Malaysian Journal of Computing, 4 (1): 225-236, 2019

eISSN: 2600-8238 online

CHEST X-RAY IMAGE CLASSIFICATION USING FASTER R-CNN

Taufik Rahmat[1], Azlan Ismail[1,2], Sharifah Aliman[1,3]

[1]*Faculty of Computer and Mathematical Sciences*
[2]*Knowledge and Software Engineering Research Group*
[3]*Advanced Analytics Engineering Centre*
Universiti Teknologi MARA (UiTM) 40450 Shah Alam,
Selangor, Malaysia
taufikrahmat84@gmail.com, azlanismail@tmsk.uitm.edu.my,
sharifahali@tmsk.uitm.edu.my

ABSTRACT

Chest x-ray image analysis is the common medical imaging exam needed to assess different pathologies. Having an automated solution for the analysis can contribute to minimizing the workloads, improve efficiency and reduce the potential of reading errors. Many methods have been proposed to address chest x-ray image classification and detection. However, the application of regional-based convolutional neural networks (CNN) is currently limited. Thus, we propose an approach to classify chest x-ray images into either one of two categories, pathological or normal based on Faster Regional-CNN model. This model utilizes Region Proposal Network (RPN) to generate region proposals and perform image classification. By applying this model, we can potentially achieve two key goals, high confidence in the classification and reducing the computation time. The results show the applied model achieved higher accuracy as compared to the medical representatives on the random chest x-ray images. The classification model is also reasonably effective in classifying between finding and normal chest x-ray image captured through a live webcam.

Keywords: *Image Classification, Chest X-ray Analysis, CNN, faster R-CNN, Region Proposal Network*

1. Introduction

Chest radiography is one of the primary diagnostic imaging procedures in medical image diagnosis. The aim is to evaluate and detect for pulmonary diseases such as cardiomegaly, pneumothorax, pleural thickening. This procedure is heavily depended on the radiologists which at certain situation may negatively affect the quality and productivity of radiologists' tasks. Furthermore, in the situation where radiologist is absent, the interpretation made by the respective medical officer may not be accurately analyzed and thus requires for the *second opinion*.

The advancement of medical image analysis has emerged into Computer-Aided Diagnosis(CAD) (Doi, 2007). The main purpose is to assist in finding the location of a lesion and also to estimate the probability of a disease in the chest x-ray image. One of the key technological aspects for CAD is classifying and detecting diseases based on machine learning.

The most promising machine learning technique for medical imaging is recognized as deep learning. For instance, the work by Agarwal et al. (Agarwal, Balasubramanian, & Jawahar, 2018) addressed multiclass classification problems with the aim to improve the accuracy of the existing deep learning models. The key idea is to breakdown the network into individual binary problems.

Deep learning is also widely researched for chest x-ray image classification (Ker, Wang, Rao, & Lim, 2018), in particular, Convolutional Neural Networks (CNN). This fact is supported by the existing related works who have studied the application of CNN for chest x-ray image classification and detection especially when a few datasets has been released including ChestX-ray14 by Wang et al. (X. Wang et al., 2017a). Among these works are the application of Deep CNN (Rajpurkar et al., 2017a), 2D CNN (Yao et al., 2017), Attention Guided CNN (Guan et al., 2018), Location Aware Dense

Networks (Gundel et al., 2018), and weakly supervised CNN (Sedai, Mahapatra, Ge, Chakravorty, & Garnavi, 2018).

Furthermore, we also noticed the advancement of CNN based on regional paradigm to tackle the object detection and classification. It begins with a model called Regional-CNN (R-CNN) (Girshick, Donahue, Darrell, & Malik, 2014) followed with Fast R-CNN (Girshick, 2015) and Faster R-CNN (Ren, He, Girshick, & Sun, 2017). *Regional* paradigm aims to improve the precision of object detection, whilst *Fast* and *Faster* targets to improve the performance of the entire models. A few works have been done that applied R-CNN and Faster R-CNN in medical imaging area. Our work complements the existing in demonstrating the application of Faster R-CNN for addressing chest x-ray classification. Thus, with the advantages of Faster R-CNN, we believe that it will bring huge benefits to the area of medical imaging analysis.

Therefore, we propose an approach to apply Faster R-CNN for chest x-ray image classification and detection. We use a subset of data from chestX-ray14 to support this study. A small dataset was taken because we have limited resource to mark the regions for each x-ray image. Existing python libraries and data analytic platform are adopted to train and test the classification model. Furthermore, we evaluate the model by observing its loss function, assessing its effectiveness, performance, and time execution. The assessment of performance and time execution involve the comparison with medical representatives. Meanwhile, the effectiveness is evaluated through live web-cam. We believe this work can potentially contribute to the state-of-the-art of chest x-ray image analysis based on CNN and its evolution.

This paper is organized as follows. Section 2 describes about CNN and related regional paradigm of CNN. Section 3 explains the methodology used to carry out this research. Section 4 presents our results. Section 5 highlights some related works. Finally, section 6 provides the conclusion and outlines the future work.

2. Background

In this section, we present the background of the applied techniques for this research.

2.1 Convolutional Neural Networks

Convolutional Neural Networks also known as CNN is a type of deep learning architectures for processing huge data and is getting widely applied for addressing medical imaging analysis challenges (Litjens et al., 2017) (Ker et al., 2018). The advantages of CNN over other machine learning methods can be viewed from two aspects (Yamashita, Nishio, Do, & Togashi, 2018). Firstly, it does not require hand-crafted feature extraction. Secondly, its architectures do not necessarily require certain segmentation by human experts.

CNN architectural model is conceptually shown as in Figure 1. It contains several building blocks, namely, the *input layer*, the *hidden layers* that contain convolution with RELU layer and *pooling layer*, the *fully connected layer* and the *output layer*. The *input layer* is the stage of loading the data input such as an image file into the computational platform. The *convolutional layer* contains convolution operation to extract features. In particular, it takes the input data and filter them using different kind of convolution filters to produce a set of feature maps. Then, the RELU is performed with a nonlinear activation function that sets negative input values to zero. Once completed, the *pooling layer* is conducted to execute down sampling operation (e.g. max pooling or global average pooling) onto the produced feature maps.

The convolution layer and pooling layer will be done in a few cycles. The outcome of each cycle of pooling layer is a flattened output, which is a one-dimensional array of numbers, or simply a vector. This vector becomes the input to form connected networks in the *fully connected layer*, where the learning process begins by determining the relationship between features. The result of the learning process is the answer of a specific class or label, as part of the *output layer*.

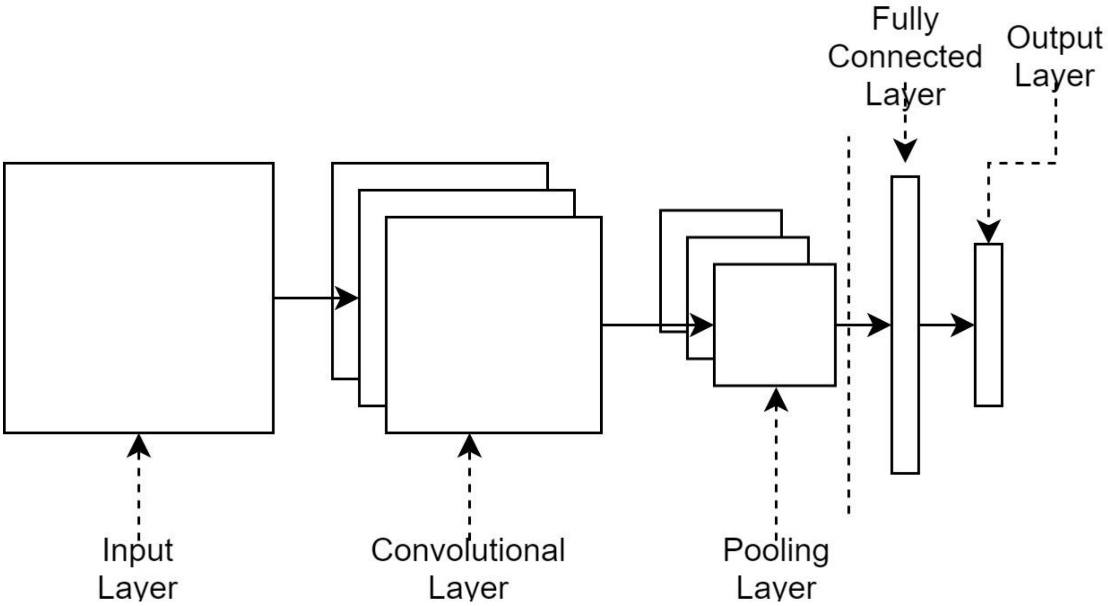

Figure 1. Convolutional Neural Network (CNN) Architecture

2.2 Regional Convolutional Neural Networks

Regional CNN (R-CNN) was introduced by Girshick et al. (Girshick et al., 2014) to address the object detection or localization problem of CNN. Previous localization approach for CNN is based on the sliding-window paradigm, however it suffers from achieving good precision for localization when dealing with higher number of convolutional layers. Hence, the authors proposed to apply the region paradigm (Chunhui Gu, Lim, Arbelaez, & Malik, 2009) to solve the CNN localization problem.

The design principle of R-CNN consists of three modules (Girshick et al., 2014). The first module aims to generate a set of category-independent region proposals based on the selective search (Uijlings, van de Sande, Gevers, & Smeulders, 2013) which is a search technique that combines the best of the intuitions of segmentation and exhaustive search.

The segmentation is needed to make use the structure of the image to generate object locations. Meanwhile, the exhaustive search is used to capture all possible object locations. The second module aims to extract features of the images in the region using the Caffe implementation of the CNN (Krizhevsky, Sutskever, & Hinton, 2012). The third module aims to tighten the coordinates of the region using linear regression model.

2.3 Fast Regional Convolutional Neural Networks

R-CNN is claimed to be slow in performance (although it works as expected) mainly because of two reasons (Girshick, 2015). Firstly, it requires a forward pass of the CNN for every single region proposal for every single image, although the regions may be overlapped between each other. Secondly, it requires the execution of three different models separately, which are the feature extraction model, the classification model and the regression model.

Fast R-CNN was proposed to overcome the above challenges (Girshick, 2015). The improvement was made by concerning at the bigger picture as well as the extraction phase. Hence, there are two key improvements of Fast R-CNN. Firstly, as opposed to R-CNN, it combines all models into one network that consists of feature extraction, classification and detection. Secondly, the need to run CNN per region has been reduced to only once per image. These CNN features (i.e. a conv feature map) are shared across all regions at a specific layer called Region of Interest (RoI) pooling layer. Technically, RoI is a rectangular window into a convolution feature map. Thus, the outcome of RoI pooling (i.e. using max pooling) is the features of each region.

2.4 Faster Regional Convolutional Neural Networks

The term *faster* reflects the evolution of Fast R-CNN that is quicker in producing the required results. For Fast R-CNN that is derived from R-CNN, the region proposals generation occurs before the convolutional layers. This specific step is claimed to cause slower performance when dealing with large images.

Hence Faster R-CNN (Ren et al., 2017) was proposed to address the performance issue by introducing Region Proposal Network (RPN) layer and removing the existing region proposals generation at earlier stage. Technically, RPN is a fully-convolutional network that simultaneously predicts object bounds and its scores. RPN computation is proposed to occur after performing feature extraction. Therefore, the CNN features map become the input of RPN that determines a set of regions with the predicted scores.

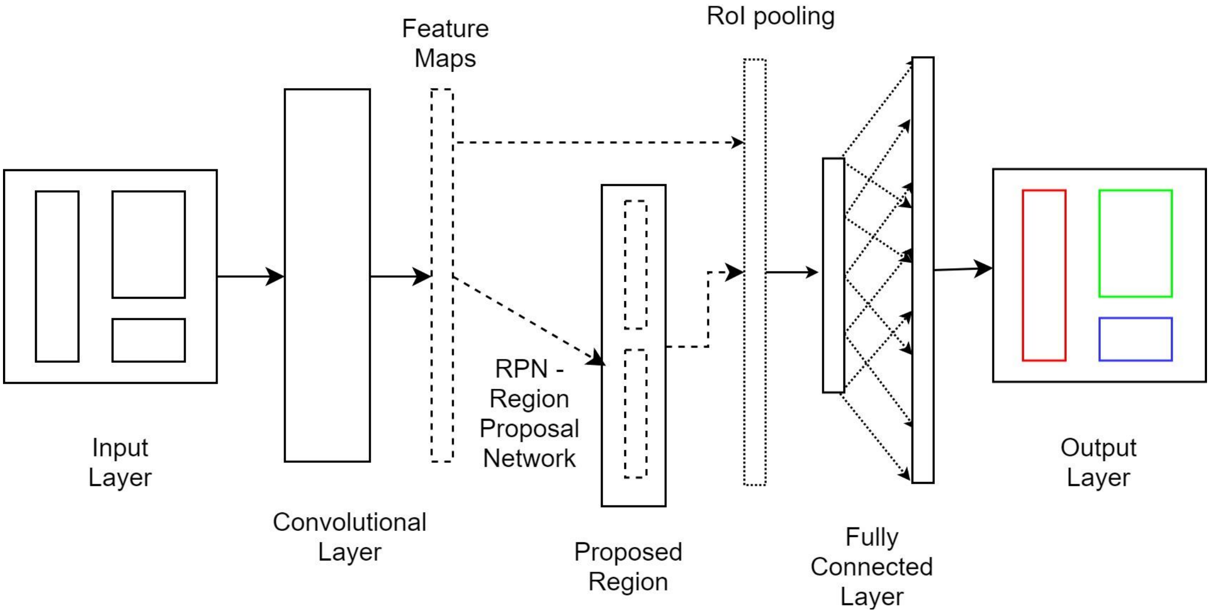

Figure 2. Faster Regional based Convolutional Neural Network Architecture

3. Background

In this section, we present the background of the applied technique for this research.

3.1 Data Preparation

A few chest x-ray image datasets can be used to research about chest x-ray image classification as mentioned in (Ajay Mittal, Hooda, & Sofat, 2017). In this work, we have chosen ChestX-ray14 dataset (X. Wang et al., 2017a) to complement the related works that have applied CNN for chest x-ray image classification and detection.

ChestX-ray14 dataset is extracted from the clinical PACS database at National Institutes of Health Clinical Center. It contains 112,120 frontal-view x-ray images of 30,805 unique patients. Each image has been annotated with one of 14 different thoratic pathology category. These images were scaled and standardized into 1024 x 1024 pixels. Then the images were changed to grayscale color since some of the images were in blue color.

In relation to this work, we have decided to select 200 chest x-ray images randomly because we have to manually set the regions' labels for the training purpose. For labeling, we have used *Labellmg* application. In addition, from 200 images, we ensure half of the images have finding, whilst another half of the images are normal. All images are in *png* format. We also have simplified them from 14 pathology categories into two categories which are *normal chest x-ray* and *pathological chest x-ray*. Table 1 shows the two categories of our simplification.

Table 1. Re-categorizing Chest X-ray14 dataset into two main categories

Existing Category	**New Category**
No finding	Normal chest x-ray
Atelectasis Consolidation Infiltration Pneumothorax Edema Emphysema Fibrosis Effusion Pneumonia Pleural thickening Cardiomegaly Nodule Mass Hernia	Pathological chest x-ray

3.2 Model Building and Training

The aim of the model is to address common classification problem (e.g. (Jamil, Mutalib, rahman, & Aziz, 2018)). For the model building and training, we have used Tensorflow library which is an open source library for high performance numerical computation. We also have used Anaconda virtual environment for training the model. In the case of Faster R-CNN, we have applied Faster R-CNN-Inception-V2 model (*Tensorflow detection model zoo - COCO-trained models*, 2017).

The dataset that we have selected is randomly split into 80% for training and 20% for testing. We then used the training dataset to train the model and testing dataset to assess the model. The loss values have recorded while the trained model progresses to understand the quality of the prediction.

Figure 3 shows a conceptual diagram on how the trained chest X-ray image classifier should work using Faster R-CNN architecture. The details of the process can be referred to section 2.4. The inputs are the chest x-ray images and the results are the classified images with predicted accuracy.

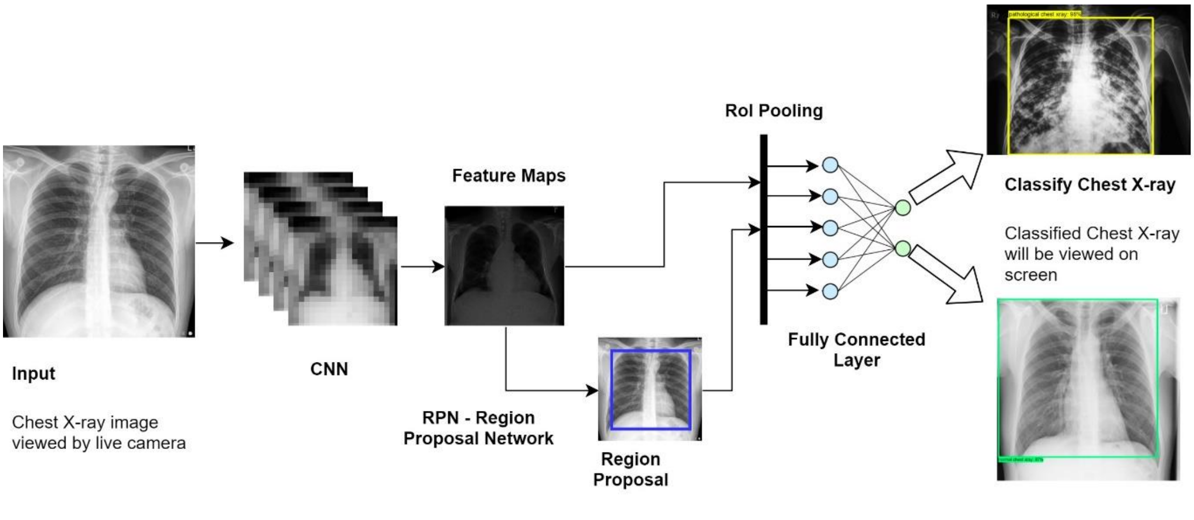

Figure 3. Chest X-ray Image Classification Process using Faster R-CNN

3.3 Model Testing and Validation

Model testing and validation phase focuses on three aspects, namely, *effectiveness, performance evaluation*, and *time comparison*. Each of them is discussed as follows:

- The *effectiveness* of the model is done through live webcam. Validated pictures will be marked with either yellow or green box with an accuracy percentage to show the confidence level of the prediction.
- For the model *performance evaluation*, the complete trained model is validated with another 100 randomly chest x-ray images taken from Chest X-ray14 dataset (X. Wang et al., 2017a). We then compute the confusion matrix to obtain four key values (Ajay Mittal et al., 2017); accuracy $AC = \frac{t_p+t_n}{t_p+t_n+f_n+f_p}$, precision $PR = \frac{t_p}{t_p+f_p}$, sensitivity $SE = \frac{t_p}{t_p+f_n}$, and specificity $SP = \frac{t_n}{t_n+f_p}$ to conclude its performance where true positive is t_p, false positive is f_p, true negative is t_n, and false negative is f_n. In addition, we also compare its performance with medical persons, specifically, one general practitioner *GP* and one medical student *MS*. These persons are chosen to represent a situation where radiologist is absence and thus require these persons to examine the chest x-ray images. They are also required to examine and classify 100 randomly chest x-ray images. We then compute their confusion matrices and compare them with the model performance.

- The *time comparison* is done by capturing the time taken by each, *model*, *GP* and *MS* to complete the examination and classification.

4. Results and Discussions

In this section, we presented the results of four aspects. We begin with the *loss observation* followed by the results of model validation, namely, *effectiveness*, *performance*, and *time comparison*.

4.1 Loss Observation

Convolutional Neural Networks also known as CNN is a type of deep learning architectures for processing huge data and is getting widely applied for addressing medical imaging analysis challenges (Litjens et al., 2017). The advantages of CNN over other machine learning methods can be viewed from two aspects (Yamashita et al., 2018). Firstly, it does not require hand-crafted feature extraction. Secondly, its architectures do not necessarily require certain segmentation by human experts.

The loss function is important to understand the quality of the prediction. Thus, we pay specific attention to three kinds of loss functions. Firstly, the *Total Loss* that shows all the loss functions involved in the training and testing model. The loss function will decrease when the training model is longer. Secondly, the *Regional Proposal Network Loss* that illustrates the performance of the model during the training phase and determines the quality of the region proposals. Thirdly, the *Classification Loss* that represents the loss function during the classification phase and determines the quality of the classification.

Figure 4 presents three loss function charts after training the model with the target learning rate of below 0.02 and reached more than 90,000 steps. The charts show a significant reduction of loss throughout training. It also shows the quality of the model accuracy is improving over the training steps.

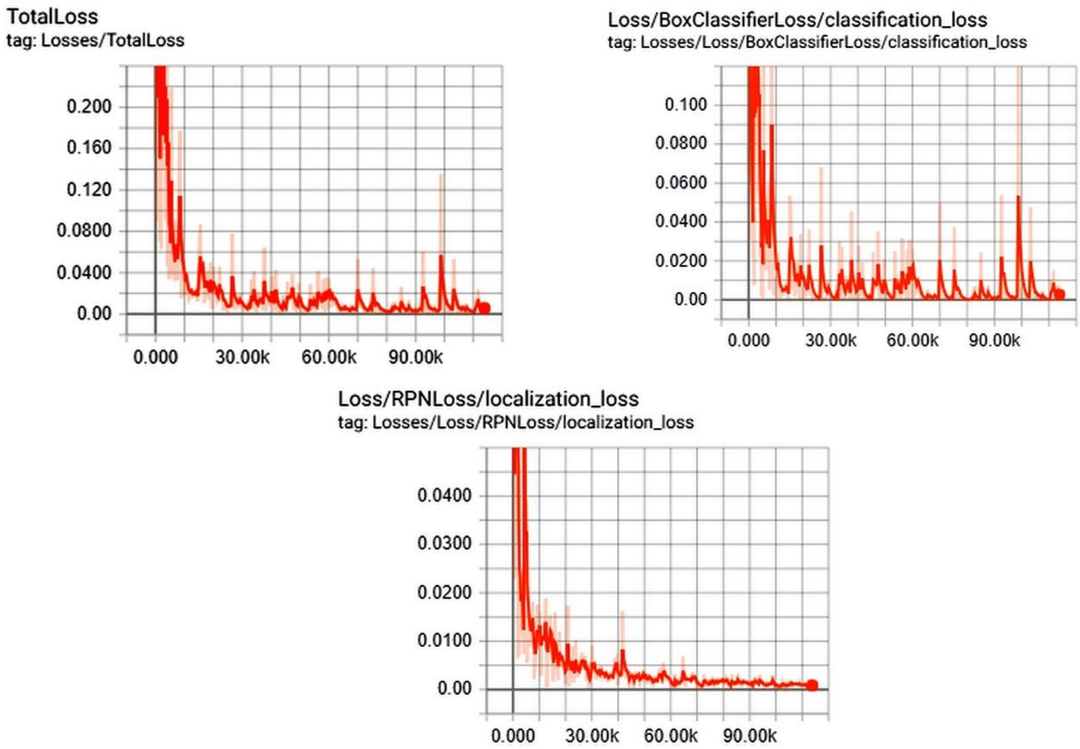

Figure 4. The charts represent the loss functions captured from Tensorboard after training the model below 0.02 learning rate and the training has reached more than 90,000 of training steps.

4.2 Effectiveness

We also evaluated the application of chest x-ray image classifier through real-time detection. For this reason, we used live webcam to capture the x-ray image followed by deciding whether there is a finding or vice versa. In the case of finding, the image is marked with yellow box. Otherwise, the image is marked with green box. In addition, the percentage of the model prediction will be shown too. Figure 5 illustrates a sample of detection outcome with the accuracy of the prediction.

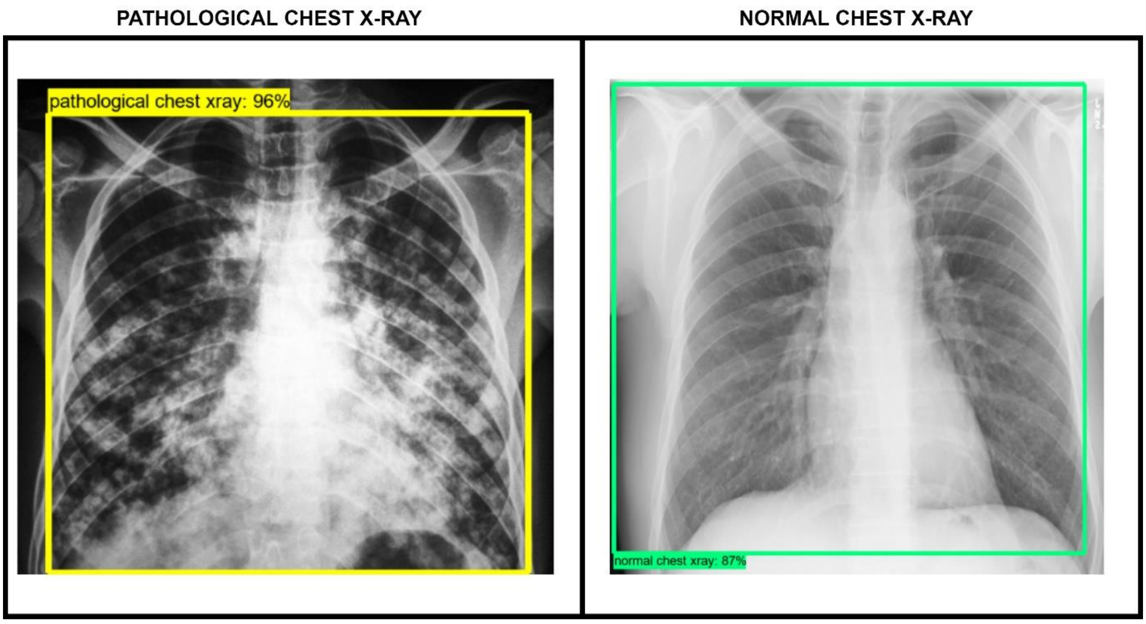

Figure 5. The images show the outcome of the classification. The left image is an x-ray image with finding, whilst the right is a normal x-ray image.

4.3 Performance

Table 2 shows the performance of the chosen model as compared to medical student *MS* and general practitioner *GP*. The results are also visualized using the radar chart as in Figure 6.

Table 2. Performance Results of *Model*, *GP* and *MS*

Parameter	Model	GP	MS
Accuracy	62%	50%	56%
Sensitivity	72.09%	93.02%	83.72%
Specificity	54.39%	17.54%	35.09%
Precision	54.39%	45.98%	49.32%

From the table, the model has the highest accuracy, specificity and positive prediction as compared to *MS* and *GP*. In terms of accuracy, *MS* is slightly better than *GP*. However, in the case of sensitivity, *GP* has higher score than the model and *MS*. This result may be caused by the cautious examination by *GP*. Typically, if *GP* has any doubt in examining chest x-ray, he/she will label the chest x-ray as pathological and ask a radiologist to get a second opinion.

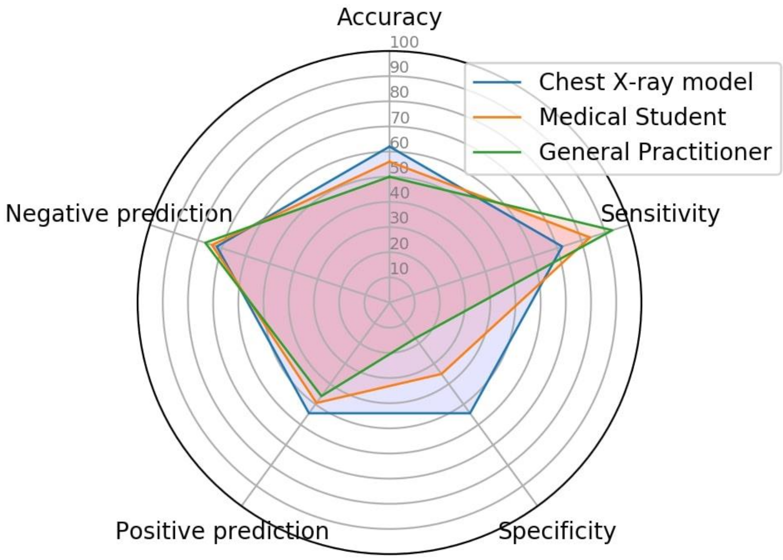

Figure 6. Performance results of *model*, *GP* and *MS*

4.4 Time Comparison

We also captured the time taken to obtain the classification outcome between model and medical representatives, as shown in Figure 7.

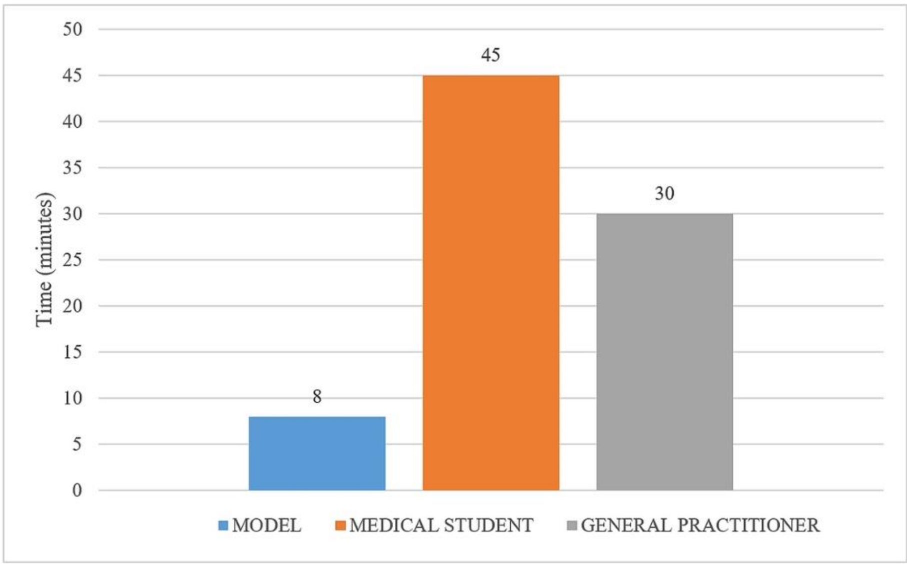

Figure 7. Time taken to classify 100 chest x-ray images into finding or normal

The figure shows that the model took an average 4.8sec per image, while *MS* required 27sec and *GP* required 18sec per image. Thus, in average, the model is faster in generating the results.

5. Related Work

In this section, we highlight a few related works that focused on chest x-ray image classification. The detailed review can be referred to previous surveys (Rahmat, Ismail, & Aliman, 2018). Herein, we briefly mention the related works that applied deep learning techniques for chest x-ray image classification and localization, and conclude to the unique contribution of this work.

A few works have applied deep learning techniques for chest X-ray images using different datasets. The work by Shin et al. (Shin et al., 2016) used CNN with Recurrent Neural Network (RNN) to classify and annotate diseases from chest x-ray images by utilizing OpenI dataset. Then, Bar et al. (Bar et al., 2016) made use of Sheba Medical Center dataset applied CNN (Donahue et al., 2014) with Kruskal–Wallis feature selection (Hollander & Wolfe, 1999). Meanwhile, Islam et al. (Islam, Aowal, Minhaz, & Ashraf, 2017) explored DCNN to localize features on chest x-ray images by utilizing several datasets from Indiana, JSRT and Shenzhen datasets.

There are several works (X. Wang et al., 2017b) ,(Rajpurkar et al., 2017b) that utilized similar dataset, namely, ChestX-ray8 dataset was released by Wang et al. (X. Wang et al., 2017b) and later, it is known as ChestX-ray14. In (X. Wang et al., 2017b), they applied Deep CNN to classify those images and predict the related eight diseases. It is then followed by several works, namely Rajpurkar et al. (Rajpurkar et al., 2017b) proposed CheXNet's algorithm consists of 121 layers of CNN, and Yao et al. (Yao et al., 2017) that proposed 2D ConvNet that combined a densely connected image encoder with a recurrent neural network model decoder. Furthermore, Guan et al. (Guan et al., 2018) proposed a three-branch Attention Guided CNN which learns from a global CNN branch, Gundel et al. (Gundel et al., 2018) proposed a Location Aware Dense Networks to classify pathologies in chest X-ray images, and Sedai et al. (Sedai et al., 2018) that proposed a weakly supervised method based on CNN which can predict the location of chest pathologies in x-ray images. Meanwhile, Abiyev et al. (Abiyev & Ma'aitah, 2018) presented some comparison between Backpropagation neural network, competitive neural network and CNN on classifying chest x-ray images.

A few enhanced models of CNN based on region proposals has been proposed to address localization problem. In particular, Girshick et al. (Girshick et al., 2014) proposed Regional-CNN to improve the precision of localization. Meanwhile, Wang et al. (L. Wang et al., 2017) they addressed guide-wire detection based on real-time approach with CNN. Furthermore, Regional Proposal Network is used to extract guide-wires features. R-CNN is then been enhanced with Fast R-CNN (Girshick, 2015) to address the performance issue.

It is then followed by Faster R-CNN (Ren et al., 2017) to further improve its performance. The work by Ruhan Sa et al. (Sa et al., 2017) has applied Faster R-CNN for detecting landmark points in lateral lumbar x-ray images. They demonstrated that a higher accuracy can be achieved even with a small dataset of x-ray images by tuning the deep learning network on natural images. Our work follows this direction by focusing on the chest x-ray images for addressing the classification.

6. Conclusion and Future Work

Region proposal network has been introduced into CNN to improve its precision in the object detection. It was then improved in terms of time execution by assembling multiple layers into a single network (i.e. Fast R-CNN) followed by removing the region proposal extraction at the earlier stage (i.e. Faster R-CNN). Hence, we have taken the initiative to apply Faster R-CNN for the chest x-ray image classification. To train and test the model, we have used a subset of ChestX-ray14 dataset and partition them into 80% for training set and 20% for testing set. We have validated the model by comparing its performance with medical representatives. The results show a reasonable accuracy (i.e. 62%) as compared to the medical representatives.

There are a few aspects that can be tackled in the future to improve the model performance. Firstly, the training dataset can be increased. For this reason, an alternative to manual labelling is needed to label the region. Secondly, we have simplified the categories into two, finding and normal. Thus, future work can consider the classification of 14 categories of diseases using Faster R-CNN model. Thirdly, the model validation can be compared with more medical representatives that may include radiologists.

References

Abiyev, R. H., & Ma'aitah, M. K. S. (2018). Deep Convolutional Neural Networks for Chest Diseases Detection. *Journal of Healthcare Engineering*, *2018*, 1–11. https://doi.org/10.1155/2018/4168538

Agarwal, N., Balasubramanian, V. N., & Jawahar, C. V. (2018). Improving multiclass classification by deep networks using DAGSVM and Triplet Loss. *Pattern Recognition Letters*, *112*, 184–190. https://doi.org/10.1016/J.PATREC.2018.06.034

Ajay Mittal, Hooda, R., & Sofat, S. (2017). Lung field segmentation in chest radiographs: a historical review, current status, and expectations from deep learning. *IET Image Processing*, *11*(11), 937–952. Retrieved from http://digital-library.theiet.org/content/journals/10.1049/iet-ipr.2016.0526

Bar, Y., Diamant, I., Wolf, L., Lieberman, S., Konen, E., & Greenspan, H. (2016). Chest pathology identification using deep feature selection with non-medical training. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, *6*(3), 259–263. https://doi.org/10.1080/21681163.2016.1138324

Chunhui Gu, Lim, J. J., Arbelaez, P., & Malik, J. (2009). Recognition using regions. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 1030–1037. https://doi.org/10.1109/CVPR.2009.5206727

Doi, K. (2007). Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized Medical Imaging and Graphics : The Official Journal of the Computerized Medical Imaging Society*, *31*(4–5), 198–211. https://doi.org/10.1016/j.compmedimag.2007.02.002

Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., & Darrell, T. (2014). DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. *Proceedings of the 31st International Conference on Machine Learning, PMLR*, 647–655. Retrieved from http://proceedings.mlr.press/v32/donahue14.html

Girshick, R. (2015). *Fast R-CNN*. Retrieved from http://arxiv.org/abs/1504.08083

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 580–587. https://doi.org/10.1109/CVPR.2014.81

Guan, Q., Huang, Y., Zhong, Z., Zheng, Z., Zheng, L., & Yang, Y. (2018). *Diagnose like a Radiologist: Attention Guided Convolutional Neural Network for Thorax Disease Classification*. Retrieved from http://arxiv.org/abs/1801.09927

Gundel, S., Grbic, S., Georgescu, B., Zhou, K., Ritschl, L., Meier, A., & Comaniciu, D. (2018). *Learning to recognize Abnormalities in Chest X-Rays with Location-Aware Dense Networks*. Retrieved from http://arxiv.org/abs/1803.04565

Hollander, M., & Wolfe, D. A. (1999). *Nonparametric statistical methods* (2nd ed.). Wiley.

Islam, M. T., Aowal, M. A., Minhaz, A. T., & Ashraf, K. (2017). *Abnormality Detection and Localization in Chest X-Rays using Deep Convolutional Neural Networks*. Retrieved from http://arxiv.org/abs/1705.09850

Jamil, M. Z. A., Mutalib, S., rahman, S. A., & Aziz, Z. A. (2018). Classification of Paddy Weed Leaf using Neuro-Fuzzy Methods. *MALAYSIAN JOURNAL OF COMPUTING*, *3*(1), 54–66. Retrieved from http://gids.mohe.gov.my/index.php/mjoc/article/view/4883

Ker, J., Wang, L., Rao, J., & Lim, T. (2018). Deep Learning Applications in Medical Image Analysis. *IEEE Access*, *6*, 9375–9389. https://doi.org/10.1109/ACCESS.2017.2788044

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, pp. 1097–1105. Retrieved from https://dl.acm.org/citation.cfm?id=2999257

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., … Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005

Rahmat, T., Ismail, A., & Aliman, S. (2018). Chest X-rays Image Classification in Medical Image Analysis. *Applied Medical Informatics*, *40*(3–4), 63–73. Retrieved from http://ami.info.umfcluj.ro/index.php/AMI/article/view/639

Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., … Ng, A. Y. (2017a). *CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning*. Retrieved from http://arxiv.org/abs/1711.05225

Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., … Ng, A. Y. (2017b). *CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning*. Retrieved from http://arxiv.org/abs/1711.05225

Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031

Sa, R., Owens, W., Wiegand, R., Studin, M., Capoferri, D., Barooha, K., … Chaudhary, V. (2017). Intervertebral disc detection in X-ray images using faster R-CNN. *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 564–567. https://doi.org/10.1109/EMBC.2017.8036887

Sedai, S., Mahapatra, D., Ge, Z., Chakravorty, R., & Garnavi, R. (2018). *Deep Multiscale Convolutional Feature Learning for Weakly Supervised Localization of Chest Pathologies in X-ray Images*. https://doi.org/10.1007/978-3-030-00919-9_31

Shin, H.-C., Roberts, K., Lu, L., Demner-Fushman, D., Yao, J., & Summers, R. M. (2016). Learning to Read Chest X-Rays: Recurrent Neural Cascade Model for Automated Image Annotation. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2497–2506. https://doi.org/10.1109/CVPR.2016.274

Tensorflow detection model zoo - COCO-trained models. (2017). Retrieved from https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/detection_m odel_zoo.md

Uijlings, J. R. R., van de Sande, K. E. A., Gevers, T., & Smeulders, A. W. M. (2013). Selective Search for Object Recognition. *International Journal of Computer Vision*, *104*(2), 154–171. https://doi.org/10.1007/s11263-013-0620-5

Wang, L., Xie, X.-L., Bian, G.-B., Hou, Z.-G., Cheng, X.-R., & Prasong, P. (2017). Guide-wire detection using region proposal network for X-ray image-guided navigation. *2017 International Joint Conference on Neural Networks (IJCNN)*, 3169–3175. https://doi.org/10.1109/IJCNN.2017.7966251

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017a). *ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases*. Retrieved from http://arxiv.org/abs/1705.02315

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017b). *ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases*. Retrieved from http://arxiv.org/abs/1705.02315

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, *9*(4), 611–629. https://doi.org/10.1007/s13244-018-0639-9

Yao, L., Poblenz, E., Dagunts, D., Covington, B., Bernard, D., & Lyman, K. (2017). Learning to diagnose from scratch by exploiting dependencies among labels. *Eprint ArXiv:1710.10501*. Retrieved from http://arxiv.org/abs/1710.10501