Original Article

Implant Thread Shape Classification by Placement Site from Dental Panoramic Images Using Deep Neural Networks

Sujin Yang, DDS, MSD, Ph.D¹⁺, Youngjin Choi, DDS²⁺, Jaeyeon Kim, DDS, MSD, Ph.D³, Ui-Won Jung, DDS, MSD, Ph.D⁴, Wonse Park, DDS, MSD, Ph.D⁵⁺

¹Clinical Professor, Department of Advanced General Dentistry, College of Dentistry, Yonsei University, Seoul, Korea

²Public Health Doctor, Wang-gung Health center, Iksan-si, Jeollabuk-do, Korea
 ³PhD Student, Department of Advanced General Dentistry, College of Dentistry, Yonsei University, Seoul, Korea
 ⁴Professor, Department of Periodontology, College of Dentistry, Yonsei University, Seoul, Korea
 ⁵Professor, Department of Advanced General Dentistry, College of Dentistry, Yonsei University, Seoul, Korea

*Corresponding author: Wonse Park, Department of Advanced General Dentistry, College of Dentistry, Yonsei University, Seoul 03722, Korea

Tel: +82-2-2228-8980. Fax: +82-2-2227-8906. E-mail: wonse@yuhs.ac

⁺These authors equally contributed to this work.

Abstract

Purpose: In this study, we aimed to classify an implant system by comparing the types of implant thread shapes shown on radiographs using various Convolutional Neural Networks (CNNs), particularly Xception, InceptionV3, ResNet50V2, and ResNet101V2. The accuracy of the CNN based on the implant site was compared.

Materials and Methods: A total of 1000 radiographic images, consisting of eight types of implants, were preprocessed by resizing and CLAHE filtering, and then augmented. CNNs were trained and validated for implant thread shape prediction. Grad-CAM was used to visualize class activation maps (CAM) on the implant threads shown within the radiographic image.

Results: Averaged over 10 validation folds, each model achieved an AUC of over 0.96: AUC of 0.961 (95% CI 0.952 – 0.970) with Xception, 0.973 (95% CI 0.966-0.980) with InceptionV3, 0.980 (95% CI 0.974-0.988) with ResNet50V2, and 0.983 (95% CI 0.975-0.992) with ResNet101V2. Accuracy was higher in the posterior region than in the anterior area in all four models. Most CAMs highlighted the implant surface where the threads were present; however, some showed responses in other areas.

Conclusion: The CNN models accurately classified implants in all areas of the oral cavity according to the thread shape, using radiographic images.

Keywords: Artificial intelligence, Convolutional neural networks, Classification, Deep learning, Implant system

www.kci.go.k

OPEN ACCESS

pISSN 2765-7833 eISSN 2765-7841 Journal of implantology and applied sciences 2024; 28(1): 18-31 https://doi.org/10.32542/implantology.2024003

Received: February 8, 2024 Revised: March 19, 2024 Accepted: March 23, 2024

ORCID

 Sujin Yang

 https://orcid.org/0000-0001-5400-2667

 Jaeyeon Kim

 https://orcid.org/0000-0002-0603-619X

 Ui-Won Jung

 https://orcid.org/0000-0001-6371-4172

 Wonse Park

 https://orcid.org/0000-0002-2081-1156

Informed Consent Statement Informed consent was obtained from the subjects involved in the study.

Conflict of Interest The authors declare no conflict of interest.

Copyright © 2024. The Korean Academy of Oral & Maxillofacial Implantology



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.

I. Introduction

With the rapid advancements in implant dentistry, dental implants are now commonly used to restore edentulous areas in elderly patients.¹ As a result, the ability to treat the elderly with dental implants has significantly improved and is expected to play a critical role in improving the quality of life for the elderly in the upcoming super-aged society. However, as the number of patients with implants increases worldwide, so does the frequency with which dentists encounter implant-related complications. Studies have revealed that the 10-year success rate of dental implants is relatively high at 95%; however, also reports frequent complications, the incidence of biological complications within 5 years being 7.1%, and the incidence of screw loosening among mechanical complications being 8.8%.² To treat these complications, implants require constant maintenance and management, and the compatibility between implant systems is crucial. This can only be achieved through accurate identification and information needs to be shared between remote countries and clinics, resulting in the additional costs and time needed to request information from a specialized company.

More than 4000 implant brands from 500 companies are available worldwide.³ The proprietary features of each manufacturer's implant fixture (straight, tapered, conical), surface treatment (machined blasted, acid-etched), thread shape (buttress, reverse buttress, V-shaped), abutment, and fixing screws differ from one another.⁴ To identify a particular implant system, a clinician uses periapical or panoramic radiographs and compares the images according to such features. However, distortion and resolution issues in radiographs, along with the numerous systems available in the market, render the identification and classification of a certain implant a difficult task.

Recently, virtual artificial intelligence (AI) has played an important role in the medical field.⁵ Among AI, convolutional neural networks (CNNs) have demonstrated outstanding performance in finding and classifying features in medical images.⁶ Consequently, several attempts have been made to apply these neural networks in dentistry, particularly for classifying dental implants.⁷ The shortcoming of these artificial neural networks is that, although they exhibit great efficiency, they can only specify the types of implants they have been trained with and cannot decipher types that have not been included in the learning process. The aim of this study was to predict the type of implant by classifying various implant characteristics using artificial intelligence and then merging these features to identify the fittest candidate, similar to human judgment pathways.

Thread shape is a characteristic of most implants and can be distinguished on a radiographic image to some extent, even with the naked eye. The aim was to classify implant systems by comparing the types of implant thread shapes shown on radiographs using various CNNs, particularly Xception, InceptionV3, ResNet50V2, and ResNet101V2. In addition, the accuracy of the CNN depending on the

implant site (maxillary incisors, maxillary molars, mandibular incisors, and mandibular molars) was compared. The null hypothesis was that there would be no significant difference in the detection performance between the CNN models, and the performance would be similar across all implant sites.

II. Materials and Methods

1. Data Collection

The participants of this study were retrospectively selected from patients who underwent implant treatment in two departments of a single institution between January 2015 and January 2020. This study was approved by the Institutional Review Board (IRB) (approval no. 2-2020-0080) and was performed in accordance with the Declaration of Helsinki. The research was performed in accordance with the relevant guidelines and regulations. Due to the retrospective nature of the study, the requirement for written informed consent was waived by the IRB. A total of 435 and 465 participants from the two departments, respectively, were included. Panoramic images were taken using a Pax-i plus (Vatech Co., Hwaseong, Korea) with standard parameters, including a tube voltage of 73 kV, tube current of 9 mA, and acquisition time of 13.5 s. The panoramic images of each participant were collected by exporting radiographic images from the Zetta Pacs Viewer database. The images were anonymized, cropped to contain only the implant fixture area, and adjusted to the Joint Photograph Expert Group (JPEG) format. The implants included in this study were those without prosthetic work (i.e., implants with only a cover screw or healing abutment) to ensure that the prediction of the CNN would not be affected by the prosthesis.

A total of 1000 images consisting of eight types of implants were used; Dentium Implantium (Implantium; Dentium, Seoul, Korea), Dentium NR line (NR line; Dentium), Osstem TS III (TS III; Osstem, Seoul, Korea), Straumann BL (Straumann Bone Level Implant; Institut Straumann AG, Basel, Switzerland), Straumann BLT (Straumann Bone Level Tapered Implant; Institut Straumann AG), Straumann Standard (Straumann Standard Implant; Institut Straumann AG), Straumann Standard Plus (Straumann AG), and Straumann Tapered Effect (Straumann Tapered Effect Implant; Institut Straumann AG).

2. Assessment of Implant Type according to Thread Shape

Implant geometry can be classified using three criteria: thread pitch, thread shape, and thread depth and width. The primary geometry used in this study was the implant thread shape. Initially, the implant thread shape was V-shaped, but it was soon modified when it became aware that the implant thread shape was affected by the stress applied to the bone surrounding the implant fixture. Currently, there are five types of implant thread shapes on the market: V-shape, square shape, buttress, reverse buttress, and spiral shape. In this study, three of the five most commonly used thread shapes (buttress, reverse buttress, and V-shape) were included.⁸ A total of 356 images of Dentium implants with buttress-shaped threads, 463 images of Straumann Standard, Straumann Standard Plus, Straumann BL, Straumann BLT, Straumann Tapered Effect with reverse buttress-shaped threads, and 181 images of Dentium NR Line and Osstem TS III with V-shaped threads were included. The images were manually annotated by a dentist with over 7 years of expertise, according to the surgical records of each patient.

3. Assessment of Implant Type according to Implant Site

To evaluate the AI accuracy according to the region within the panoramic image, the implant placement site was divided into four areas: maxillary incisor, maxillary molar, mandibular incisor, and mandibular molar (incisor refers to the inter-canine area, while molar refers to the premolar and molar areas). There were 90 images per site for the maxillary incisors, 365 for the maxillary molars, 40 for the mandibular incisors, and 485 for the mandibular molars. Table 1 summarizes the baseline characteristics of the implants included in the study.

Subject Characteristics				
Implant thread shape				
Buttress-shaped				
Dentium Implantium	356			
Reverse buttress-shaped				
Straumann BL	197			
Straumann BLT	16			
Straumann Standard	32			
Straumann Standard Plus	199			
Straumann Tapered Effect	19			
V-shaped				
Dentium NRLine	64			
Osstem TS III	117			
Sum	1000			
Implant site				
Maxillary incisor	94			
Maxillary molar	371			
Mandibular incisor	41			
Mandibular molar	494			
Sum www.kci.go.kr	1000			

Table 1. Baseline characteristics of the implants according to implant type, thread type, and site

4. Preprocessing and Data Augmentation

Selected radiographic images were resized according to the requirements of each model. Xception was modified to $229 \times 229 \times 24b$, InceptionV3 to $299 \times 299 \times 24b$, ResNet50V2, and ResNet101V2 to $224 \times 224 \times 24b$. In addition, a contrast-limited adaptive histogram equalization (CLAHE) filter was applied to increase the image visibility.

Next, a ten-fold cross-validation procedure was performed. Cross-sampling is a resampling procedure used to evaluate machine learning models with limited dataset, and aims to overcome deviations in the datasets used to train a deep learning model for image classification.⁹ The datasets were randomly split into ten partitions, with one part used as the testing dataset and the other nine parts used as the training dataset. Care was taken to ensure that the training and test data did not contain samples from the same tooth or patient. Data augmentation was used to increase the number and diversity of the images. This was randomly performed with a vertical flap of the image, rotation of -10 degrees to +10 degrees, shift from -10% to +10%, and distortion of 10%.

5. The Architecture of the Deep CNN Algorithm

Various neural network models are now readily available from the Keras in TensorFlow. In this study, Xception, ResNet50V2, ResNet101V2, and InceptionV3 were selected. A GlobalAveragePooling2D layer and four dense layers were added at the end of each model to visualize the model and classification. The output numbers of the four dense layers were 512, 256, 128, and 3, and the activation mode was set to ReLU for the upper 3 and Softmax for the lowest. All the dense layers were connected to a fully connected layer. The model architectures and hyperparameters were optimized using a randomized search. Fig. 1 shows the overall flow starting from image preprocessing and the schematic structure of the model.



Fig. 1. The overall scheme of the pre-trained convolution neural network. Image preprocessing was done following the acquirements of each model and applying a Contrast Limited Adapted Histogram equalization (CLAHE) filter for enhanced visibility.

	Training			Test		
Total count	900			100		
Fold	Buttress	Reverse-Buttress	V-shaped	Buttress	Reverse-Buttress	V-shaped
1	323	414	163	33	49	18
2	319	418	163	37	45	18
3	319	415	166	37	48	15
4	323	417	160	33	46	21
5	325	410	165	31	53	16
6	309	426	165	47	37	16
7	326	417	157	30	46	24
8	317	423	160	39	40	21
9	321	411	168	35	52	13
10	322	416	162	34	47	19

Table 2. Training, Test, and Validation of dataset according to implant thread shape using 10-fold cross-validation. The diagnostic performance of each fold was obtained and the average of the 10-fold procedure was used as the estimated performance

The dataset classification according to the type of implant thread followed by a 10-fold cross-validation is shown in Table 2.

The algorithms were run on an Intel[®] CoreTM i7-8750H CPU @ 2.20GHz, with a RAM of 16GB, on a 64-bit operating system, and using an NVIDIA GeForce GTX 1050 Ti GPU.

Categorical cross-entropy was used as the loss function for learning, and RMSprop was used as the optimizer. The learning process was conducted over 100 epochs. The learning rate was set to 2×10^{-5} and the batch size was set to 32.

6. Statistical Analysis

The accuracy and area under the curve (AUC, calculated from receiver operating characteristic (ROC) analysis) of each of the four models were obtained. Based on the diagnostic performance of each model from each fold, the accuracy and loss of each model were compared using the Mann-Whitney U Test, and the AUC values were compared using McNemar's chi-square analysis. The level of significance was set at p < .05.

III. Results

1. Accuracy and Loss in Training and Validation Datasets

Fig. 2 shows the average accuracy and loss plots of the training process for each model. For the

training dataset, all four models exhibited convergence of accuracy and loss.

According to the results from the validation dataset, the accuracy of ResNet101V2 was the highest at 98%, and it exhibited the lowest loss. Xception exhibited an accuracy of 96% and the lowest loss was 0.003973%. Each model achieved an AUC of > 0.96, 0.961 (95% CI 0.952–0.970) with Xception, 0.973 (95% CI 0.966-0.980) with InceptionV3, an AUC of 0.980 (95% CI 0.974-0.988) with ResNet50V2, and 0.983 (95% CI 0.975-0.992) respectively. When comparing the accuracy between the



Fig. 2. The average accuracy and loss during 100 epochs from the 10-fold cross-validation procedure. The X-axis represents the number of learning epochs, and the Y-axis represents the accuracy and loss respectively. The red line indicates the training dataset and blue the test dataset.

Model	Accuracy (%, 95% CI)	Loss (%, 95% CI)	AUC (%, 95% CI)	Accuracy by implant site(%, 95% CI)			
				Maxilla anterior	Maxilla posterior	Mandible anterior	Mandible posterior
Xception	0.9600	0.1742	0.9613	0.8958	0.9789	0.8467	0.9681
	(0.9483-0.9716)	(0.1051-0.2432)	(0.9523-0.9702)	(0.8299-0.9616)	(0.9691-0.9887)	(0.7352-0.9581)	(0.9559-0.9803)
Inception	0.9680	0.1637	0.9729	0.9166	0.9865	0.8583	0.9735
V3	(0.9593-0.9766)	(0.1038-0.2236)	(0.9655-0.9802)	(0.8785-0.9547)	(0.9776-0.9954)	(0.7639-0.9528)	(0.9650-0.9821)
ResNet50	0.9780	0.1368	0.9809	0.9361	0.9892	0.9083	0.9817
V2	(0.9675-0.9884)	(0.0593-0.2143)	(0.9738-0.9881)	(0.8818-0.9903)	(0.9751-1.0000)	(0.8158-1.0000)	(0.9724-0.9909)
ResNet101	0.9800	0.1090	0.9834	0.9476	0.9812	0.9292	0.9918
V2	(0.9690-0.9909)	(0.0407-0.1760)	(0.9749-0.9919)	(0.9031-0.9920)	(0.9704-0.9920)	(0.8521-1.0000)	(0.9828-1.0000)
<i>p</i> -value	.0426*	.5579	.0021*				

Table 3. Diagnostic performance between deep learning AI and accuracy according to the implant site

CI: confidence interval, AUC: Area under the receiver operating characteristic curve, Mx ant: maxillary anterior, Mx post: maxillary posterior, Mn ant: mandibular anterior, Mn post: mandibular posterior.



Fig. 3. Multiclass classification confusion matrix using (A) Xception, (B) Inception, (C) ResNet50V2, (D) ResNet101V2 without normalization. The diagonal elements represent the number of points where the predicted label matches the actual label, while the non-diagonal elements show the incorrect detections made by the classifier. The higher the diagonal value and the darker the shade of blue, the more accurate the diagnosis.

different placement sites, the accuracy was higher in the posterior region than in the anterior area in all four models. ResNet50V2 demonstrates an accuracy of > 85% in all areas. The diagnostic performance measures for determining the implant type for each deep learning model are presented in Table 3. The p-value between the four models considering accuracy was 0.0426, loss was 0.5579, and 0.0021 for AUC. Multiclass classification confusion matrices were drawn from the results, as shown in Fig. 3.

2. Visualization of Model Classification

Deep CNNs demonstrate high predictions, yet do not explain how these predictions are made. Thus, to compare model performance based on human perception and add interpretability to the model, Grad-CAM was applied to visualize class activation maps (CAM) on the implant threads shown in the radiographic image. A Class Activation Map can be drawn using a Global Average Pooling Layer.¹⁰



Fig. 4. Radiograph examples and activation maps. (Top two rows) Examples of predictions obtained from ResNet101V2 according to implant thread shape. (Lower two rows) Class Activation Map (CAM)s drawn from the Global Average Pooling (GAP) layer. The model generally focuses on the implant threads, but in some cases, it highlights the connection area.

	17	Tooth number	41
		Ground truth	V-shaped
		Prediction with score	Buttress 99.6%
		Tooth number	14
		Ground truth	Reverse-Buttress
	23	Prediction with score	V-shaped 99.9%
		Tooth number	16
		Crewed truth	
		Grouna truth	BUTTRESS
		Prediction with score	Reverse-Buttress 100%

Fig. 5. Failure analysis of incorrectly classified cases. Heat maps are visible on the healing abutment or the connection area instead of the threaded portion of the implant, implying overfitting.

When the CNN visualizes the extracted features, most CAMs highlight the implant surface where the threads are present. However, some CAMs show a response in the healing abutment or the connection between the healing abutment and the fixture, rather than the threaded portion of the implant. Examples of the prediction accuracy and CAM obtained from ResNet101V2, which exhibited the highest accuracy among the four models, are shown in Fig. 4. A failure analysis of the incorrectly classified cases is shown in Fig. 5.

IV. Discussion

Research to identify various dental implant systems by employing artificial intelligence algorithms is actively underway. One report showed an accuracy of 93-98% in classifying four types of implants using 801 periapical radiographs and five pre-trained CNN models,¹¹ while another reported an accuracy of 86-93% in distinguishing 11 implant types among 8859 panoramic images using five CNN models.¹² These attempts used deep learning algorithms, a form of data-driven AI built on a bottom-up concept that trains mathematical models with an image database.¹³ The neural networks used in these studies, each using small or large datasets, implement a learning process that adapts image features from the input dataset and performs object detection and classification. The diagnostic accuracies of these studies are relatively high; however, they fail to demonstrate the process by which the actual decision was made. Furthermore, the process cannot be interpreted in a medically acknowledged context and fails to provide interpretability and transparency.¹⁴ To add interpretability to the algorithm, a knowledge-based AI approach can be added to the deep learning process. Dental clinicians classify implant systems using a top-down method by recognizing the distinguishing features of the implant shown in the radiograph, such as thread shape, and adding other features to identify the fittest candidate; this logic was expected to be conjugated to the AI classification hierarchy. Therefore, this study was a trial application to construct a model that distinguishes the critical features of an implant system, and in further studies, synthesizes the features to predict a specific implant system or at least present a list of possible systems, rather than specifically predicting all known implant systems. By using this logic, the model is not required to learn every single implant system available in the market; rather, through a less exhausting learning procedure of certain features conjugated with a premade backup of a database of possible implant systems, it can predict the most eligible candidate in a clinician's decision-making process.

The structure of a deep CNN algorithm is critical for achieving superior AI performance. The pre-trained deep learning models used in this study were Xception, ResNet50V2, ResNet101V2, and InceptionV3, all of which were provided by Keras. Pre-trained models have the advantage of requiring less training time and data.¹⁵ However, reports suggest that pre-trained models accelerate convergence

early in training but do not necessarily provide regularization or improve the final target task accuracy.¹⁶ All four models showed an accuracy of over 92%, indicating that performance may be improved by fine-tuning and additional transfer learning techniques.¹⁷

The size of the model varied depending on the number of nodes and weights of each model, and the accuracy varied depending on the configuration of the layers. These four models are in the size range of 88–171 MB but have been reported to show over 0.93 in top-5 accuracy.¹⁸ Owing to limitations in computing power, these four models were chosen among many others for this study. According to Sunny et al., insufficient computational resources in data processing have become obstacles, constraining the efficiency of AI.¹⁹ Upgrades in computing power are expected to result in accelerated AI performance.

The number of images prepared in this study was more than 10,000 for each group after data augmentation, and a 10-fold cross-validation was applied because of the limited dataset; however, the quantity of the training dataset used in this study may still be considered insufficient. Although deep CNN architectures are highly useful in cases where the available training sets are limited, small datasets can be a bottleneck for the further advancement of computer-aided detection.²⁰ For high AI performance, the use of a high-quality training dataset is crucial, and this can be achieved by implementing high-quality, well-annotated datasets.

Image filtering is implemented to increase the visibility of images used for medical diagnosis. Attia et al. compared the effect on prediction accuracy of three image filtering methods: Imadjust, Histogram Equalization, and Contrast Limited Adaptive Histogram Equalization (CLAHE). When image enhancement was applied to a chest X-ray and compared using image quality measurement techniques, such as Mean Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR), CLAHE proved to be the most effective.²¹ Thus, CLAHE was chosen as the filtering method for this study. This is expected to compensate for the limitations caused by downscaling the resolution of radiographic images to some extent, since downscaling was unavoidable due to limitations in computational costs and operating space.²² However, other methods can be expected to be applied in future studies for enhanced results.

Overcoming the black-box nature of deep learning algorithms and ensuring reasonable interpretation to reflect the rationale behind predictions is a challenge for the application of AI in dentistry. Class activation maps highlight the areas focused on prediction, strengthening the model's reliability through visualization capacity. Zhou et al. reported that adding a Global Average Pooling (GAP) layer to the last layer of a neural network instead of a flattening layer can build a generic localizable deep representation that exposes the implicit attention of CNNs on an image.¹⁰ Most CAMs highlight the implant surface where threads are present; however, some localize the healing abutment or connection area instead of the threaded portion of the implant. This is presumably because when the CNN clusters implant types according to the thread shape, it focuses on areas other than the implant surface, implying overfitting of

the model. More data is required to ensure increased accuracy and transparency.

Among the four models, ResNet101V2 exhibited the highest accuracy in the validation dataset, while all models showed lower accuracy at the anterior than at the posterior sites. This could be the result of the overlap of various structures in the anterior area, along with the increased blurring and magnification of this area, which is often observed in panoramic images. Fluctuations are observed in the loss plot during the learning process. This, in addition to the unfavorable localization of CAMs mentioned above, reflects the likelihood of overfitting. Overfitting in this study can be explained by insufficient data and image distortion. Data shortages occurred because the types of implants were biased towards a particular model. Clinicians show preferences in selecting implants;²³ thus, it is difficult to collect an equal number of cases from various systems. Among the eight types of implant systems investigated in this study, dentin implants accounted for 35% of the total number of implants. The image distortion in panoramic radiographs was greater than that in periapical images. Therefore, the thread shape of the implant is more likely to be distorted, making it difficult for the AI to recognize. The distortion rate was highest in the maxillary incisor area and lowest in the mandibular second premolar and first molar areas, regardless of the patient's head position.²⁴ This coincides with the accuracy according to the implant site, and the lower accuracy rate in the anterior portion is expected to contribute to overfitting.

This study has notable limitations. First, the sample size used in this study was relatively small compared with other CNN-related studies. This is because the radiographic images were obtained from patients who underwent implant surgery at a single institution; resulting in a limited number of patients enrolled for data collection. The number of images prepared in this study was multiplied through data augmentation, and 10-fold cross-validation was used; however, the quantity of the training dataset used in this study may still be considered insufficient. Although deep CNN algorithms with pre-trained weights are highly useful in cases where available training sets are limited, small datasets can be a bottleneck for the further advancement of computer-aided detection.²⁰ For high AI performance, the use of a high-quality training dataset is crucial, which can be achieved by implementing high-quality, well-annotated datasets. Second, data homogeneity was present because the data were collected from a single institution. Although pre-trained deep CNN models using large image datasets are effective for general image classification because of their pre-trained weights, the possibility of overfitting cannot be excluded when using a small dataset with a singular character originating from a limited data source. External validation is required to confirm that multicenter studies are necessary to increase the classification accuracy. With further improvements made using additional refined data, the fine-tuning of the models can be expected to be optimized to distinguish among the various implants used worldwide. Finally, among the various implant design features, only the thread shape was considered; this was further divided into three categories. However, other features of implant design can be subdivided

into numerous other categories. Implant systems vary in fixture shape (straight, tapered, or conical), emergence, screw, or abutment design. Such variables can be implemented to advance the accurate classification and proposal of system candidates.

V. Conclusion

In conclusion, the CNN models used in this study could accurately classify implants in all areas of the oral cavity according to the thread shape using radiographic images. All models exhibited higher performance in the posterior area, and the overall performance was the highest when using ResNet101V2. These models can support automated implant classification in educational fields and clinics.

References

- Adell R, Eriksson B, Lekholm U, Brånemark PI, Jemt T. A long-term follow-up study of osseointegrated implants in the treatment of totally edentulous jaws. Int J Oral Maxillofac Implants 1990;5:347-59.
- Jung RE, Zembic A, Pjetursson BE, Zwahlen M, Thoma DS. Systematic review of the survival rate and the incidence of biological, technical, and aesthetic complications of single crowns on implants reported in longitudinal studies with a mean follow-up of 5 years. Clin Oral Implants Res 2012;23:2-21.
- Jokstad A, Ganeles J. Systematic review of clinical and patient-reported outcomes following oral rehabilitation on dental implants with a tapered compared to a non-tapered implant design. Clin Oral Implants Res 2018;29:41-54.
- 4. Steigenga JT, Al-Shammari KF, Nociti FH, Misch CE, Wang HL. Dental implant design and its relationship to long-term implant success. Implant Dent 2003;12:306-17.
- 5. Hwang JJ, Jung YH, Cho BH, Heo MS. An overview of deep learning in the field of dentistry. Imaging Sci Dent 2019;49:1-7.
- 6. Shan T, Tay F, Gu L. Application of artificial intelligence in dentistry. J Dent Res 2021;100: 232-44.
- 7. Lim HK, Kwon YJ, Lee ES. Application of artificial intelligence in identification of dental implants system: literature review. J Dent Implant Res 2020;39:48-52.
- Manikyamba YJB, Sajjan SMC, Raju RAV, Rao B, Nair CK. Implant thread designs: An overview. Trends Prosthodont Dent Implantol 2017;8:11-9.
- 9. Rodriguez JD, Perez A, Lozano JA. Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE Trans Pattern Anal Mach Intell 2009;32:569-75.
- Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016. p. 2921-9.

www.kci.go.

- 11. Kim JE, Nam NE, Shim JS, Jung YH, Cho BH, Hwang JJ. Transfer learning via deep neural networks for implant fixture system classification using periapical radiographs. J Clin Med 2020;9:1117.
- 12. Sukegawa S, Yoshii K, Hara T, Yamashita K, Nakano K, Yamamoto N, et al. Deep neural networks for dental implant system classification. Biomolecules 2020;10:984.
- Montani S, Striani M. Artificial intelligence in clinical decision support: a focused literature survey. Yearb Med Inform 2019;28:120-7.
- 14. Magrabi F, Ammenwerth E, McNair JB, De Keizer NF, Hyppönen H, Nykänen P, et al. Artificial Intelligence in Clinical Decision Support: Challenges for Evaluating AI and Practical Implications: A Position Paper from the IMIA Technology Assessment & Quality Development in Health Informatics Working Group and the EFMI Working Group for Assessment of Health Information Systems. Yearb Med Inform 2019;28:128-34.
- 15. Simon M, Rodner E, Denzler J. Imagenet pre-trained models with batch normalization. arXiv preprint arXiv: 1612.01452, 2016.
- 16. He K, Girshick R, Dollár P. Rethinking imagenet pre-training. Proceedings of the IEEE International Conference on Computer Vision. 2019; 4918-27.
- 17. Keskar NS, Mudigere D, Nocedal J, Smelyanskiy M, Tang PTP. On large-batch training for deep learning: Generalization gap and sharp minima. arXiv preprint arXiv: 1609.04836, 2016.
- 18. Keras Applications.
- Sunny S, Baby A, James BL, Balaji D, NV A, Rana MH, et al. A smart tele-cytology point-of-care platform for oral cancer screening. PLoS ONE 2019;14:e0224885.
- Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 2016;35:1285-98.
- 21. Attia SJ. Enhancement of chest X-ray Images for diagnosis purposes. J Nat Sci Res 2016;6:43-6.
- 22. Jadoon MM, Zhang Q, Haq IU, Butt S, Jadoon A. Three-class mammogram classification based on descriptive CNN features. Biomed Res Int 2017;2017:3640901.
- 23. Al-Wahadni A, Barakat MS, Abu Afifeh K, Khader Y. Dentists' most common practices when selecting an implant system. J Prosthodont 2018;27:250-9.
- 24. Zarch SH, Bagherpour A, Langaroodi AJ, Yazdi AA, Safaei A. Evaluation of the accuracy of panoramic radiography in linear measurements of the jaws. Iran J Radiol 2011;8:97-102.

