Deep Learning in Dental Radiographic Imaging

Hyuntae Kim
Department of Pediatric Dentistry, Seoul National University Dental Hospital, Seoul, Republic of Korea

Abstract
Deep learning algorithms are becoming more prevalent in dental research because they are utilized in everyday activities. However, dental researchers and clinicians find it challenging to interpret deep learning studies. This review aimed to provide an overview of the general concept of deep learning and current deep learning research in dental radiographic image analysis. In addition, the process of implementing deep learning research is described. Deep-learning-based algorithmic models perform well in classification, object detection, and segmentation tasks, making it possible to automatically diagnose oral lesions and anatomical structures. The deep learning model can enhance the decision-making process for researchers and clinicians. This review may be useful to dental researchers who are currently evaluating and assessing deep learning studies in the field of dentistry. [J Korean Acad Pediatr Dent 2024;51(1):1-10]

Keywords
Artificial intelligence, Deep learning, Dentistry, Image classification, Image segmentation

Introduction
Since the start of the 20th century, extraordinary advances in medical technologies have extended the human lifespan, leading to improvements in human well-being. As life expectancy has increased, dental care has become more important, and it is widely recognized that maintaining healthy natural teeth improves the quality of life[1-3]. Teeth play a critical role in food consumption by breaking down food into smaller fragments to facilitate its digestion in the digestive system. The initial stage of the digestive system occurs in the oral cavity, which includes teeth and adjacent tissues[4].

ORCID
Hyuntae Kim
https://orcid.org/0000-0003-2915-8584

Article history
Received February 11, 2024
Revised February 17, 2024
Accepted February 17, 2024

Corresponding author: Hyuntae Kim
Department of Pediatric Dentistry, Seoul National University Dental Hospital, 101 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea
Tel: +82-2-6256-3265 / Fax: +82-2-6256-3266 / E-mail: kht131@snu.ac.kr

© 2024 Korean Academy of Pediatric Dentistry
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
However, as the human oral cavity harbors a diverse array of microorganisms[5-7], dental caries[8,9] and periodontal disease[10,11] have persisted despite rapid advances in modern dentistry. Therefore, dental caries, periodontal disease, and the resulting tooth loss are the primary challenges faced by individuals. As it is difficult to visually inspect the oral cavity oneself, many people only visit a dentist when they experience pain or discomfort.

When a patient with dental caries or another oral disease arrives at a clinic, the dentist diagnoses the severity of the oral disease and formulates a treatment plan. Early intervention is important to minimize patient discomfort and damage to teeth and oral tissues. Dental radiographic imaging techniques provide information imperceptible to the visual and tactile senses, allowing dentists to diagnose oral diseases with a high degree of sensitivity.

Computer-aided diagnosis (CAD) has evolved significantly with the digitization of radiographic information. Recently, with the active advancement of artificial intelligence (AI) research, computer-based diagnostic procedures have become increasingly powerful and effective[12].

Deep learning, which is a subset of machine learning, is a computer technology that has triggered the development of AI applications. The aim is to create machines with human-like intelligence based on deep neural networks. In this review, I briefly overview the basic concepts of deep learning and describe the current applications of deep learning in dental radiographic imaging.

Overview of deep learning

Deep learning is a subset of machine learning that consists of deep layers of artificial neural networks[13]. The name “deep” refers to the architecture of the model, which has multiple layers (Fig. 1). Unlike traditional machine learning, deep learning does not require manual extraction of features, but rather the training of models from large amounts of data (Table 1)[14]. Different deep learning networks are used for different cases and data types. For example, recurrent neural networks are often used for natural language processing and speech recognition, while convolutional neural networks (CNNs) are more commonly used for classification and computer vision tasks. CNNs are composed of three main types of layers: convolutional layer, pooling layer, and fully connected layer. With each layer, the CNN extracts image features such as edges and colors, then begins to recognize larger elements or shapes of the objects, until it finally identifies the intended object[15]. CNNs are distinguished from other neural networks by their superior performance in processing pixel data, and have shown promising results in computer vision and medical diagnosis[16].

As computing power has increased dramatically over the past 10 years, neural networks have become more sophisticated, with multiple layers and connections. This is called “deep learning” (DL).
What can deep learning do?

1. Classification

Classification is the task of determining the structure belonging to a particular category. This usually involves categorizing the presence of a lesion. For supervised learning, human experts need to ensure accurate classification and reference standard settings (Fig. 2).

ResNet and GoogLeNet are the main deep learning architectures used in the classification tasks. ResNet is an advanced CNN with many deep layers incorporated into its framework. It is useful for problems that require extremely deep structures, hence it is also used for object detection and segmentation[17]. GoogLeNet utilizes the inception module, which optimizes the utilization of computational resources by incorporating multiple filter sizes within a single layer. It is built on a 22-layer deep neural network architecture[18].

Deep learning-based classification models can be used to diagnose various intraoral lesions or anatomical structures. These include impacted mesiodens[19], dental caries[20], dental implant[21], and temporo-mandibular joints[22].

2. Object detection

This task involves identifying an object that exists at a particular location. A bounding box is used to indicate a specific location within an image. The deep learning model utilizes a substantial dataset of region of interests (ROIs) that have been manually annotated by expert dentists. It then autonomously identifies and defines the ROIs in the objective images. Using an object detection model, it is possible to simultaneously identify the class to which an object belongs. For example, it is possible to classify each tooth separately on a periapical radiograph and determine the category to which the tooth belongs (Fig. 3)[23].

R-CNN object detection architecture extracts 2000 region proposals from an image and feeds them to a convolutional neural network. These features are then passed to a support vector machine for the purpose of classifying the presence of an object within each candidate region proposal[24]. Since R-CNN models use different CNN networks for each of the 2000 different candidate regions, they are costly and require long training times. Fast and faster R-CNN are modification of R-CNN using single CNN for the proposed regions and region proposal networks[25,26]. You Only Look Once (YOLO) model, on the other hand, focuses on parts of the image that have a high probability of containing the object and provides faster results. However, it encounters difficulties in detecting small objects due to the spatial limitations of the algorithm[27].

Object detection models can be applied to various oral lesions and their anatomical structures. In panoramic radiographs, deep learning models can detect teeth and

Fig. 2. Example of the classification task. The impacted mesiodens classification model denotes (A) no mesiodens, and (B) mesiodens.

Fig. 3. Example of object detection task. The model automatically detects teeth in the panoramic radiograph.
their numbering[28], impacted mesiodens[29], and cleft jaw and palate[30].

3. Segmentation

The boundaries of an object can be segmented and classified by identifying the characteristics of each pixel in the image. It is also possible to determine the class to which an object belongs. The annotation of segmentation tasks requires tracing the outline of the target, whereas object detection only encloses the target using a rectangular bounding box. Therefore, the segmentation task is generally more informative because it provides detailed information about anatomical structures (Fig. 4).

For pixelwise segmentation, Mask R-CNN and U-Net are mainly used. Mask R-CNN is a two-stage framework that builds on the Faster R-CNN object detection model by adding a mask prediction in parallel with the existing object detection branch[31,32]. While U-Net was initially developed to segment biomedical images, it can be applied to other images as well. It is an encoder-decoder architecture that uses skip connections to transfer low-level features from the encoder to the decoder[33].

In panoramic and intraoral radiographs, a segmentation task is performed to extract periapical lesions[34], the mandibular canal and maxillary sinus[35], and dental caries[36,37]. Image segmentation can be performed using computed tomographic images[38,39]. Nozawa et al.[40] conducted a study on the segmentation of the temporomandibular joint disc using MRI images.

Deep learning procedure

1. Data processing

In supervised learning, data annotation refers to the process of labeling structures in an image and creating specifications for a reference standard. Mohammad-Rahimi et al.[41] categorized reference standards into five subgroups.

- Gold standard: Histological assessment of dental hard tissues (caries) or soft tissues (oral mucosal lesions) is a reliable reference.
- Consensus: The diagnosis of a condition is agreed upon by a group of clinicians.
- Majority voting: An expert panel votes on the diagnosis
- Intersection: Pixels affected by a condition are labeled independently by two or more clinicians. The intersection of the segmentation masks serves as the reference standard.
- Union: All pixels segmented by all clinicians can serve as the reference standard.

The total datasets must be separated into training, validation, and test sets. The training set is used for model training and parameter optimization. During training, model performance is optimized using the validation dataset. To test the trained model, separate hold data (not included in the training process) should be used to evaluate the performance of deep learning models. This is the only method that expresses the generalizability of the trained model.

**Fig. 4.** Example segmentation task of dental caries. (A) An intraoral digital X-ray, (B) The segmentation model delineates dental caries in the digital X-ray.
2. Deep learning workflow

Rahimi’s deep learning workflow is presented below. This is a modification of the work of Montagon et al.[42], which is required for medical and dental deep learning projects (Table 2)[41].

First, the clinical application that is to be applied to the model should be defined. Second, each project should have a project manager with basic knowledge of the technical and clinical aspects. Third, ethical approval and informed consent need to be obtained from the patients. Fourth, funding of human and hardware resources is required. Next, data collection, model development, and model assessment are performed using team-based models. Finally, this model can be implemented for clinical adoption and requires regular monitoring. A model can be retrained by using additional data to update its performance.

3. Model evaluation metrics

The model’s performance is measured using different modalities of AI models. For classification, the accuracy, sensitivity, specificity, precision, and F1-score are evaluated. For the segmentation and object detection tasks, the Jaccard index and Dice score are used for evaluation (Table 3).

<table>
<thead>
<tr>
<th>Table 2. Deep learning project checklists</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step title</td>
</tr>
<tr>
<td>Defining project scope</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Team building</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Ethics</td>
</tr>
<tr>
<td>Hardware and software</td>
</tr>
<tr>
<td>Data collection</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Model development</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Model assessment</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Regulatory</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Clinical adoption</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3. Model evaluation metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metric</td>
</tr>
</tbody>
</table>
| Accuracy                                | Ratio between correct predictions and total      | \[
|                                          | predictions                                      | \[
|                                          | \[
|                                          | \[
| Sensitivity (Recall)                    | Proportion of actual positives predicted as       | \[
|                                          | positives                                        | \[
| Specificity                             | Proportion of actual negatives predicted as       | \[
|                                          | negatives                                        | \[
| Precision                               | How many of the predicted positives are actually  | \[
|                                          | positive                                         | \[
| F1-score                                | A combination overview of precision and recall    | \[
|                                          | metrics                                          | \[
| Jaccard index                           | Area of intersection of two segmentations divided | \[
|                                          | by the area of union between them                 | \[
| Dice score                              | Area of intersection of two segmentation masks    | \[
|                                          | multiplied by two, then divided by their total     | \[
|                                          | area                                             | \[

TP: true positive; TN: true negative; FP: false positive; FN: false negative; GT: ground truth; MS: machine segment.
Application in the field of dentistry

1. Dental caries diagnosis

Dental caries are chronic oral diseases that pose a significant threat to oral and human health[43]. Lee et al.[20] introduced a CNN algorithm with GoogLeNet Inception v3 for caries detection on periapical radiographs, and the diagnostic accuracies of premolars and molars were 89.0% and 88.0%, respectively. Lee et al.[44] used a segmentation model called U-Net for caries detection with an F1-score of 65.02%. Although the algorithm did not outperform the dentist group, the overall diagnostic performance of all clinicians improved significantly with the help of the results from the trained model. Cantu et al.[45] used a U-Net model to segment dental caries on bitewing radiographs. Their model exhibited an accuracy of 80%. Generally speaking, in deep learning, a good accuracy score would be above 70%, and if the accuracy is between 60 - 70%, the model can be considered acceptable[46].

2. Orthodontics

For a successful orthodontic treatment, it is important to establish a treatment plan based on an accurate diagnosis. Cephalometric landmarks play an important role in orthodontic diagnosis.

Park et al.[47] evaluated the performance of a deep learning model that automatically measures landmarks on a lateral cephalogram. The YOLO v3 model exhibited an accuracy of 80.4 - 96.2%, which is 5% higher than that of other methods.

Yu et al.[48] constructed a multimodal CNN model based on 5,890 lateral cephalograms to test skeletal classification models. The accuracy, sensitivity, and specificity were greater than 90%.

Deep learning models can assist dentists in deciding whether to perform extraction or non-extraction treatments. These models use numerical data from cephalometric analysis and demonstrate relatively higher accuracy[49,50].

3. Endodontics

Accurate identification of the complex anatomical structures of the root canal system has a significant impact on the success of endodontic treatment. Diagnosing the presence or absence of a periapical lesion at an appropriate time and initiating treatment are important for the success of root canal treatment.

Fukuda et al.[51] used a CNN model to detect vertical root fractures in cone-beam computed tomography (CBCT) images. The model effectively detected 267 out of 330 vertical root fractures. The deep learning model exhibited a precision of 0.93 and an F-score of 0.83, indicating that it is clinically applicable.

Hiraiwa et al.[52] used a deep-learning algorithm to detect extra roots in mandibular first molars. CBCT images were used as the reference standard, and the diagnostic accuracy of the deep learning model was 86.9%.

4. Periodontology and Implantology

Periodontitis is an inflammatory disease caused by bacteria that leads to the formation of periodontal pockets, alveolar bone loss, and tooth loosening. Periodontal disease is the most common oral disease that occurs throughout a human's lifespan, and it is considered the most common cause of tooth loss throughout a human's lifespan.

Krois et al.[53] applied deep CNN models to detect periodontal bone loss in panoramic dental radiographs using 2001 image segments. The trained CNN model demonstrated diagnostic performance comparable to that of dentists.

Chen et al.[54] devised a new deep-learning ensemble model based on CNN algorithms to predict tooth position, shape, remaining interproximal bone level, and radiographic bone loss using periapical and bitewing radiographs. The new ensemble model is based on YOLOv5, VGG16, and U-Net. The model accuracy was approximately 90% for periapical radiographs, and its performance was superior to that of dentists.

Chang et al.[55] developed a model that automatically
determines the stage of periodontal disease on panoramic radiographs using a deep learning algorithm. The correlation index between the results of the deep learning model and the diagnosis results between radiologists was 0.73, and the intraclass correlation index was 0.73. It was found to be 0.91, showing high accuracy and reliability and providing effective results in automatic diagnosis of the degree of alveolar bone loss and stage of periodontal disease.

Kong et al.[56] used data from 14,037 implant images in the training and test datasets using YOLO v5 and v7, which are object detection models, and the mean average precision (mAP) was evaluated. The mAP of YOLO v7 was higher than that of YOLOv5, with the best value being 0.984.

In a multicenter study on the classification of dental implant systems, the deep learning model outperformed most participating dental professionals, with an AUC of 0.954[57].

5. Oral and maxillofacial surgery

Deep learning algorithms are used to diagnose lesions in the oral and maxillofacial areas and establish surgical plans. Studies have been performed to isolate and evaluate the important anatomical structures that must be considered during surgical procedures.

Poedjiastoeti and Suebnukarn[58] created a convolutional neural network using VGG-16 and trained ameloblastoma and keratocystic odontogenic tumors in 500 digital panoramic images. The accuracy of the trained model was 83.0%, and the diagnostic time was much faster than that of oral and maxillofacial specialists.

Jung et al.[59] segmented the maxillary sinus into the maxillary bone, air, and lesion and evaluated its accuracy by comparing it with human experts. They adopted the 3D-UNet model in CBCT, and the performance measure was the Dice coefficient. They concluded that a deep learning model could alleviate annotation efforts and costs by efficiently training CBCT datasets.

Conclusion

Deep learning has been widely applied in dentistry and is expected to have a significant impact on dental and oral healthcare in the future. This review provides an overview of deep learning and the recent research in dentistry. It also provides practical guidance on deep learning for dental researchers and clinicians.

Deep learning models have shown great potential for improving diagnosis in dentistry. However, because of the inherent complexity of the models, humans cannot evaluate how a model arrives at a particular decision. As the complexity of deep-learning models increases, their interpretability decreases. This is why deep learning models are called “black boxes.” In addition, a large amount of data is required to obtain a highly accurate model, which requires considerable time and effort.

This review mainly focused on the deep learning-based analysis of dental radiographs. From a broader perspective, deep learning is not limited to image analysis. Natural language processing (medical interviews and electronic medical record data) and predictive analysis (epidemics of infectious diseases, number of patients, and prognosis of diseases) can be another focus of deep learning. There is still ample scope for deep learning in dentistry, and a concerted effort is needed to follow the latest technological advances and develop new clinical and research applications.

Conflict of Interest

The author has no potential conflicts of interest to disclose.

References

2. Park HE, Song HY, Han K, Cho KH, Kim YH : Number of remaining teeth and health-related quality of


45. Cantu AG, Gehring S, Krois J, Chaurasia A, Rossi JG,


