

Pneumothorax Segmentation: Deep Learning Image Segmentation to predict Pneumothorax

Karan Jakhar, Rohit Bajaj, Ruchika Gupta
Computer Science and Engineering Department
Chandigarh University
Chandigarh, India

karanjakhar49@gmail.com, rohitbajaj.cse@cumail.in, rgupt009@gmail.com

Abstract—Computer vision has shown promising results in the medical image processing. Pneumothorax is a deadly condition and if not diagnosed and treated at time then it causes death. It can be diagnosed from chest X-ray images. We need an expert and experienced radiologist to predict whether a person is suffering from pneumothorax or not by looking at the chest X-ray images. Everyone does not have access to such facility. Moreover, in some cases we need quick diagnoses. So we propose a image segmentation model to predict and give output a mask which will assist the doctor in taking this crucial decision. Deep Learning has proved their worth in many areas and outperformed many state-of-the-art models. We want to use the power of these deep learning model to solve this problem. We have used U-net [13] architecture with ResNet [17] as backbone and achieved promising results. U-net [13] performs very well in medical image processing and semantic segmentation. Our problem falls in the semantic segmentation category.

Index Terms—Image Segmentation; Pneumothorax; Deep Learning; Semantic Segmentation; Medical Image Processing; Computer Vision;

I. INTRODUCTION

Pneumothorax is a life-threatening condition which is present in around 30-39% of patients who are suffering from chest trauma [1], [2]. One way of diagnosing it with chest radiography. Pneumothorax is a condition in which air from lungs leaks to space between lungs and chest wall. This exert a force on the lungs and stop them from expanding. Due to this the person can not breathe and if not treated then it causes death. However, it is a very challenge task and often fail to identify a pneumothorax and even reaching false-negative rate to 50% [3], [4]. If we fail to treat an enlarging pneumothorax in time then it may result in patient death. Chest X-ray is a way to diagnose if a patient is have pneumothorax or not. But to do so we need well qualified and experienced radiologists. Everyone can not have access to such facility. So we want to use technology here. Today, technology have power to solve any problem. We just need to use it in that way. Advancement in Deep Learning in recent year make it is to work with image data. Chest X-ray images will be processed by our model and it will predict whether a person is suffering from pneumothorax or not. We also want to perform image segmentation on it and give a mask as output so that doctor could know which area is problematic. This also gives a sense to doctor why the model is predicting so. In recent years, deep convolution networks have outperformed other the state

of the art model in many visual recognition tasks, e.g. [7], [8]. Though convolutional networks were already there for a long time [9] but their success was limited as we didn't have large amount of data and computation power to handle large networks. In 2012 Krizhevsky et al. [10] make a breakthrough by supervised training of a large network having 8 layers and millions of parameters on the ImageNet dataset which have 1 million training images. After that many deeper networks have been introduced and a lot of work has been done on it [11]. Having a model which can predict the weather a person is have Pneumothorax problem or not is very helpful for doctors to start the treatment at the earliest. Deep learning is a very powerful tool and we should use its power to solve the problems which can save some lives. We prepared a model which can be easily deployed and used by the doctors to get real time results. It will not only predict weather a person is have pneumothorax but also indicated the problem area. To predict if a patient is suffering from Pneumothorax just by read the X-rays required a lot of expertise in the field. We need experience and well qualified radiologists to do this. But everyone doesnt have access to a radiologist and it is expensive also [12]. This process of prediction by a human being is very time consuming.

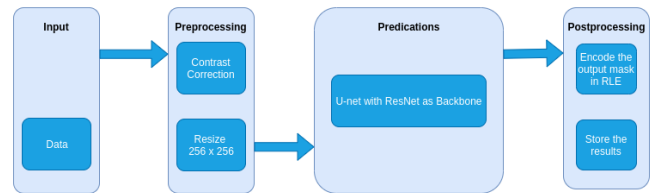


Fig. 1. Proposed Framework

Using technology to solve medical problems is great way to utilize its power. Today, we have such strong models and hardware to solve many problems which seems impossible without them. We should develop such systems which can assist doctors to take crucial decisions. We propose a system which will help the doctors to take decision whether a patient has pneumothorax or not. Our model will give a mask showing the problematic area due to which our model is predicting that patient has pneumothorax. After having a look at the mask the

doctor can easily take decision whether the patient is suffering from pneumothorax and what treatment is required.

II. RELATED WORK

There is a lot of work done in the field of Image Segmentation and Pneumothorax independently. Image Segmentation is used before in medical imaging but it is not yet used for Pneumothorax. We have U-net [13] which is used in ISBI challenge and won the competition. The task was to segment neuronal structures present in electron microscopic stacks. Segmentation of a 256 X 256 or greater can be done in microseconds on a recent GPU. Our problem falls in category of semantic segmentation [24]. Before deep learning methods there were many techniques used to segment an object in an image for example region based, edge detection etc. RCNN [25] in 2015 used deep learning architecture to successfully achieve state-of-the-art results at that time. Its main focus was on object detection and drawing bounding boxes around the object. Later on it was improved further with introduction of Fast-RCNN [26] after that we had Faster-RCNN [27]. In 2018 FAIR(Facebook AI Research) introduced Mask-RCNN [28] which has amazing results on COCO dataset [29]. Its main focus was on Instance Segmentation. There are many image segmentation techniques are available in the literature. U-net has been used a lot in medical imaging as it shows promising results. U-net [13] uses concept of skip connection as described in ResNet model architecture [17]. With skip connection we avoid problem of vanishing gradient as we go deeper and also preserve the more information when upsampling. U-net works on the concept of downsampling and upsampling like encoder-decoder models. Mask R-CNN is simple and addition to Faster R-CNN which generate two outputs corresponding to each candidate object, a class label and a bounding-box offset and in case of Mask R-CNN we add a third branch which generates the object mask. I am combining best work from both fields and produce useful results which can be used by doctors to diagnose Pneumothorax as early as possible and keeping false-negative low. U-net is more related to the generative models with little difference that is instead of learning to generate the same image, it learns to generate the mask. A lot of work and experimentation is done on generative models. We have guidelines to improve generative models. Pneumothorax is a well known disease and it is very tough in some situations for a doctor to come to a conclusion that the person is suffering from Pneumothorax. Thats why having a deep learning model assisting the doctor in taking decision is very helpful. As this result will be a mask at the problematic area so the doctor can consider his decision easily.

III. MATERIAL AND METHODS

A. Proposed Framework

The output of U-net model with pretrained weights of the ResNet is fast and accurate. This helps doctors to start the treatment at the earliest. U-net is consist of an encoder and decoder. Encoder narrow down the information of the image to a latent space. Then decoder taking input latent space

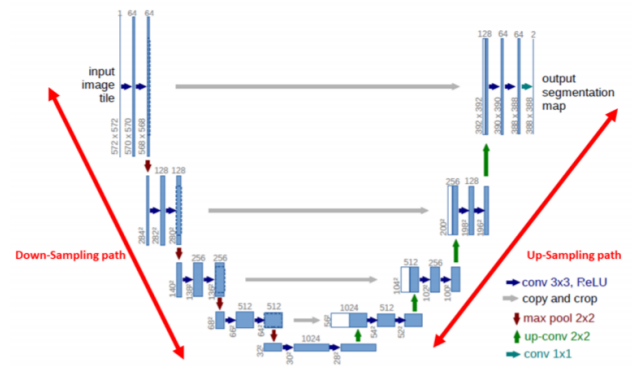


Fig. 2. Classical U-net Architecture

regenerates the image. Here the model will be generating the mask. The mask tells us were is Pneumothorax if it exists. Image preprocessing is an important part here. As real world images are not of the same size and more sometime are corrupted. In preprocessing we have to resize the image in 256 X 256. Then the image is put into correct color range. The image will be normalized. For training we filtered out the corrupted images and done random crop to remove unwanted boundary. As X-ray images usually have a lot of boundary noise which is not required for our model. The output mask is a black and white image. In which white area indicated the position of the problem. We then map this mask image to out original image to get exact position on the X-ray image. The encoder part downscale the images and learn a latent space with help of which we generate the masks. This is kind of mapping an image to a mask and training our model to learn to generalize the mapping. When the model successfully learns to generalize it, we can feed any X-ray image and get the mask for it. The mask will indicate the position of the part which is potentially causing Pneumothorax. If the mask is blank means there is no problem and patient is safe. Prediction should be very fast because if it take too long then it can not help the doctor in needed time. U-net is pretty small network and very effective also. We are using ResNet as encoder and we are going to use the weights from pretrained network. Using weights from pretrained network helps to boost the result and saves us time. Skip connections of U-net help to retain useful features learned in encoder layer while upsampling by decoder. ResNet concept of skip connection helps fight vanishing gradient problem. Thats why decided to try the ResNet pretrained model as backbone. It improves the result.

B. Segmentation Techniques

There has been a lot of shift in approaches and techniques used to perform segmentation. Before Deep Learning was a thing, various methods were adapted which were not efficient and were not giving the required results. With Deep Learning entering into the picture various architectures have been developed. Unet is one such architecture concept which works well in medical image segmentation. Not every architecture works best for every problem. There have been many shifts in state-

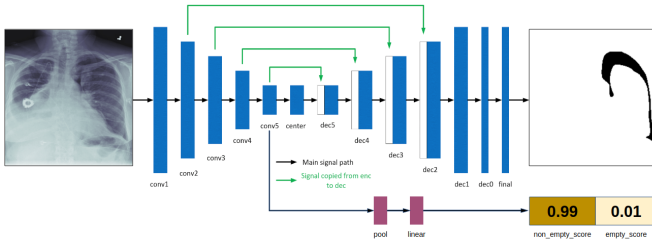


Fig. 3. U-net Architecture Used

of-the-art models in image segmentation after the arrival of deep learning. Image Segmentation falls into object detection problem. Earlier researchers found solutions which will find the object in the image, predict its class and draw a bounding box around it. In Image Segmentation we predict a masks for an object in the image. Below we will discuss a few of the approaches used for segmentation. There are different type of segmentation such as semantic segmentation and instance segmentation. Our problem falls into semantic segmentation.

a) *Semantic Segmentation*: In this we classify each and every pixel of the image into a class. Object belonging to same class will be given same colour. It does not classify objects belonging to same class differently. Instance Segmentation It is different then Semantic Segmentation in context that if classify each object in the image differently whether it belongs to same class or not. Every object gets a label to it. It is also we tough to achieve and require a lot of different techniques to get desired results.

b) *Region based segmentation*: In this type of segmentation we divide regions with its pixel value. This works well when we have a sharp contrast between object and background. We can set a threshold value and pixel values falling below and above it can to categorized accordingly. This technique is called threshold segmentation [30].

c) *Edge Detection*: We use filters and convolutions to identify the edge or boundary between two region. When we move from one region to another in an image they have a edge that separates them. If we can find that edge then it is very easy to separate them. We use filters for that. We have different types of filters which can detect a horizontal, vertical or both edges.

d) *Clustering*: In this we divided the polulation into different groups. We can divided pixels in different groups and like this we can segment them. For example we can use K-Means cluster algorithm and can easily form different clusters of similar data points or pixels. It is not suitable for large amount of data.

e) *FCN(fully convolutional network)* : This type of model uses convolutional and pooling layers to map an image to masks. These masks are used to produce the segmented image. The concept is simple. First image is downsampled using convolutional and pooling layers then upsampled in to the same size but generating the masks [14].

f) *Mask R-CNN*: This is state-of-the-art model for instance segmentation. It extends Faster R-CNN[16]. Faster R-CNN [16] gives bounding box and label for the class but Mask R-CNN [15] has added another branch which will give mask of the object. The concept is simple first we extract feature out of image then regions are proposed. Then we find region of interest. Going according to this process the model generates label, bounding box, and mask for the object.

IV. EXPERIMENTAL INVESTIGATION

A. Dataset

The dataset download from Kaggle [18] which has training image and their masks. The masks are encoded in RLE(Run Lenght Encoding) and we have used the same algorithm to decode the mask image. We kept some of the training images for testing and validating the model after training. Dataset has good amount of the images. In the dataset we have 12047 images and same number of masks. Images are in shape 256 x 256. If an image does not have Pneumothorax then its mask is all black. The all black mask image teaches the model which X-ray images does not have Pneumothorax. We will be keeping 20% of the images for testing and validating the model.

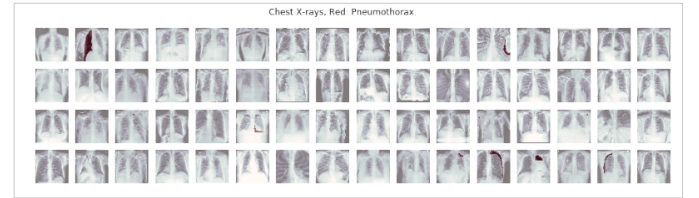


Fig. 4. Chest X-ray Images

B. Experimental Setting

Python programming language is simple, fast and efficient way to implement deep learning models. We used Keras [22] library in Python which is high level API. Keras is fast and at the backend it uses Tensorflow [21]. Due to Keras we need not to write a lot of repetitive code. Image processing is also a great part of the experiment. We used OpenCV [23] to process the image. The major operation we performed is contrast correction. The code becomes concise and easy to interpret. The model is implemented in parts. First encoder, then decoder is implemented. We used pretrained weights of ResNet model. Therefore, ResNet is used as encoder model. Without pretrained weights the result is not good. Deep learning models are data hungry and need a lot of computation power to train. As we are training deep learning model we need a lot of computation power. We need good GPU. It is very time consuming to train the model on a CPU. We used Adam optimizer with learning rate 0.0001, binary cross-entropy loss and early stop validation. To get a clear output mask we used a threshold techniques to remove the noise and save only highly supported pixel values. Without setting threshold and filtering pixels we get little noise in some mask and output masks are

not that sharp. Careful post-processing is also required. We used OpenCV [23] to perform this process.

V. RESULTS AND DISCUSSION

A. Performance Evaluation

Choosing right metric is very important to evaluate a model. We evaluated our model on Dice Coefficient [19], IOU [20] which are considered best metrics to evaluate a segmentation result. There are various metrics to evaluate segmentation. But we are using dice coefficient as described in [19]. Checking result by visualizing is good but when we are train we need some metrics to monitor the model. We use dice and IoU because are proven best when we have to check similarity between two images. And here we want to check how much our output mask is similar to real mask. The formulae of these score metrics is easy to understand and are shown in figures 5(a) 5(b).

$$D = \frac{2|X \cap Y|}{|X| + |Y|} \quad (1)$$

$$IoU = \frac{|X \cap Y|}{|X \cup Y|} \quad (2)$$

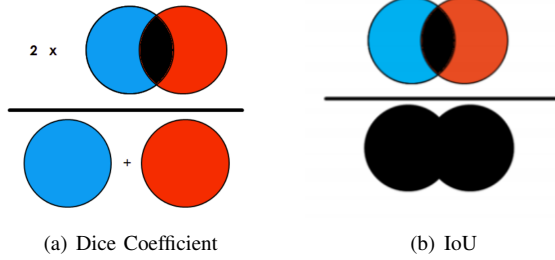


Fig. 5. Metrics Illustrations

B. Experimental Results

Our model is showing promising results as we can see the masks predicted by the model in figures 8(a) 8(b) 8(c) match closely to the real annotation. Our Dice score and IoU score are 84.3 and 82.6 respectively. Both of these metrics are used to indicate the similarity between images. And in figures 9(a) 9(b) 9(c) 10(a) 10(b) 10(c) we can see the results are very good and insightful. Both Dice score and IoU indicates that model is strong. Dice score and IoU both are positively correlated. It means if one say a model is strong then other will also support the decision. Using ResNet as backbone in U-net architecture helps converse the model fast and improves the results. The implementation is also easy in Keras. There are other metrics also but these metrics are considered best to test similarity between two images.

TABLE I
SCORES

| Dice Coefficient | IoU |
|------------------|------|
| 84.3 | 82.6 |

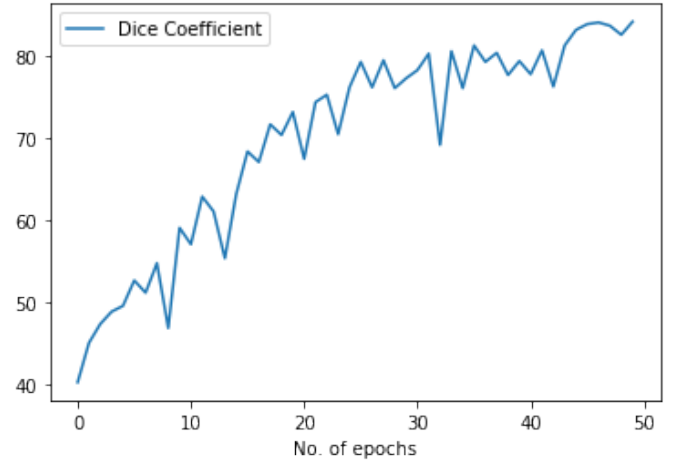


Fig. 6. Dice Coefficient Validation score

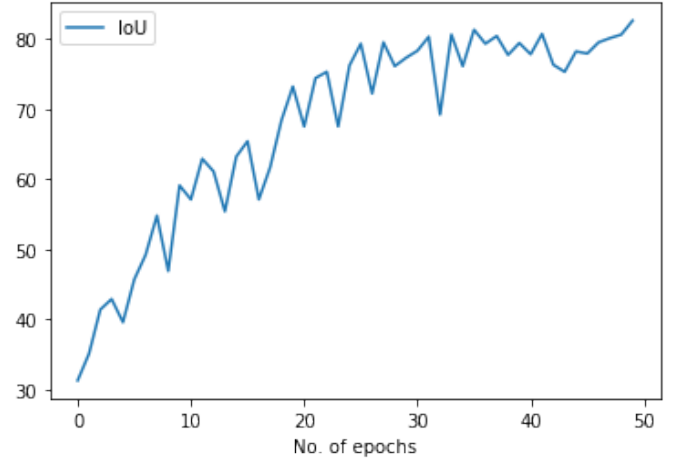


Fig. 7. IoU Validation score

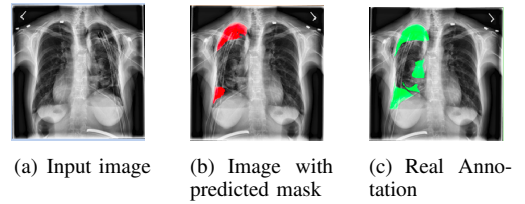


Fig. 8. Patient 1

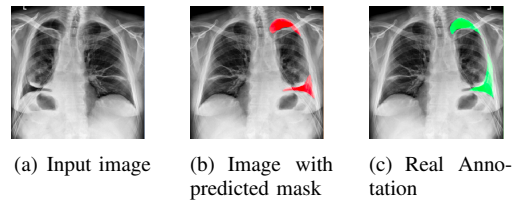


Fig. 9. Patient 2

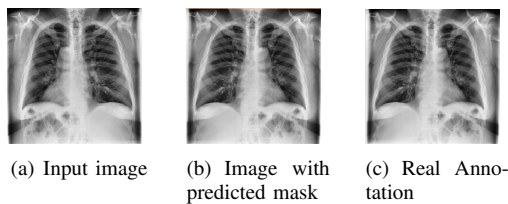


Fig. 10. Patient 3

CONCLUSION AND SCOPE

Deep learning is a powerful tool. If it can be used effectively then it could help us to tackle a lot of problem. Using segmentation to predict Pneumothorax provides a great help to doctors to deal with it. Many people die due to Pneumothorax because it took a lot of time to diagnose the problem. With help of this great technology we can diagnose faster and with great accuracy. Our proposed model is fast and accurate. It can easily be deployed and used. Using U-net with pretrained ResNet weights is an effective way to get a great model in less time. With advancement in hardware we can make this prediction in seconds. This accurate and fast result will help doctors a lot to take decision and start their treatment. The model performs very well and training time is also decreased due to ResNet pretrained weights. Using ResNet as backbone helps model to converse fast and it also deals with various problems also like vanishing gradient. Computer vision has many applications in medical field this is just one of them. There are many to explore further.

REFERENCES

- [1] Shorr RM, Crittenden M, Indeck M, Hartunian SL, Rodriguez A. Blunt thoracic trauma: analysis of 515 patients. *Ann Surg* 1987; 206:200-205 [Crossref] [Medline] [Google Scholar]
- [2] Tekinbas C, Erolu A, Krkolu IC, Trkylmaz A, Yekeler E, Karaolanolu N. Chest trauma: analysis of 592 cases [in Turkish]. *Ulus Travma Acil Cerrahi Derg* 2003; 9:275-280 [Medline] [Google Scholar]
- [3] Farooq U, Raza W, Zia N, Hanif M, Khan MM. Classification and management of chest trauma. *J Coll Physicians Surg Pak* 2006; 16:101-103 [Medline] [Google Scholar]
- [4] Trupka A, Waydhas C, Hallfeldt KK, Nast-Kolb D, Pfeifer KJ, Schweiberer L. Value of thoracic computed tomography in the first assessment of severely injured patients with blunt chest trauma: results of a prospective study. *J Trauma* 1997; 43:405-411; discussion 411-412 [Crossref] [Medline] [Google Scholar]
- [5] Ball CG, Kirkpatrick AW, Laupland KB, et al. Incidence, risk factors, and outcomes for occult pneumothoraces in victims of major trauma. *J Trauma* 2005; 59:917-924; discussion 924-925 [Crossref] [Medline] [Google Scholar]
- [6] Ball CG, Kirkpatrick AW, Laupland KB, et al. Factors related to the failure of radiographic recognition of occult posttraumatic pneumothoraces. *Am J Surg* 2005; 189:541-546; discussion 546 [Crossref] [Medline] [Google Scholar]
- [7] Bridges KG, Welch G, Silver M, Schinco MA, Esposito B. CT detection of occult pneumothorax in multiple trauma patients. *J Emerg Med* 1993; 11:179-186 [Crossref] [Medline] [Google Scholar]
- [8] Kanz KG, Korner M, Linsenmaier U, et al. Priority-oriented shock trauma room management with the integration of multiple-view spiral computed tomography [in German]. *Der Unfallchirurg* 2004; 107:937-944 [Crossref] [Medline] [Google Scholar]
- [9] Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: NIPS. pp. 11061114 (2012)
- [10] Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2014)
- [11] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition (2014), arXiv:1409.1556 [cs.CV]
- [12] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W.Jackel, L.D.: Backpropagation applied to handwritten zip code recognition. *Neural Computation* 1(4), 541551 (1989)
- [13] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical image computing and computer- assisted intervention*. Springer, Cham, 2015.
- [14] J. Long, E. Shelhamer and T. Darrell, "Fully convolutional networks for semantic segmentation," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 3431-3440. doi: 10.1109/CVPR.2015.7298965
- [15] He, Kaiming, et al. "Mask r-cnn." *Proceedings of the IEEE international conference on computer vision*. 2017.
- [16] Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems*. 2015.
- [17] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep Residual Learning for Image Recognition, arXiv:1512.03385 [cs.CV]
- [18] <https://www.kaggle.com/jesperdramsch/siim-acr-pneumothorax-segmentation-data>
- [19] Tustison, N. J., and J. C. Gee. "Introducing Dice, Jaccard, and other label overlap measures to ITK." *Insight J* 2 (2009).
- [20] Thada, Vikas, and Vivek Jaglan. "Comparison of jaccard, dice, cosine similarity coefficient to find best fitness value for web retrieved documents using genetic algorithm." *International Journal of Innovations in Engineering and Technology* 2.4 (2013): 202-205.
- [21] <https://www.tensorflow.org/>
- [22] <https://keras.io/>
- [23] <https://opencv.org/>
- [24] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
- [25] Lenc, Karel, and Andrea Vedaldi. "R-cnn minus r." *arXiv preprint arXiv:1506.06981* (2015).
- [26] Girshick, Ross. "Fast r-cnn." *Proceedings of the IEEE international conference on computer vision*. 2015.
- [27] Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems*. 2015.
- [28] He, Kaiming, et al. "Mask r-cnn." *Proceedings of the IEEE international conference on computer vision*. 2017.
- [29] <http://cocodataset.org/>
- [30] Buie, Helen R., et al. "Automatic segmentation of cortical and trabecular compartments based on a dual threshold technique for in vivo micro-CT bone analysis." *Bone* 41.4 (2007): 505-515.