Faster R-CNN and YoloV3 Based Landmark Detection Of Cephalometric Landmarks

L.C Tabata
3678013@myuwc.ac.za
Computer Science Department
University of the Western Cape
Bellville, Western Cape

ABSTRACT

The objective of this study is to compare the results obtained between two deep-learning algorithms in identifying cephalometric landmarks on 2D cephalograms (X-rays) using an automated identification system, Artificial Intelligence (AI). The comparison is performed on two deep-neural networks a proposal algorithm Faster Region-based Convolutional Neural Network (Faster R-CNN) and a regression-based algorithm You Only Look Once version 3 (Yolov3). Both algorithms are trained and tested on a dataset provided by the University of the Western Cape, Department of Dentistry and a dataset presented in the IEEE International Symposium on Biomedical Imaging Challenge (ISBI-2015) combined. The detection was based on their performances, in terms of accuracy under the accepted accuracy range of 2 mm and the detection rate for clinical use. Null hypothesis was that Faster R-CNN will have a high precision rate and Yolov3 will have a higher succession rate.

KEYWORDS

Artificial Intelligence, Convolutional Neural Network, Yolov3, Faster R-CNN

1 INTRODUCTION

Cephalometric landmarks are tracings of anatomical features of a skull soft and hard tissues used by orthodontists for diagnosis, physicians to detect pathologies and oral maxillofacial treatment planning [1], which includes 32 soft tissue landmarks and 48 skeletal landmarks, see figure 1. Before application of AI in landmark detection, qualifying dentistry practitioners had to identify the landmarks manually through tracings. The manual process was not efficient it was time-consuming, the obtained results were biased and lead to observation and analysis errors. The application of deep-learning algorithms in the medical field in imaging is emerging. Researchers claim that Convolutional Neural Network (CNN) demonstrates as a cost-effective, labour-saving and powerful method in cephalometric metric landmark identification [2]. An increase in computational capacity and development of new advanced CNN algorithms improved reliability, accuracy and efficiency [3]. Present studies proposed frameworks of deep-learning algorithms in CNN namely Faster R-CNN and Yolov3. Faster R-CNN is based on Region Proposal Network (RPN) single network that is a combination of region proposal mechanism and CNN classifier. Faster R-CNN relies on powerful GPU computing power [4] and has a high detecting precision rate but lower computing speed. Yolov3 is not based on RPN, however predicts the bounding-boxes without generating object proposals. Yolov3 has a lower detecting precision rate but faster computing speed due to the network structure. However, previous studies in landmark detection [5] propose that an increase in dataset might improve the performances of the algorithms, Faster R-CNN and Yolov3. The aim of this study is to compare the results obtained between two deep-neural networks Faster R-CNN and Yolov3 in identifying 80 cephalometric landmarks based on performance categorized into accuracy under the clinically accepted range 2mm and detection rate.

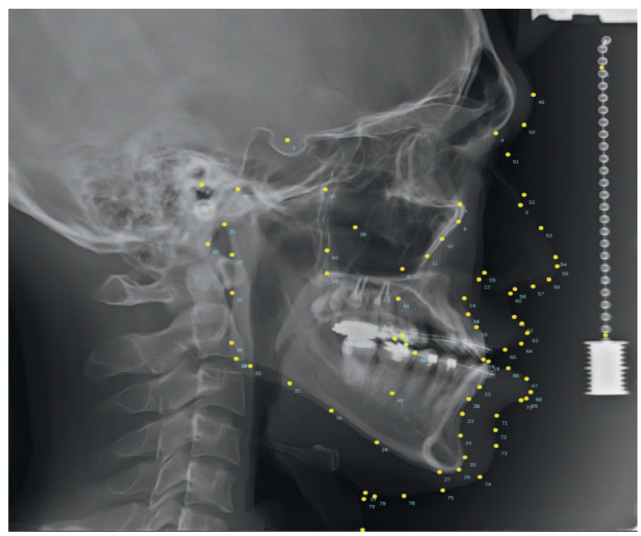

Figure 1: Image of a skeleton with identified 80 landmarks, taken from an article by Park et al. [3].

2 LITERATURE REVIEW

Recent studies, believe errors AI encounters in the process of detecting the landmarks are the same as of a human which maybe caused by poor image quality, unclear images or bones overlapping [6, 7] . Use of AI suggests that deep-learning is divided into two branches:

Two-stage object detection method: Li et al. [8] elaborates more on how R-CNN was improved. It began with R-CNN, region based CNN detector, which extracted features on proposed regions using CNN, and feed the output to SVM classifiers respectively and to the regressor to tighten an objects bounding-box. Because R-CNN was too slow, Fast R-CNN enhanced the rate by extracting features before proposing regions and replacing SVM classifier algorithm with the softmax layer, expanding the networks predictions. However the slow production of region proposal was improved by the RPN combined with Fast R-CNN algorithm in Faster R-CNN [9]. Faster R-CNN had a higher precision rate but slow computing speed, because of low computing speed it cannot be classed as a real-time algorithm.

One-stage object detection method: According to Kathuria [10], it begins with Yolo, which generated multi bounding-boxes and class probabilities at the same time, but new models outperform it in terms of accuracy. Yolov2 introduced Darknet-19, 19 layered network, omitting the fully-connected layer, to improve accuracy, but struggled to detect small features. Yolov3 introduced Darknet-53, 53 layered network, increasing accuracy but slower than Yolov2 [11]. However, Yolov3 still had a low precision rate especially targeting smaller features and high computing speed due to its structure [3]. This method also includes SSD, but Xu [7] concluded it was outperformed by Yolov3 in terms of speed and accuracy in detecting cephalometric landmarks.

Recent studies revealed that an improved Faster R-CNN method, CephaNet is 6 % more accurate than other deep learning methods including Faster R-CNN in cephalometric landmark detection [2] and a research based on landmark detection revealed that the Random Forest method identified 19 landmarks in a small span of time [12]. Reinders et al. [13] concluded that the Machine Learning good enough landmark detection algorithm Random Forest-Voting [3] and Faster R-CNN produced adequate results and they had an ability to detect in-depth even smaller features even with small datasets. Shadab et al. [5] concluded both algorithms , Faster R-CNN and Yolov3 can successfully detect features on objects and suggested a large dataset would improve the results. This study will assist in the research of obtaining a good enough cephalometric landmark detection model to simplify the dentistry life saving time and producing results are not biased. The aim of this study is to compare the results obtained between two deep learning neural networks Faster R-CNN and Yolov3 in detecting 80 cephalometric landmarks.

3 RESEARCH FOCUS

Cephalometric landmark detection is an analysis of tracings on cephalograms, utilized in ortho medicine and related fields. Since automated landmark identification algorithms and methods were introduced within the 1980's [14], several deep-learning methods are suggested and approved for clinical use like advanced Faster R-CNN and Yolov3.

From the literature review, recent studies are using both machine learning and deep-learning algorithms so detect the landmarks on cephalograms. However, it has been found that Faster R-CNN is more accurate in landmark detection than all the methods within the same framework. Faster R-CNN uses CNN to extract feature maps from the image once off, the RPN generates predictions from the maps. Region of Interest (RoI) pooling is applied to extract features that correspond to the landmarks, then that information is assessed and classified and also the bounding boxes are preserved and attached to the image boundary, see figure 2. Park et al. [3] concluded that Yolov3 is better than SSD when it comes to accuracy and detection rate making it the better ones in the framework. Yolov3 uses base feature extractor Darknet-53 to extract features from the image, CNN layers added to the extractor predicts the bounding boxes and class predictions, see figure 3. Faster R-CNN has a high precision rate and Yolov3 has high computing speed. Qian et al. [2] in the results showed that they have compared the CephaNet method with Faster R-CNN and the accuracy difference was 6 %, not a big difference. From the surveyed literature review it is evident that Faster R-CNN is more accurate than Yolov3 and Yolov3 has a higher computation speed than Faster R-CNN in landmark detection. Reinders et al. [13] suggested that Faster R-CNN and Random Forest Voting algorithm produced adequate results.

Thus in this study, we compare both Faster R-CNN and Yolov3 because it has not been done in the literature and the results we obtain can be used to find a method that is good enough to detect cephalometric landmarks on cephalograms. We will learn and test the dataset of cephalograms on the models and weight the results.

4 PROPOSED APPROACH

4.1 Dataset

In this study we'll be employing a dataset of 100 cephalometric X-ray images obtained from the University of the Western Cape, Dentistry department combined with the dataset of 400 cephalometric X-ray images from the IEEE International Symposium on Biomedical Imaging Challenge (ISBI-2015) [**?**]ollected from 400 patients, totaling to a dataset of 500 images. The landmarks on the dataset were manually identified by qualified orthodontists and use of a landmark identification software V-Ceph. The dataset was split into 80/20, were 80 % of the cephalometric images (400) are used for learning and 20 % of the cephalometric images (100) are used for testing.

4.2 Systems

The two systems one based on Faster R-CNN [16] and the other based on Yolov3 [17], were coded in Python on Ubuntu 18.04.4 LTS running server. Learning data of of 400 images for the learning algorithms. 80 landmarks manually recorded served as inputs in the learning processes.

The images were resized from the original sizes to the targeted size of 608 x 608 pixels for best deep-learning. During the learning processes each cephalogram image is processed through CNN with its corresponding landmark label for both algorithms, Faster R-CNN and Yolov3.

4.3 Procedures

To test for detection rates and accuracies between the systems, the 20 % testing dataset is used. The accuracies obtained are recorded as point-to-point errors by calculating the square distance values between the corresponding identified landmark and the ground truth positions. We present the obtained errors with 95 % confidence ellipses based on chi-square distribution [18–20] and a two-dimensional scatter-gram for each landmark. The detection rates were reported as the average processing speeds to identify the 80 cephalometric landmarks on a cephalogram image and for accuracy, the results obtained were compared with those obtained and used

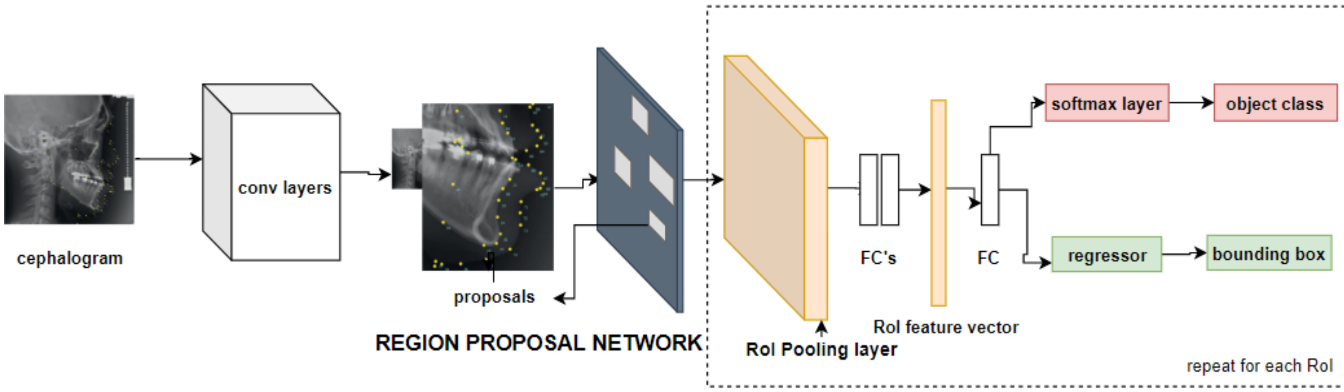

Figure 2: Faster R-CNN process.

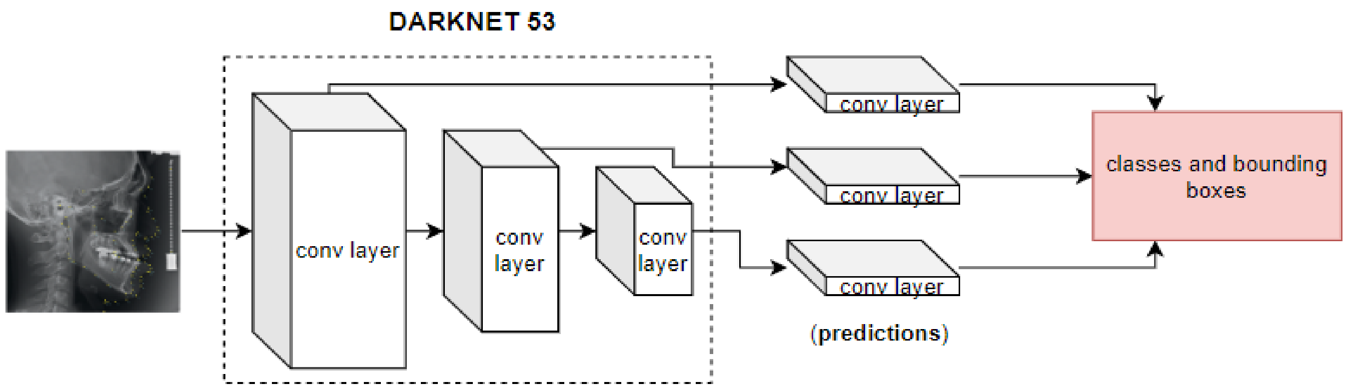

Figure 3: Yolov3 precess.

in the literature in identifying 19 cephalometric landmarks, (2, 2.5, 3, 4) mm ranges [12]. Statistical analysis were obtained using the R language. We applied a t-test with the probability of 0.05 with correction of alpha errors Bonferroni to compare the difference in the errors obtained from testing data between the algorithms Faster R-CNN and Yolov3.

REFERENCES

[1] Calderon-Sastre Enrique, Lopez-Ramirez Jose Luis, Beltran-Hernandez Jose Guadalupe, Ibarra-Manzano Mario Alberto, and Ruiz-Pinales Jose. Automatic estimation of cephalometric landmarks for lateral skull x-rays using cnn. 2019.

[2] Jiahong Qian, Ming Cheng, Yubo Tao, Jun Lin, and Hai Lin. CephaNet: An improved faster r-cnn for cephalometric landmark detection. *Proceedings - International Symposium on Biomedical Imaging*, 2019-April(Isbi):868–871, 2019. ISSN 19458452. doi: 10.1109/ISBI.2019.8759437.

[3] Ji-Hoon Park, Hye-Won Hwang, Jun-Ho Moon, Youngsung Yu, Hansuk Kim, Soo-Bok Her, Girish Srinivasan, Mohammed Noori A Aljanabi, Richard E Donatelli, and Shin-Jae Lee. Automated identification of cephalometric landmarks: Part 1—comparisons between the latest deep-learning methods yolov3 and ssd. *The Angle Orthodontist*, 89(6):903–909, 2019.

[4] Qi-Chao Mao, Hong-Mei Sun, Yan-Bo Liu, and Rui-Sheng Jia. Mini-yolov3: Real-time object detector for embedded applications. *IEEE Access*, 7:133529–133538, 2019.

[5] Malik Hashmat Shadab, Mahavir Dwivedi, SN Omkar, Tahir Javed, Abdul Bakey, Mohammad Raqib Pala, and Akshay Chakravarthy. Disease recognition in sugarcane crop using deep learning.

[6] Hye-Won Hwang, Ji-Hoon Park, Jun-Ho Moon, Youngsung Yu, Hansuk Kim, Soo-Bok Her, Girish Srinivasan, Mohammed Noori A Aljanabi, Richard E Donatelli, and Shin-Jae Lee. Automated identification of cephalometric landmarks: Part 2-might it be better than human? *The Angle Orthodontist*, 90(1):69–76, 2020.

[7] Joyce Xu. Deep Learning for Object Detection: A Comprehensive Review. https://towardsdatascience.com/deep-learning-for-object-detection-a-comprehensive-review-73930816d8d9, 2017. Accessed February 22, 2020.

[8] Kaidong Li, Wenchi Ma, Usman Sajid, Yuanwei Wu, and Guanghui Wang. Object detection with convolutional neural networks. *arXiv preprint arXiv:1912.01844*, 2019.

[9] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1137–1149, 2017. ISSN 01628828. doi: 10.1109/TPAMI.2016.2577031.

[10] Ayoosh Kathuria. What's new in YOLO v3? https://towardsdatascience.com/yolo-v3-object-detection-53fb7d3bfe6b, 2018. Accessed Feb 29.

[11] Wanzeng Kong, Jichen Hong, Mingyang Jia, Jinliang Yao, Weihua Cong, Hua Hu, and Haigang Zhang. Yolov3-dpfin: A dual-path feature fusion neural network for robust real-time sonar target detection. *IEEE Sensors Journal*, 2019.

[12] Ching-Wei Wang, Cheng-Ta Huang, Jia-Hong Lee, Chung-Hsing Li, Sheng-Wei Chang, Ming-Jhih Siao, Tat-Ming Lai, Bulat Ibragimov, Tomaž Vrtovec, Olaf Ronneberger, Philipp Fischer, Tim Cootes, and Claudia Lindner. A benchmark for comparison of dental radiography analysis algorithms. *Medical Image Analysis*, 31, 02 2016. doi: 10.1016/j.media.2016.02.004.

[13] Christoph Reinders, Hanno Ackermann, Michael Ying Yang, and Bodo Rosenhahn. Learning convolutional neural networks for object detection with very little training data. In *Multimodal Scene Understanding*, pages 65–100. Elsevier, 2019.

[14] AM Cohen, HH-S Ip, and AD Linney. A preliminary study of computer recognition and identification of skeletal landmarks as a new method of cephalometric analysis. *British journal of orthodontics*, 11(3):143–154, 1984.

[15] Ching-Wei Wang, Cheng-Ta Huang, Jia-Hong Lee, Chung-Hsing Li, Sheng-Wei Chang, Ming-Jhih Siao, Tat-Ming Lai, Bulat Ibragimov, Tomaž Vrtovec, Olaf Ronneberger, et al. A benchmark for comparison of dental radiography analysis

algorithms. *Medical image analysis*, 31:63–76, 2016.
[16] Sercan Ö. Arik, Bulat Ibragimov, and Lei Xing. Fully automated quantitative cephalometry using convolutional neural networks. *Journal of Medical Imaging*, 4(1):014501, 2017. ISSN 2329-4302. doi: 10.1117/1.jmi.4.1.014501.
[17] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *CoRR*, abs/1804.02767, 2018. URL http://arxiv.org/abs/1804.02767.
[18] Richard E Donatelli and Shin-Jae Lee. How to report reliability in orthodontic research: Part 1, 2013.
[19] Richard E Donatelli and Shin-Jae Lee. How to test validity in orthodontic research: A mixed dentition analysis example, 2015.
[20] Richard E Donatelli and Shin-Jae Lee. How to report reliability in orthodontic research: Part 2, 2013.