The validity of an artificial intelligence application for assessment of orthodontic treatment need from clinical images

Sameh Talaat, Ahmed Kaboudan, Wael Talaat, Budi Kusnoto, Flavio Sanchez, Mohammed H. Elnagar, Christoph Bourauel, and Ahmed Ghoneima

Aim: **To assess the validity of a Convolutional Neural Network (CNN) digital model to detect and localize orthodontic malocclusions from intraoral clinical images. *Materials and methods:* The sample of this study consisted of the intraoral images of 700 Subjects. All images were intraoral clinical images, in one of the following views: Left Occlusion, Right Occlusion, Front Occlusion, Upper Occlusal, and Lower Occlusal. The following malocclusion conditions were localized: crowding, spacing, increased overjet, cross bite, open bite, deep bite. The images annotations were repeated by the same investigator (S.T) with a one week interval (ICC $\geq$ 0.9). The CNN model used for this research study was the "You Only Look Once" model. This model can detect and localize multiple objects or multiple instances of the same object in each image. It is a fully convolutional deep neural network; 24 convolutional layers followed by 2 fully connected layers. This model was implemented using the TensorFlow framework freely available from Google. *Results:* The created CNN model was able to detect and localize the malocclusions with an accuracy of 99.99%, precision of 99.79%, and a recall of 100%. *Conclusions:* The use of computational deep convolutional neural networks to identify and localize orthodontic problems from clinical images proved valid. The built AI engine accurately detected and localized malocclusion from different views of intra-oral clinical images. (Semin Orthod 2021; 27:164–171)**

Introduction

Malocclusion is the third most prevalent oral disease following dental caries and periodontal disease.[1] It affects periodontal health and increases the risk of dental caries, traumatic dental injuries, and temporomandibular joint problems.[2] In addition, it negatively

Assistant Professor of Orthodontics, Department of Orthodontics, Faculty of Dental Medicine, Future University, Egypt; Oral Technology, University of Bonn, Germany; DigiBrain4 Inc., Chicago, IL, United States; Associate Professor of Oral and Maxillofacial Surgery, Department of Oral and Craniofacial Health Sciences, College of Dental Medicine, University of Sharjah, United Arab Emirates; Assistant Professor of Oral Surgery, Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Suez Canal University, Egypt; Professor of Orthodontics, Department of Orthodontics, College of Dentistry, University of Illinois at Chicago, United States; Assistant Professor of Orthodontics, Department of Orthodontics, College of Dentistry, University of Illinois at Chicago, United States; Department of Orthodontics, Faculty of Dentistry, Tanta University, Egypt; Professor and Head of Oral Technology Department, University of Bonn, Germany; Professor, Chair and Graduate Program Director, Department of Orthodontics, H Hamdan Bin Mohammed College of Dental Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, United Arab Emirates; Adjunct Faculty, Department of Orthodontics, Faculty of Dental Medicine, Future University, Egypt.

Corresponding author at: Professor, Chair and Graduate Program Director, Department of Orthodontics, H Hamdan Bin Mohammed College of Dental Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, United Arab Emirates. E-mail: ahmed.ghoneima@mbru.ac.ae

1073-8746/12/1801-$30.00/0
https://doi.org/10.1053/j.sodo.2021.05.012

interferes with patients' quality of life, hinders their social interaction, and affects their psychological well-being.[3] Consequently, the Demand for orthodontic treatment is rising.[4] The purpose of orthodontic treatment is to create a healthy, functional occlusion, reduce decay gum disease, injury, promote oral health and general physical health. Orthodontic treatment also brings about an attractive smile and balanced facial pattern, which have emotional benefits related to self-confidence and self-esteem.[5] However, malocclusion varies from mild to moderate to severe; therefore, the level of treatment need varies widely.[6] Timing of orthodontic treatment varies as well depending on the type of malocclusion. Early diagnosis could intercept developing malocclusion, provides more favorable outcomes. More importantly, proper diagnosis, localization, and assessment of malocclusion are crucial for oral health.

Detection and assessment of malocclusion and orthodontic treatment need could be done by direct examination by a specialist or on the diagnostic cast.[7] Alternatively, studies showed that orthodontists could rely on digital images for assessing malocclusion.[8,9] However, provider shortages, geographic remoteness, busy lifestyle, lack of transportation, and cost are common barriers to access to care and screening for malocclusion.[10]

In this research study, the first fully automated system for localizing, detecting, assessing malocclusion and orthodontic treatment need is proposed. Using one of the rapidly developing deep learning techniques - convolutional neural networks (CNNs).[11] The major applications for CNNs are computer vision and recognition applications, image classification, and image segmentation.[12,13] Biologically inspired by humans' capability to glance at an image and instantly identify what objects are in the image, where they are, and how they interact. CNNs simulate the action of interconnected neurons in the human cortex to allow us to identify and locate objects in an image. Identification is the process of inferring the proper label describing the object, while localization is the process of deducing the actual bounding box, image coordinates, and surrounding (localizing) the detected object; it can be used to identify and count objects in a scene while accurately labeling them.[14]

In dentistry, deep learning techniques have been used for automatic cephalometric landmark detection and teeth segmentation in a panoramic x-ray.[15,16] The widespread use of intraoral photos as a screening tool for malocclusion facilitate communication between different dental specialties for assessment of patient treatment needs.[8] The literature lacks enough supportive information on accuracy of automated tools in detecting malocclusion and assessing treatment needs from photos or other diagnostic images.The purpose of this study was to develop and validate an AI application for automatic localization of malocclusion and assessing the treatment need using intraoral raw images.

Materials and methods

Ethical approval (IRP Number 20193360) for this study was obtained from the research ethics committee of WIRB-Copernicus. All experiments were carried out at the department of Oral Technology at Bonn University (Germany) in accordance with approved guidelines.

Dataset

Data collection

The primary dataset composed of intraoral images of 700 Subjects. All images were intraoral clinical patient images, in one of the following views: Left Occlusion, Right Occlusion, Front Occlusion, Upper Occlusal, and Lower Occlusal.

Data inspection and cleansing

Images weere reviewed for quality, accuracy, consistency, and uniformity. Images not conforming with the study protocol were excluded at this stage, and the resulting set was only 576 images (Fig. 1).

Training, validation, and test datasets

The data set was divided into three groups in the ratio 80:15:05. Training set, 80% (460) images, used for training. Validation set, 15% (86) images, used to validate (test) how well training is advancing. Test set, 5% (30) images, used to test the fully trained network.

Figure 1. Example of an excluded image because of being motion-blurred.

Annotating the dataset

Each image was annotated independently by an investigator and orthodontic specialist (S.T.). For every image, the malocclusion was identified and localized. The following malocclusion conditions were identified: crowding, spacing, increased overjet, crossbite, open bite, deep bite. The annotations were repeated by the same investigator with a one-week interval in between. There was a strong agreement between the two malocclusion annotations rounds taken with a one-week interval (ICC $\geq$ 0.9).

Each image in the dataset was manually annotated. The malocclusion of interest was identified and recorded with its unique label by a drawn box. The coordinates of the box surrounding that malocclusion was recorded. A free software application, Microsoft Vott,[17] was used to perform this step (Fig. 2).

Export of annotations

Pascal VOC format is a standard, human-readable, computer file format for recording the labels and the coordinates of the objects of interest. It is based on the XML format, and it is wildly used in computer software. A single file was generated for each image. Total number of annotation files = Total number of images = 576. Fig. 3 shows a sample of a single VOC file.

The neural network model

The model that was chosen to be used for this research study was the "You Only Look Once" model.[14] This model can detect and localize multiple objects or multiple instances of the same object in each image. This object detector uses features learned by a deep convolutional neural network to detect objects. It is a fully convolutional deep neural network, 24 convolutional layers followed by 2 fully connected layers. Alternating 1×1 convolutional layers reduce the

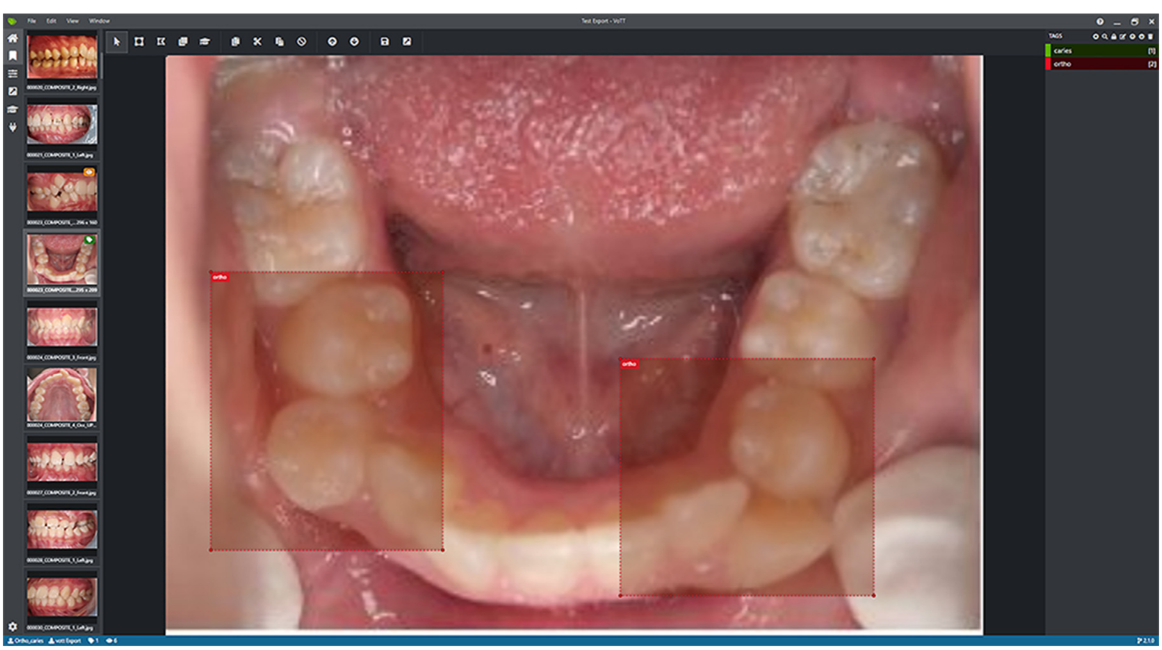

Figure 2. A snapshot during the manual annotation using VOTT.

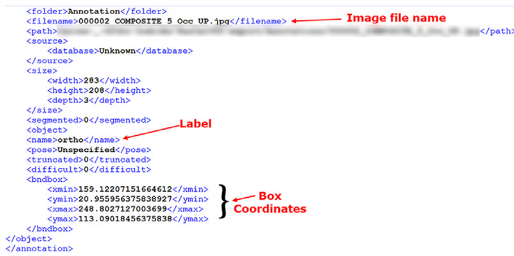

Figure 3. Export of annotations in pascal VOC format.

features space from preceding layers (US Patent no: 10,878,954 B2) (Fig. 4).

This model was implemented using the TensorFlow framework freely available from Google.[18] The object detection and localization model divides each image into an S x S grid. If the center of an object falls into a grid cell, that grid cell is responsible for detecting that object (Fig. 5). Each grid cell detects multiple bonding boxes around the object of interest, each box with its relative confidence number (Fig. 6).

Five different box sizes were used to get the best localization of the object. Confidence of multiple boxes participates in a single score in order to get the final coordinates of the best fitting boxes. This technique is named IoU (intersection over Union).Intersection over Union is an evaluation metric used to measure an object detector's accuracy on a particular dataset. In the numerator, we compute the area of overlap between the predicted bounding box and the ground-truth bounding box. The denominator is the Area of Union, or more simply, the area encompassed by both the predicted bounding box and the ground-truth bounding box. Dividing the area of overlap by the Area of Union yields our final score - the Intersection over Union (Figs. 7 and 8).

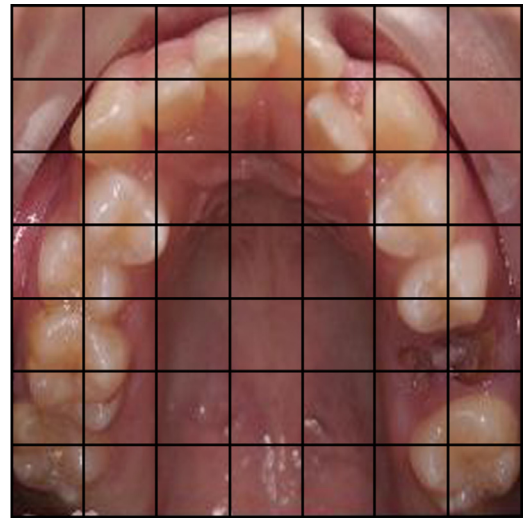

Figure 5. This model divides each image to an S x S grid.

Training

Training is a very time-consuming procedure for this type of deep networks. We used our in-house GPUs setup, based on 2 NVIDIA Volta GPUs, for a total of 24 GB memory, 10,240 CUDA Cores, and 1280 Tensor Cores. This setup can reach more than 200 teraflops (Tera=10^{12} floating-point operations per second). Images were all pre-scaled to 416×416 and were all transformed to jpg format. After many trials, the best learning rate was set to: 0.0001. Batch size was set to 32, a momentum of

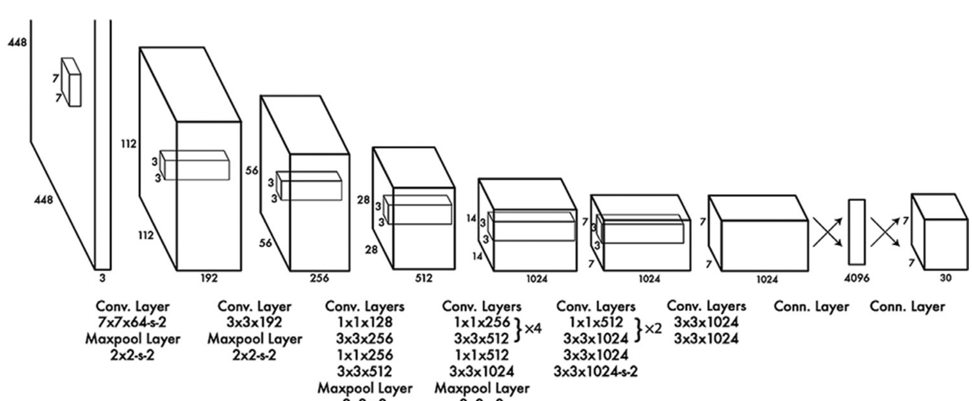

Figure 4. The neural network model.

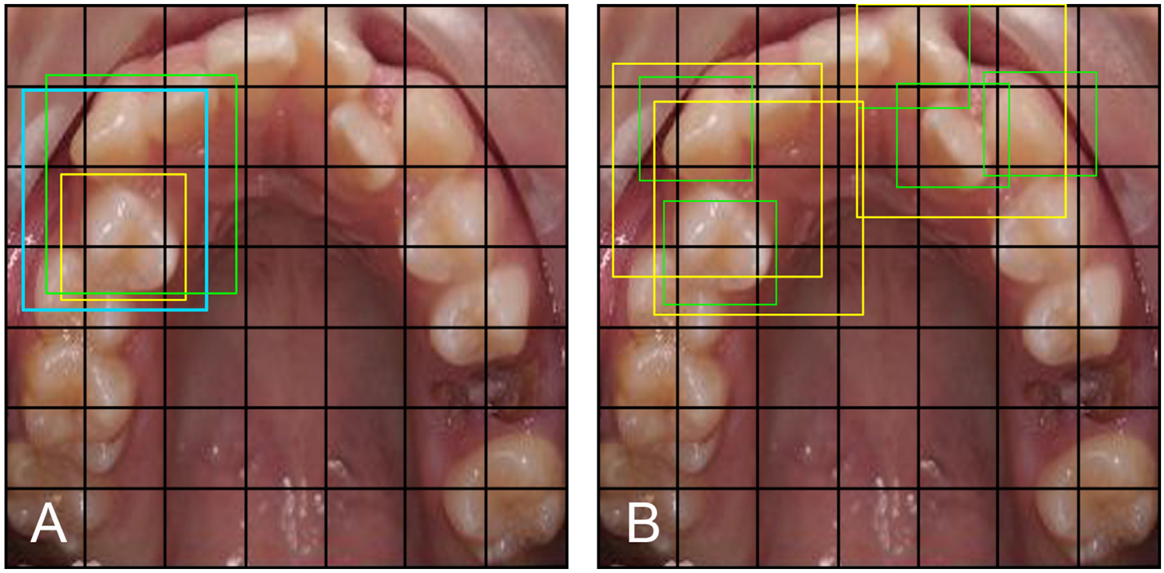

Figure 6. Grid cells detecting bonding boxes.

0.9 and a decay of 0.0005. The training was run for 300 epochs.

Results

Accuracy metrics

Considering the possible predictions, they can be arranged as follows:

True Positives	False Positives
False Negatives	True Negatives

A True Positive (TP) is an observation correctly predicted for a given label, while a False Positive (FP) is an observation incorrectly predicted for that label. Both True and False Negatives (FN and TN) are those predicted for a different label.

Accuracy is then defined as the sum of the number of TH and TN divided by the total number of examples (where # means 'number of', and TP stands for True Positive, etc.):[19]

$$\text{Accuracy} = \frac{TP + FP}{TP + FP + TN + FN}$$

Recall **is** the ratio of the number of correct positive examples out of those that were classified as positive, while Precision is the ratio of correct positive examples to the number of actual positive examples:

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

Combined together, both Accuracy and Recall measures give more information than just the accuracy. They can be combined to give a single measure, the F1 measure,[19] which can be written in terms of precision and recall as:

$$F_1 = 2\frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

$$F_1 = \frac{TP}{TP + (FN + FP)/2}$$

Based on the former formulas, results can be summarized as in Table 1. (Fig. 9)

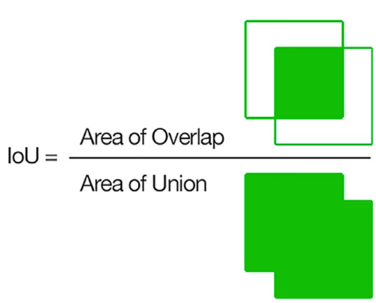

Figure 7. Calculating IoU metric.

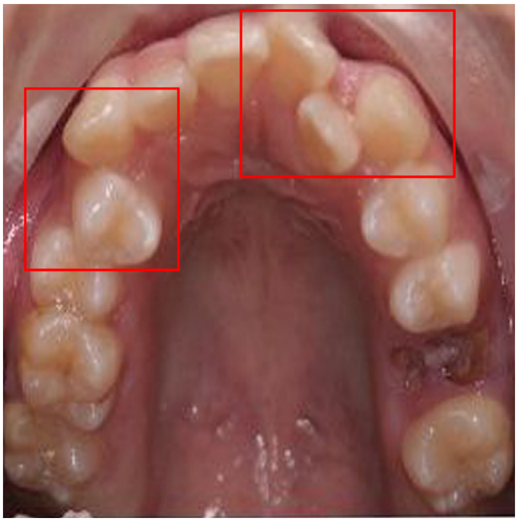

Figure 8. Joining the confidence of multiple boxes yielding the best fitting boxes.

Table 1. Summary of findings.

Accuracy	0.9992 (99.99%)
Precision	0.9979 (99.79%)
Recall	1.00
F1 score	1.00
True Positive	~ 98% and False Positive ~ 0.9%.

Discussion

The cognitive processes that underlie medical reasoning are complex; such processes can be affected by some amount of "irreducible uncertainty" that easily results in failure.[20] Rational medical thinking states that doctors must collect, weigh, and summarize all relevant information to arrive at a conclusive clinical judgment. Cliniciansdo so automatically by integrating clinical perceptions within the context of their medical knowledge base, involving the recognition of patterns and schemes that they have previously seen.[21] Practitioners must make decisions in an environment of uncertainty and under the constraint of the limited time available. Emergency physicians, for example, generate 75% of diagnostic hypotheses in the first 5 min of the clinical encounter,[22] in a domain with over 10,000 known medical diagnoses, and over 40 cognitive biases that may impact the diagnostic outcome.[20]

Machine learning is a subdomain of AI, which allows a computer to be trained to recognize patterns using large data sets. The use of ML in medicine is currently directed toward improving work efficiency, saving time, and achieving more accurate results using computational technologies. In the last two decades, computer-aided technologies have significantly evolved and are currently being used to assist clinicians and radiologists in diagnosing several diseases in different medical fields.[23] CNNs have been used in various medical applications like in detection of colitis on computed tomography and the detection of brain lesions on magnetic resonance images.[24,25] However, the applications in dentistry are still limited.[26] In the orthodontic specialty, the practitioner has to handle clinical information under the pressure of time, heavy workload, and insufficient knowledge. In the present study, the authors trained a CNNs model for a fully automated orthodontic assessment of intraoral raw clinical images.

In the current study, a CNN model was created, tested, and was able to detect and localize different types of malocclusions from variable intraoral clinical images. All images were intraoral clinical images, in one of the next views: Left Occlusion, Right Occlusion, Front Occlusion, Upper Occlusal, and Lower Occlusal. The following malocclusion conditions were localized: crowding, spacing, increased overjet, crossbite, open bite, deep bite. The annotations were repeated by the same investigator (S.T) with a one-week interval. There was a strong agreement between the 2 malocclusion annotations rounds taken with a one-week interval (ICC $\geq$ 0.9). The created CNN model was able to detect and localize the malocclusions with an accuracy of 99.99%, precision of 99.79% and a recall of 100%.

This opens a new window for home care automated assessment, done by the patient himself, by presenting images produced by his mobile, according to specific instructions, and immediately get a summary of the problems and a treatment difficulty hint. The current work focused on orthodontic problems, but is valid for almost all types of dental and oral pathosis, extending the horizon of automated dental care. This may also present a valuable tool for health/dental insurance companies, allowing for automated pre-authorization and pre/post-treatment evaluation.

Conclusion

The use of computational deep convolutional neural networks to identify and localize dental problems from clinical images proved valid. The built

Figure 9. Resulting testing images showing fully automated prediction boxes depicted in thin red marking the malocclusions.

AI engine was capable of detecting and localizing malocclusion from clinical images accurately.

References

1. Suliano AA, Rodrigues MJ, Caldas ADF, Da Fonte PP, Porto-Carreiro CDF. Prevalence of malocclusion and its association with functional alterations of the stomatognathic system in schoolchildren. *Cad Saude Publ*. 2007;23(8):1913–1923.
2. Anthony SN, Zimba K, Subramanian B. Impact of malocclusions on the oral health-related quality of life of early adolescents in Ndola, Zambia. *Int J Dent*. 2018. 2018.
3. Marques LS, Ramos-Jorge ML, Paiva SM, Pordeus IA. Malocclusion: esthetic impact and quality of life among Brazilian schoolchildren. *Am J Orthod Dentofac Orthop*. 2006;129(3):424–427.
4. Jawad Z, Bates C, Hodge T. Who needs orthodontic treatment? Who gets it? And who wants it? *Br Dent J*. 2015;218(3):99–103.
5. Importance of Orthodontics for Children Parents guide to orthodontics. Accessed September 30, 2020. https://www.aaoinfo.org/blog/parent-s-guide-post/importance-of-orthodontics/.
6. Doan AA, Sari E, Uskun E. Ahin Salam AM. comparison of orthodontic treatment need by professionals and parents with different socio-demographic characteristics. *Eur J Orthod*. 2010;32(6):672–676.
7. Abdul Rahim FS, Mohamed AM, Nor MM, Saub R. Malocclusion and orthodontic treatment need evaluated among subjects with down syndrome using the dental Aesthetic index (DAI). *Angle Orthod*. 2014;84(4):600–606.
8. Mok CW, Zhou L, McGrath C, Hägg U, Bendeus M. Digital images as an alternative to orthodontic casts in assessing malocclusion and orthodontic treatment need. *Acta Odontol Scand*. 2007;65(6):362–368.
9. Jackson TH, Kirk CJ, Phillips C, Koroluk LD. Diagnostic accuracy of intraoral photographic orthodontic records. *J Esthet Restor Dent*. 2019;31(1):64–71.
10. Morris RS, Hoye LN, Elnagar MH, et al. Accuracy of dental monitoring 3D digital dental models using

photograph and video mode. *Am J Orthod Dentofac Orthop.* 2019;156(3):420–428.

11. A.G. Howard, M. Zhu, B. Chen, et al. MobileNets: efficient convolutional neural networks for mobile vision applications.
12. Yang X, Wu N, Cheng G, et al. Automated segmentation of the parotid gland based on atlas registration and machine learning: a longitudinal mri study in head-and-neck radiation therapy. *Int J Radiat Oncol Biol Phys.* 2014;90(5):1225–1233.
13. O. Ronneberger, P. Fischer, T. Brox U-Net: convolutional networks for biomedical image segmentation. http://lmb.informatik.uni-freiburg.de/.
14. J. Redmon, S. Divvala, R. Girshick, A. Farhadi You only look once: unified, real-time object detection. Accessed October 2, 2020. http://pjreddie.com/yolo/.
15. Y. Song, X. Qiao, Y. Iwmoto, Y.W. Chen Semi-automatic cephalometric landmark detection on X-ray images using deep learning method. In:; 2020:585−592.
16. Wirtz A, Mirashi SG, Wesarg S. Automatic teeth segmentation in panoramic X-Ray images using a coupled shape model in combination with a neural network. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).* Springer Verlag; 2018:712–719. Vol 11073 LNCS.
17. GitHub - microsoft/VoTT: Visual object tagging tool: an electron app for building end to end object detection models from images and videos. Accessed October 2, 2020. https://github.com/microsoft/VoTT.
18. TensorFlow. Accessed October 4, 2020. https://www.tensorflow.org/.
19. Stephen M. *Machine Learning: An Algorithmic Perspective, Second Edition - Stephen Marsland - © 2015 By Taylor & Francis Group, LLC.* CRC Press Is an Imprint of Taylor & Francis Group, an Informa BusinessGoogle Books; 2020... Accessed October 4; https://books.google.com/books/about/Machine_Learning.html?id=6GvSBQAAQBAJ.
20. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med.* 2003;78:775–780.
21. Croskerry P. Achieving quality in clinical decision making: cognitive strategies and detection of bias. *Acad Emerg Med.* 2002;9:1184–1204.
22. Wegwarth O, GaissmaierW GG. Smart strategies for doctors and doctors-in-training: heuristics in medicine. *Med Educ.* 2009;43:721–728.
23. Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Comput Med Imaging Graph.* 2007;31(4−5):198–211.
24. Liu J, Wang D, Lu L, Wei Z, Kim L, Turkbey EB, et al. Detection and diagnosis of colitis on computed tomography using deep convolutional neural networks. *Med Phys.* 2017;44:4630–4642.
25. Hsieh YZ, Luo YC, Pan C, Su MC, Chen CJ, Hsieh KL. Cerebral small vessel disease biomarkers detection on MRI-sensor-based image and deep learning. *Sensors.* 2019;19(11):2573.
26. Talaat S, Kaboudan A, Talaat W, et al. Improving the accuracy of publicly available search engines in recognizing and classifying dental visual assets using convolutional neural networks. *Int J Comput Dent.* 2020;23(3):211–218.