Artificial Intelligence in Detecting Temporomandibular Joint Osteoarthritis on Orthopantomogram

Eunhye Choi

School of Dentistry and Dental Research Institute, Seoul National University

Donghyun Kim

Yonsei University College of Dentistry

Jeong-Yun Lee

Seoul Cheongchoon Dental Clinic

Hee-Kyung Park (✉ dentopark@snu.ac.kr)

School of Dentistry and Dental Research Institute, Seoul National University

Research Article

Keywords: temporomandibular osteoarthritis, orthopantomogram, artificial intelligence, deep learning

Posted Date: December 29th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-133347/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Orthopantomogram (OPG) is still important in primary diagnosis of temporomandibular joint osteoarthritis (TMJOA), because of cost and radiation of computed tomogram (CT). The aims of this study were to develop an artificial intelligence (AI) model and compare its TMJOA diagnostic performance on OPG with that of an oromaxillofacial radiology (OMFR) expert. An AI model was developed using Karas' ResNet model and trained to classify images into three categories: normal, indeterminate OA, and OA. This study included 1189 OPG images confirmed by cone-beam CT and evaluated the results by model (accuracy, precision, recall, and F1 score) and diagnostic performance (accuracy, sensitivity, and specificity). The model performance was unsatisfying when AI developed with 3 categories. After the indeterminate OA images were reclassified as normal, OA, or omission, the AI diagnosed TMJOA in a similar manner to the expert and was most accordant with CBCT read when the indeterminate OA was omitted (accuracy: 0.78, sensitivity: 0.73, and specificity: 0.82). Our deep learning model showed a sensitivity equivalent to that of an expert, with a better balance between sensitivity and specificity, which implies that AI can play an important role in primary diagnosis of TMJOA on OPGs in most general practice clinics where OMFR experts or CT are not available.

Introduction

Temporomandibular joint osteoarthritis (TMJOA) is an important subtype of temporomandibular disorders (TMDs) and may lead to substantial joint pain, dysfunction, dental malocclusion and reduced health-related quality of life¹. Osteoarthritis (OA) is a degenerative disease of joints caused by a series of degenerative processes including gradual loss of joint cartilage, remodeling and hardening of subchondral bone, and formation of osteoproliferative bodies². TMJOA is confirmed by structural bony changes observed on radiographic examination. Because of the ability to show the minute change of integrity of the cortical and subcortical bone of the TMJ, it is widely accepted that computed tomography (CT) is the reference standard to diagnose TMJOA³. Cone beam CT (CBCT), which has benefit of lower radiation exposure than conventional CT, is reportedly as accurate for detecting TMJOA⁴. However, CBCT is not the first of choice for TMJOA examination in normal clinical apparatus yet because it still has higher radiation dose and cost than plain radiographs. Normally, orthopantomogram (OPG) is the most common examination for screening various lesions and conditions in the maxillofacial region, while it is difficult to identify bony change of the TMJ structure small in size and overlapped by other skull structures⁵. It makes OPG just used as screening examination in a way that the experienced expert such as oromaxillofacial radiology (OMFR) or orofacial pain specialist reads and then recommends, if necessary, an additional CBCT as confirmative examination. In a situation where such experts are not available, the patient's TMJOA can be overlooked or misread. On this point an AI was developed and got to learn how to read TMJOA on OPGs based on CBCT results already confirmed by experts. Various studies have been reported to apply AI algorithms to read OPG for clinical conditions such as tooth segmentation⁶, age estimation⁷, third molar and inferior alveolar nerve detection⁸, cyst and tumors of jaw⁹, osteoporosis¹⁰, and maxillary sinusitis¹¹. However, there is few AI studies on TMJOA diagnosis, and a study using CBCT and a study based on OPG have been reported^{12,13}.

This study aimed to investigate the clinical usability of an AI diagnostic tool developed for TMJOA diagnosis on OPG in terms of deep learning by comparing its read with that of an expert.

Results

The results of AI developed based on 3 categories, normal, indeterminate, OA, is shown in Table 2. Because overall accuracy was not satisfying, model development process was modified and reevaluated. On the point that indeterminate TMJOA could've compromised AI training because of its vagueness on radiographic reading^{14,15}. AI was trained, validated, and tested again in 3 ways for modification. Indeterminate OA was considered as normal in trial 1, OA in trial 2,

and omitted from whole development process in trial 3. The accuracy of the model performance was best in Trial 1 (0.80) followed by Trial 3 (0.78) and Trial 2 (0.74), but precision and recall were evenest in Trial 3 (Table 3).

Table 1
Clinical and demographic characteristics of the OPG dataset.

		Normal		Indeterminate		OA		Total	
Numbers of joints	Female	800	619	779	656	799	683	2378	1958
	Male		181		123		116		420
Mean age (±SD)		34.06 (±14.44)		33.11 (±14.29)		40.23 (±16.90)		35.82 (±15.58)	
Abbreviations: Indeterminate, indeterminate temporomandibular joint osteoarthritis; OA, temporomandibular joint osteoarthritis.									

Table 2
Confusion Matrix and model performance of AI developed initially.

Confusion Matrix				Model Performance					
	Predicted			Precision	Recall	Accuracy	Weighted average precision	Weighted average recall	F1 score
Actual	Normal	Indeterminate	OA						
Normal	57	46	47	0.72	0.38	0.51	0.55	0.51	0.53
Indeterminate	14	53	83	0.44	0.35				
OA	8	22	120	0.48	0.80				
Abbreviations: Indeterminate, indeterminate temporomandibular joint osteoarthritis; OA, temporomandibular joint osteoarthritis.									

Table 3
Confusion matrix and model performance in each Trial.

Confusion Matrix				Model Performance					
		Predicted		Precision	Recall	Accuracy	Weighted average precision	Weighted average recall	F1 score
	Actual	Normal	OA						
Trial 1	Normal	283	17	0.80	0.94	0.80	0.81	0.80	0.80
	OA	72	78	0.82	0.52				
Trial 2	Normal	35	115	0.81	0.23	0.73	0.75	0.73	0.74
	OA	8	292	0.72	0.97				
Trial 3	Normal	119	26	0.78	0.82	0.78	0.78	0.78	0.78
	OA	34	93	0.78	0.73				
Abbreviations: OA, temporomandibular joint osteoarthritis.									

Diagnostic performance of AI and expert and agreement of their OPG read with CBCT read is shown in Table 4. AI in Trial 1 (0.80) and expert in Trial 3 (0.85) are most accurate, respectively. However, taking sensitivity and specificity into consideration collectively, it can be said that Trial 3 is most accurate (0.85 for expert and 0.78 for AI). Cohen's kappa is

highest as well in Trial 3 showing substantial level of agreement for expert (0.69) and moderate one for AI (0.56). In all 3 trials, expert is likely to read more accurate than AI. However, the result of McNemar's test shows that AI reads more in concordance with CBCT ($p = 0.336$) in Trial 3 where TMJOA was diagnosed dichotomically.

Table 4
Diagnostic performance and level of agreement in each Trial.

		Diagnostic Performance			Cohen's kappa	Kappa index	McNemar's test
		Accuracy	Sensitivity	Specificity			
Trial 1	Expert	0.81	0.61	0.91	0.54	moderate	.001
	AI	0.80	0.52	0.94	0.51	moderate	.000
Trial 2	Expert	0.69	0.57	0.93	0.42	moderate	.000
	AI	0.73	0.97	0.23	0.25	fair	.000
Trial 3	Expert	0.85	0.72	0.97	0.69	substantial	.000
	AI	0.78	0.73	0.82	0.56	moderate	.366
Abbreviations: AI, artificial intelligence.							

Discussion

OPG is the most widely used plain radiograph in primary diagnosis of TMJOA. However, OMFR experts or CBCT is not always available in most general practice clinics and screening of TMJOA in primary examination is easily compromised. As reported previously, age, pain, and TMJ noise showed no correlation with TMJOA on CBCT, while its high incidence rate of OA changes in TMD patients, 27.3%, was observed¹⁶. It was also reported that 24% of patients who did not show significant condyle bone changes on OPG had degenerative bone changes on CBCT¹⁷. Moreover, it is also very well known that the accuracy, sensitivity, and specificity of OPG is not good for the diagnosis of TMJOA, even when OMFR radiologist read it. OPG showed lower sensitivity (0.26) and higher specificity (0.99) compared to CT in TMJOA¹⁴. The diagnostic accuracy of OPG in detecting cortical erosion of mandibular condyle is lesser (0.55–0.64) than that of CBCT (0.77–0.95)¹⁸. Because the bone tissue must be demineralized sufficiently before getting noticeable on OPG, which usually takes more than 6 months, the mandibular condyle may appear normal on OPG in the early stages of TMJOA^{19,20}. According to the position paper from American Academy of Oral and Maxillofacial Radiology (AAOMR), OPG is only useful in diagnosing advanced TMJOA due to low sensitivity²¹. Image distortion and overlap on OPG always matter as well²². However, considering the prime role of OPG in primary examination, any supplementary diagnostic tool to screen TMJOA on OPG would be practically very helpful in clinics.

AI developed in this study showed sufficient ability as sensitive as OMFS experts in primary diagnosis of TMJOA on OPG. The diagnostic performance (accuracy, 0.51, and F1 score, 0.53) was not satisfying in initial trial when AI trained with 3 categories of TMJOA: normal, indeterminate TMJOA, and TMJOA. The Cohen's kappa value of the AI diagnosis on OPG was 0.27, which is less than the that of the expert, 0.38, while they all represent fair agreement. It implies that it is still difficult to distinguish subtle changes of TMJOA on OPG even for an expert. Besides, multi-label image classification is more challenging than single-label one²³ as shown in a previous study that reported low performance (mean accuracy, 0.51) of AI classifying lower third molar development into multiple stages for age estimations⁷. Based on this initial result, classification of images was done in different ways.

When indeterminate OA was taken as normal in Trial 1, the AI model performance was best, showing highest accuracy, precision, recall, and F1 score, 0.80, 0.81, 0.80, and 0.80, respectively. But, sensitivity in diagnostic performance was

lowest (0.52). In Trial 2 when indeterminate OA was taken as OA, the model performance (accuracy, 0.73, precision, 0.75, recall, 0.73, and F1 score, 0.74) and specificity (0.23) in diagnostic performance were lowest, while sensitivity was highest (0.97). In Trial 3 when indeterminate OA was omitted in the development process, the model performance (accuracy, 0.78, precision, 0.78, recall, 0.78, and F1 score, 0.78) comes close to that in Trial 1, while sensitivity (0.73) and specificity (0.82) were good and more balanced in diagnostic performance. Based on this result, AI performs best for TMJOA diagnosis on OPG when indeterminate OA was omitted during it trained and verified. The pilot study showed similar accuracy (0.77–0.84) with various AI models when indeterminate OA was omitted¹³.

Indeterminate OA was recategorized into normal in previous studies^{12,14} as like Trial 1 in this study. Sensitivity of AI (52.0%) and expert (60.7%) was higher and specificity was lower (94.3% for AI and 90.7% for expert) than those of experts in the previous study (sensitivity, 26.2% and specificity, 99.3%)¹⁴. The model performance (accuracy, 0.86 and F1 score, 0.84) in a recent study¹² was better than in Trial 1 of this study (accuracy, 0.80 and F1 score, 0.80) but the materials for AI development was CBCT sagittal images in that study. CBCT images have higher detail and fewer artifacts at the anatomical boundaries of the ROI and background than OPGs²⁴. Superiority of CBCT to OPG in performance of DCNN model has already been reported²⁵.

Consequently, AI in Trial 3 is most appropriate for TMJOA diagnosis on OPG, showing best balance between sensitivity and specificity among 3 trials and equivalent diagnostic performance to expert. Moreover, no statistical difference was observed between AI diagnosis on OPGs and expert diagnosis on CBCT only in trial 3 (McNemar's test, $p = 0.336$), which must be because of better balance between sensitivity and specificity. It implies that AI model is more likely to diagnose TMJOA on OPG in accordance with diagnosis on CBCT made by experts when indeterminate TMJOA is excluded from training. When AI in trial 3 read untested 115 OPG images of indeterminate TMJOA (data not shown), AI was more likely read indeterminate TMJOA as TMJOA (41 OPGs, 35.7%) than expert (23 OPGs, 20.0%). Indeterminate TMJOA may be considered normal variation, aging, physiologic remodeling, or a precursor to TMJOA¹⁵, which means it can be diagnosed as normal and OA at the same time. Because benefit of high sensitivity may exceed loss of low specificity in diagnosis of TMJOA where early detection is important, it would be clinically more beneficial for a screening tool for TMJOA to read indeterminate OA as OA. In a randomized controlled study, osseous condylar changes in adolescents/young adults with early-stage TMJOA showed repair and even regeneration after conservative and splint therapy²⁶, which may emphasize early detection and management.

AI is an up-and-coming issue in radiology, where a large amount of homogeneous image data is available, showing strength in segmenting, detecting, or classifying organs and lesions. Despite the lack of quality and number of studies, AI has been proved constantly to be equivalent to health-care professionals in diagnostic performance²⁷. On the other hand, AI is more likely to show better results when it trains with a ROI in image than with a whole image²⁸. However, ROIs were extracted manually in most studies, which can compromise its practical usability. In this study, AI was designed to extract ROI from OPG by an algorithm based on object detection technique. The average precision of mandibular condyle detection was reported as 99.4% on right side and 100% on left side in previous study¹³. The ROI extracted automatically for the study covers mandibular condyle, articular fossa and eminence, which is very important because TMJOA is not defined only by bony change of the condyle. It is obvious that OA change occurs in other structures of the TMJ such as fossa and eminence²⁹, so that the ROI for AI diagnosis should cover all structures of the joint, rather not just the condyle. But, most of studies on TMJOA usually focus on the condyle and only one AI study based on TMJ CBCT also used the mandibular condyle extracted from the sagittal Sect. ¹².

This study has several limitations. Firstly, the absolute size of the training dataset was rather small. In order to overcome the limitation of small size dataset, data augmentation by image modification and transfer learning technique with fine-tuning based on imagenet database (<http://image-net.org>) were used. It has been reported that the accuracy of transfer

learning increases to some extent for medical images other than general objects³⁰. Although AI does not exceed expert in accuracy in this study, it may improve in further study by means of larger size of data, learning OPG and CBCT images together, elaborating deep learning models using landmarks, or ensemble between multiple AI models. Secondly, it is well known that the external test dataset in multiple medical centers secure high performance of AI for reproducibility in general use. None the less, this study was based on the internal dataset, but, recently, there is an unnoticeable difference in performance among OPG equipment from various manufacturers. Lastly, deep learning algorithms have inherent uncertainty³¹. Compared to machine learning methods based on handcrafted features, the trained deep learning model is so like a “black-box” as to the inference results are not explainable³².

Conclusions

Collectively, an AI can read OPGs to diagnose TMJOA as well as can OMFR experts in regard to sensitivity and has greater accordance with CBCT interpretations than do OMFR experts, which implies that AI can play an important role in diagnosing TMJOA primarily on OPGs in most general practice clinics, where OMFR experts or CBCT are not available.

Methods

Materials.

The written documentation of informed consent was waived and approved by the decision of the Institutional Review Board of School of Dentistry, Seoul National University (S-D20200004) and ethics committee approval for the study in the same institute was also obtained. All methods were performed in accordance with relevant guidelines and regulations. Radiographic images of the patients who visited the orofacial pain clinic of Seoul National University Dental Hospital complaining of TMD related symptoms and took OPG (Orthopantomograph OP, 100D, Instumentarium corporation, Tuusula, Finland) and TMJ CBCT (SOMATOM Sensation 10, Siemens, Erlangen, Germany) which were read by OMFR specialists from January, 2015 to October, 2019 were reviewed. Records of patients under 18 years of age or with a history of orthognathic surgery, macrotrauma, and systemic diseases that could cause joint deformity or with temporal difference more than 3 months between OPG and CBCT were excluded. AI algorithm was trained, validated, and tested with 1189 OPGs, all of which had been confirmed by additional CBCT examination, selected randomly and classified by an orofacial pain specialist in terms of image analysis criteria for the diagnosis of temporomandibular disorder (research diagnostic criteria for TMD¹⁴; diagnostic criteria for TMD¹⁵: no TMJOA (normal), indeterminate for TMJOA (indeterminate), and TMJOA (OA). Among 2378 joints, 800 were diagnosed as normal, 779 as indeterminate, and 799 as OA based on the read of CBCT (Table 1).

AI model development.

Firstly, an algorithm to extract region of interest (ROI) including the mandibular condyle and surrounding structures from each OPG by object detection technique was developed (Fig. 1). The objective detection technique was based on the Faster Regions with Convolutional Neural Networks (RCNN) using Inception V3 model as categorizing algorithm in which Region Proposal Network (RPN) and Image Classification Network (ICN) work simultaneously and make it faster. In the next step, extracted ROI images were classified into normal, indeterminate, and OA by means of the Keras' ResNet model based on Convolutional Neural Network (CNN). The ROI images of 2378 joints were divided randomly into training (1478 images), validation (450 images), and test sets (450 images), with which AI was developed using Keras' ResNet model. Test set consisted of 150 normal, 150 indeterminate, and 150 OA images. Data augmentation was done by image rotation ± 5 degrees, image shift $\pm 10\%$, brightness $\pm 10\%$, and contrast $\pm 10\%$ to compensate the disadvantage of the small number of the data for robustness of model. Training and validation were repeated 35000 times (700 epochs) with

augmented data. The learning rate of the model was 1.0×10^{-6} and Adam optimizer was used. After 700 training epochs, the validation loss of the model decreased from 12.2 to 0.1 (Fig. 2).

Model and statistical analysis.

Accuracy, precision, recall, and F1 score were calculated for the model performance. Accuracy is defined as the ratio of correct predictions. Precision is the ratio of true positives to true positives and false positives. Recall is the ratio of true positives to true positives and false negatives. Finally, the F1 score is a harmonic mean of precision and recall: $(2 \times \text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$. Accuracy, specificity, and sensitivity were calculated for diagnostic performance, Cohen's kappa was calculated to estimate the agreement of TMJOA diagnosis between OPG and CBCT read, and McNemar's test was done to evaluate significance of difference. For evaluation of clinical usability of AI, the results between OPG reads by AI and expert were compared. All p-values < 0.05 was considered to be different statistically. The Python programming language (v. 3.6), tensorflow (v. 2.0) and a graphic card (GeForce GTX 2080) were used for analysis.

Declarations

Data availability

The data that support the findings of this study are available from the authors upon reasonable request.

Conflicts of interest

The authors have no conflict of interest relevant to this article.

References

1. Wang, X., Zhang, J., Gan, Y. & Zhou, Y. Current understanding of pathogenesis and treatment of TMJ osteoarthritis. *J Dent Res.* **94**, 666-673 (2015).
2. Buckwalter, J. A., Mankin, H. J. & Grodzinsky, A. Articular cartilage: degeneration and osteoarthritis. *Instr Course Lect.* **54**, 465 (2005).
3. Schiffman, E. L. *et al.* The research diagnostic criteria for temporomandibular disorders. I: overview and methodology for assessment of validity. *J. Oral Facial Pain Headache.* **24**, 7 (2010).
4. Larheim, T., Abrahamsson, A., Kristensen, M. & Arvidsson, L. Temporomandibular joint diagnostics using CBCT. *Dentomaxillofacial Radiol.* **44**, 20140235 (2015).
5. White SC, Pharoah MJ. Oral radiology: principles and interpretation(4th edition) 493-498 (Mosby, 2009).
6. Wirtz, A., Mirashi, S. G. & Wesarg, S. Automatic teeth segmentation in panoramic X-ray images using a coupled shape model in combination with a neural network. *Med Image Comput Comput Assist Interv.* 712-719 (2018).
7. De Tobel, J., Radesh, P., Vandermeulen, D. & Thevissen, P. W. An automated technique to stage lower third molar development on panoramic radiographs for age estimation: a pilot study. *J Forensic odontostomatol.* **35**, 42 (2017).
8. Vinayahalingam, S., Xi, T., Bergé, S., Maal, T. & de Jong, G. Automated detection of third molars and mandibular nerve by deep learning. *Sci Rep.* **9**, 1-7 (2019).
9. Yang, H. *et al.* Deep learning for automated detection of cyst and tumors of the jaw in panoramic radiographs. *J Clin Med.* **9**, 1839 (2020).
10. Lee, K.-S., Jung, S.-K., Ryu, J.-J., Shin, S.-W. & Choi, J. Evaluation of transfer learning with deep convolutional neural networks for screening osteoporosis in dental panoramic radiographs. *J Clin Med.* **9**, 392 (2020).

11. Murata, M. *et al.* Deep-learning classification using convolutional neural network for evaluation of maxillary sinusitis on panoramic radiography. *Oral radiol.* **35**, 301-307 (2019).
12. Lee, K. *et al.* Automated detection of TMJ osteoarthritis based on artificial intelligence. *J Dent Res.* **99**, 1363-1367 (2020).
13. Kim, D., Choi, E., Jeong, H. G., Chang, J. & Youm, S. J. Expert system for mandibular condylar detection and osteoarthritis classification in panoramic imaging using R-CNN and CNN. *App Sci.* **10**, 7464 (2020).
14. Ahmad, M. *et al.* Research diagnostic criteria for temporomandibular disorders (RDC/TMD): development of image analysis criteria and examiner reliability for image analysis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* **107**, 844-860 (2009).
15. Schiffman, E. *et al.* Diagnostic criteria for temporomandibular disorders (DC/TMD) for clinical and research applications: recommendations of the International RDC/TMD consortium network and orofacial pain special interest group. *J Oral Facial Pain Headache.* **28**, 6 (2014).
16. Kim, K., Wojczyńska, A. & Lee, J.-Y. The incidence of osteoarthritic change on computed tomography of Korean temporomandibular disorder patients diagnosed by RDC/TMD; a retrospective study. *Acta Odontol Scand.* **74**, 337-342 (2016).
17. Jeon, Y.-M. *et al.* The validity of computed tomography in diagnosis of temporomandibular joint osteoarthritis. *J Oral Med.* **33**, 195-204 (2008).
18. Honey, O. B. *et al.* Accuracy of cone-beam computed tomography imaging of the temporomandibular joint: comparisons with panoramic radiology and linear tomography. *Am J Orthod Dentofacial Orthop.* **132**, 429-438 (2007).
19. Kamelchuk, L. S. & Major, P. W. Degenerative disease of the temporomandibular joint. *J Orofac Pain.* **9** (1995).
20. Cibere, J. Do we need radiographs to diagnose osteoarthritis? *Best Pract Res Clin Rheumatol.* **20**, 27-38 (2006).
21. Brooks, S. L. *et al.* Imaging of the temporomandibular joint: a position paper of the American Academy of Oral and Maxillofacial Radiology. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* **83**, 609-618 (1997).
22. Mawani, F. *et al.* Condylar shape analysis using panoramic radiography units and conventional tomography. *Oral Surg Oral Med Oral Pathol Oral Radiol.* **99**, 341-348 (2005).
23. Deng, J. *et al.* Imagenet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition.* 248-255 (2009).
24. Scarfe, W. C., Farman, A. G. & Sukovic, P. Clinical applications of cone-beam computed tomography in dental practice. *J Can Dent Assoc.* **72**, 75 (2006).
25. Lee, J. H., Kim, D. H. & Jeong, S. N. Diagnosis of cystic lesions using panoramic and cone beam computed tomographic images based on deep learning neural network. *Oral Dis.* **26**, 152-158 (2020).
26. Lei, J., Yap, A. U. J., Liu, M. Q. & Fu, K. Y. Condylar repair and regeneration in adolescents/young adults with early-stage degenerative temporomandibular joint disease: A randomised controlled study. *J Oral Rehab.* **46**, 704-714 (2019).
27. Liu, X. *et al.* A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Health.* **1**, e271-e297 (2019).
28. Do, S., Song, K. D. & Chung, J. W. Basics of Deep Learning: A Radiologist's Guide to Understanding Published Radiology Articles on Deep Learning. *Korean J Radiol.* **21**, 33-41 (2020).
29. O'Ryan, F. & Epker, B. N. Temporomandibular joint fuction and morphology: observations on the spectra of normalcy. *Oral Surg Oral Med Oral Pathol.* **58**, 272-279 (1984).
30. Poedjiastoeti, W. & Suebnukarn, S. Application of convolutional neural network in the diagnosis of jaw tumors. *Health Inform Res.* **24**, 236-241 (2018).

31. Gal, Y. Uncertainty in deep learning. *Los Altos: IEEE/ACM Transactions on Audio, Speech, and Language Processing*. (2017).
32. Holzinger, A. From machine learning to explainable AI. *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA). Piscataway: IEEE*. 55-66 (2018).

Figures

Figure 1

Result of Extraction of ROI, 300x300 pixels.

Figure 2

Change in validation loss as a result of model training. Validation loss decreased as the epochs increased.