., & Passarella, R. (2020). Accurate Detection of Septal Defects With Fetal Ultrasonography Images Using Deep Learning-Based Multiclass Instance Segmentation. *IEEE Access*, 8, 196160–196174. <https://doi.org/10.1109/ACCESS.2020.3034367>
- [18] Zuo, L., He, P., Zhang, C., & Zhang, Z. (2020). A robust approach to reading recognition of pointer meters based on improved mask-RCNN. *Neurocomputing*, 388, 90–101. <https://doi.org/10.1016/j.neucom.2020.01.032>
- [19] Gamage, H. V. L. C., Wijesinghe, W. O. K. I. S., & Perera, I. (2019). Instance-Based Segmentation for Boundary Detection of Neuropathic Ulcers Through Mask-RCNN. *Lecture Notes in Computer Science (Including Subseries*



- Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11731 LNCS, 511–522. [https://doi.org/10.1007/978-3-030-30493-5\\_49](https://doi.org/10.1007/978-3-030-30493-5_49)
- [20] Zhang, Z., Yin, X., & Yan, Z. (2022). Rapid data annotation for sand-like granular instance segmentation using mask-RCNN. *Automation in Construction*, 133. <https://doi.org/10.1016/j.autcon.2021.103994>
- [21] Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2016). Feature Pyramid Networks for Object Detection. <http://arxiv.org/abs/1612.03144>
- [22] Evain, E., Raynaud, C., Ciofolo-Veit, C., Popoff, A., Caramella, T., Kbaier, P., Balleyguier, C., Harguem-Zayani, S., Dapvriil, H., Ceugnart, L., Monroc, M., Chamming's, F., Doutriaux-Dumoulin, I., Thomassin-Naggara, I., Haquin, A., Charlot, M., Orabona, J., Fourquet, T., Bousaid, I., ... Olivier, A. (2021). Breast nodule classification with two-dimensional ultrasound using Mask-RCNN ensemble aggregation. *Diagnostic and Interventional Imaging*, 102(11), 653–658. <https://doi.org/10.1016/j.diii.2021.09.002>
- [23] Khan, M. A., Akram, T., Zhang, Y. D., & Sharif, M. (2021). Attributes based skin lesion detection and recognition: A mask RCNN and transfer learning-based deep learning framework. *Pattern Recognition Letters*, 143, 58–66. <https://doi.org/10.1016/j.patrec.2020.12.015>
- [24] Shenavarmasouleh, F., & Arabnia, H. R. (2020). DRDr: Automatic Masking of Exudates and Microaneurysms Caused By Diabetic Retinopathy Using Mask R-CNN and Transfer Learning. <http://arxiv.org/abs/2007.02026>
- [25] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. <http://arxiv.org/abs/1512.03385>