J Periodontal Implant Sci. 2023 Feb;53(1):38-53
https://doi.org/10.5051/jpis.2201060053
pISSN 2093-2278·eISSN 2093-2286

JPIS
Journal of Periodontal & Implant Science

Research Article
Periodontal Science

Check for updates

Determination of the stage and grade of periodontitis according to the current classification of periodontal and peri-implant diseases and conditions (2018) using machine learning algorithms

Kübra Ertaş [1], Ihsan Pence [2], Melike Siseci Cesmeli [2], Zuhal Yetkin Ay [1,*]

[1]Department of Periodontology, Faculty of Dentistry, Süleyman Demirel University, Isparta, Turkey
[2]Department of Software Engineering, Bucak Faculty of Technology, Burdur Mehmet Akif Ersoy University, Burdur, Turkey

OPEN ACCESS

Received: Feb 18, 2022
Revised: Jun 23, 2022
Accepted: Jul 26, 2022
Published online: Sep 6, 2022

***Correspondence:**
Zuhal Yetkin Ay
Department of Periodontology, Faculty of Dentistry, Süleyman Demirel University, Çünür, Isparta 32260, Turkey.
Email: zuhalyetkin@yahoo.com
Tel: +90 505 2642564
Fax: +90 246 2370607

ORCID iDs
Kübra Ertaş
https://orcid.org/0000-0001-9605-3462
Ihsan Pence
https://orcid.org/0000-0003-0734-3869
Melike Siseci Cesmeli
https://orcid.org/0000-0001-9541-2590
Zuhal Yetkin Ay
https://orcid.org/0000-0003-4716-9404

ABSTRACT

Purpose: The current Classification of Periodontal and Peri-Implant Diseases and Conditions, published and disseminated in 2018, involves some difficulties and causes diagnostic conflicts due to its criteria, especially for inexperienced clinicians. The aim of this study was to design a decision system based on machine learning algorithms by using clinical measurements and radiographic images in order to determine and facilitate the staging and grading of periodontitis.
Methods: In the first part of this study, machine learning models were created using the Python programming language based on clinical data from 144 individuals who presented to the Department of Periodontology, Faculty of Dentistry, Süleyman Demirel University. In the second part, panoramic radiographic images were processed and classification was carried out with deep learning algorithms.
Results: Using clinical data, the accuracy of staging with the tree algorithm reached 97.2%, while the random forest and k-nearest neighbor algorithms reached 98.6% accuracy. The best staging accuracy for processing panoramic radiographic images was provided by a hybrid network model algorithm combining the proposed ResNet50 architecture and the support vector machine algorithm. For this, the images were preprocessed, and high success was obtained, with a classification accuracy of 88.2% for staging. However, in general, it was observed that the radiographic images provided a low level of success, in terms of accuracy, for modeling the grading of periodontitis.
Conclusions: The machine learning-based decision system presented herein can facilitate periodontal diagnoses despite its current limitations. Further studies are planned to optimize the algorithm and improve the results.

Keywords: Classification; Deep learning; Machine learning; Periodontitis

Conflict of Interest
No potential conflict of interest relevant to this article was reported.

Author Contributions
Conceptualization: Ihsan Pence, Zuhal Yetkin Ay. Formal Analysis: Ihsan Pence, Melike Siseci Cesmeli, Zuhal Yetkin Ay. Investigation: Kübra Ertaş, Zuhal Yetkin Ay. Methodology: Kübra Ertaş, Zuhal Yetkin Ay. Writing – Original Draft: Kübra Ertaş, Ihsan Pence, Melike Siseci Cesmeli, Zuhal Yetkin Ay Writing – Review & Editing: Kübra Ertaş, Ihsan Pence, Melike Siseci Cesmeli, Zuhal Yetkin Ay.

INTRODUCTION

The classification of periodontal and peri-implant diseases and conditions, which are very common worldwide, is necessary and useful for clinicians to diagnose and appropriately treat these diseases [1]. In the past, many classifications were created for periodontal diseases and conditions, but in 2017, the Classification of Periodontal and Peri-implant Diseases and Conditions was created by the American Academy of Periodontology and the European Federation of Periodontology in the World Workshop. This new classification, which was published and disseminated in 2018, created a single periodontitis group that includes staging and grading dimensions, requiring detailed clinical and radiographic examinations [2]. However, the staging and grading procedures for periodontitis have not yet been fully adopted because it was not found practical by many clinicians for rapid diagnoses in daily practice [3].

Periodontal probes and panoramic/periapical radiographs are still widely used as diagnostic tools for periodontitis and as prediction methods in traditional periodontal examinations. Computer-assisted diagnosis, which is a new field of research, has developed very rapidly in the field of dentistry and has provided impressive results in terms of diagnosis and prediction in radiological and pathological studies [4]. Artificial intelligence (AI) and its branches are now widely used in many fields of dentistry, especially for the processing of radiographic images (periodontal bone loss, maxillary sinusitis, osteoporosis detection, etc.) [5,6].

Deep convolutional neural networks (CNNs), a sub-branch of AI, are a xlink:type of machine learning used in various fields, especially for image and voice recognition. CNNs consist of convolutional layers, a pooling layer, and a fully connected layer that mimics the connectivity patterns of neurons in the animal visual cortex. Unlike traditional image classification algorithms and other deep learning algorithms, deep CNNs can also learn the filters used on images to soften handcrafted images or highlight the edges [5]. Like other branches of AI, deep CNNs have been used to process large and complex images [7]. Nonetheless, there are a limited number of studies in the literature that have used AI algorithms to perform diagnoses according to the 2018 Classification of Periodontal and Peri-Implant Diseases and Conditions [8].

This study aimed to perform the staging and grading classification of periodontitis with machine learning algorithms and a deep CNN design, using both clinical attributes and radiographic images. Thus, it is hoped that this computer-aided algorithm will be used to diagnose patients in daily practice and facilitate the utilization of the new classification in periodontal epidemiological survey studies, especially for inexperienced clinicians.

MATERIALS AND METHODS

This study analyzed clinical and radiographic data from 280 patients who visited the Department of Periodontology, Faculty of Dentistry, Süleyman Demirel University and Geriatrics Clinic between 2017 and 2019 taken from the archive with permission and approval from the Süleyman Demirel University Faculty of Medicine Clinical Research Ethics Committee (decision number 29/410, dated 30/12/2020). This study was conducted in accordance with the 1975 Declaration of Helsinki, as revised in 2013.

JPIS Journal of Periodontal & Implant Science

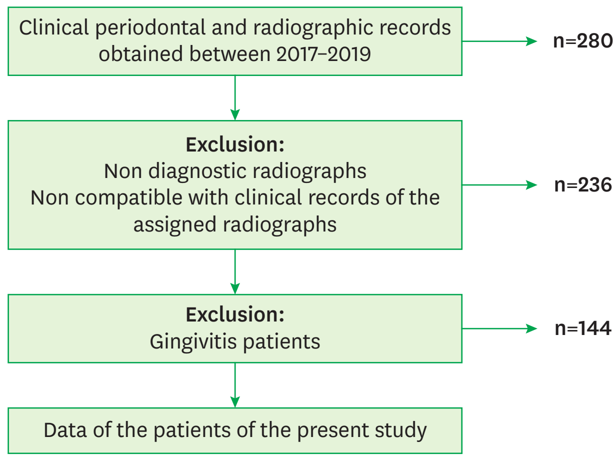

Figure 1. Diagram of the data allocation from the main data obtained between 2017 and 2019.

The inclusion criteria were the availability of periodontal recordings, including the attributes used for the staging and grading of periodontitis according to the current classification, and orthopantomographs with diagnostic criteria.

Patients whose orthopantomographs did not meet the diagnostic criteria and whose clinical records were not compatible with their orthopantomographs were excluded from the study (n=44). Of the 280 individuals whose records were screened, 236 were included in the study, and the data of 144 individuals with periodontitis were used in accordance with the aim of the present study (**Figure 1**). In the study, all clinical and radiographic data that had been previously evaluated according to the earlier classification [9] were re-evaluated, and the staging and grading of periodontitis were performed in accordance with the current classification as detailed below [10].

Before staging and grading, the diagnosis of periodontitis should be conducted using the following criteria: interdental clinical attachment loss (CAL) at ≥2 non-adjacent teeth, or buccal/oral CAL ≥3 mm with a probing depth (PD) >3 mm at ≥2 teeth, and the observed CAL should not be ascribed to non-periodontal causes [11].

A multidimensional evaluation of periodontitis was performed with staging and grading according to the current classification. Staging aims to determine the extent and severity of periodontitis and the complexity of its management based on the amount of periodontitis-induced tissue destruction and specific factors, the interdental CAL at the site of greatest loss, radiographic bone loss (expressed as a percentage), tooth loss due to periodontitis, PD, bone loss pattern (horizontal/vertical), furcation involvement, ridge defects, and the need for complex rehabilitation due to masticatory dysfunction, secondary occlusal trauma, bite collapse, drifting, or flaring [11].

Grading aims to determine the rate of disease progression and the response to standard periodontal therapy using direct and indirect evidence, such as radiographic bone loss or CAL over 5 years, the percentage of bone loss/age, biofilm deposits (in particular, whether they are commensurate with the levels of destruction), and the presence of risk factors such as diabetes and/or smoking [11].

Staging may be summarized as follows [11].

Stage I: PD ≤4 mm, CAL ≤1–2 mm, horizontal bone loss, and no tooth loss due to periodontitis

Stage II: PD ≤5 mm, CAL ≤3–4 mm, horizontal bone loss, and no tooth loss due to periodontitis

Stage III: PD ≥6 mm, CAL ≥5 mm, and may have vertical bone loss and/or furcation involvement of class II or III, loss of ≤4 teeth due to periodontitis

Stage IV: PD ≥6 mm, CAL ≥5 mm, and may have vertical bone loss and/or furcation involvement of class II or III, <20 teeth may be present, and there is the potential for loss of ≥5 teeth due to periodontitis

Grading may be summarized as follows [11].

Grade A: no bone loss over 5 years, the percentage of bone loss/age <0.25, heavy biofilm deposits with low levels of destruction, non-smokers, and normoglycemic/non-diabetic status

Grade B: bone loss <2 mm over 5 years, percentage of bone loss/age between 0.25 and 1.0 mm, destruction commensurate with biofilm deposits, smoking <10 cigarettes a day, hemoglobin A1c (HbA1c) <7.0% in patients with diabetes

Grade C: bone loss ≥ 2 mm over 5 years, percentage of bone loss/age >1.0 mm, destruction exceeding expectations given biofilm deposits; specific clinical patterns suggestive of periods of rapid progression and/or early-onset disease, smoking ≥10 cigarettes a day, HbA1c ≥7.0% in patients with diabetes

The study consists of 2 parts: in the first part, machine learning algorithms were applied to the clinical and radiological attributes used for staging and grading of periodontitis (the presence of fewer than 20 teeth, percentage of bone loss/age, radiographic bone loss, age, diabetes, smoking, tooth loss, vertical bone loss, furcation problems of grades 2–3, and interdental CAL).

In the second part, the orthopantomographs in JPEG format were used to train a deep CNN, and the aim was to perform the staging and grading of periodontitis only using photographs (the flow of the processes performed in this study is shown in **Supplementary Video 1**). Another aim was to reduce the number of attributes used, as well as to achieve high classification accuracy (CA) in staging and grading with the algorithm.

All orthopantomographs were obtained using the same digital panoramic machine, Promax (Planmeca, Helsinki, Finland), operating with settings of 66 kV, 8 mA, and an exposure time of 16 seconds. Each digital radiograph was exported with a resolution of 96 dpi at a size of approximately 1,100×550 pixels and saved as a JPEG-formatted image file that was coded according to the periodontitis group it belongs. These image files were saved in a way that did not allow access to personal information such as patients' name, sex, or age.

While the orthopantomographs were obtained with the digital panoramic machine, none of the images were preprocessed, the margins were not removed, the tooth region was not focused, and the resolution or size was not changed. However, before the images are given as input to the deep CNN model, some preprocessing steps were applied (**Figure 2**).

Determination of the stage and grade of periodontitis with clinical and radiographic features using machine learning algorithms

The machine learning model that gives the best results was established using various

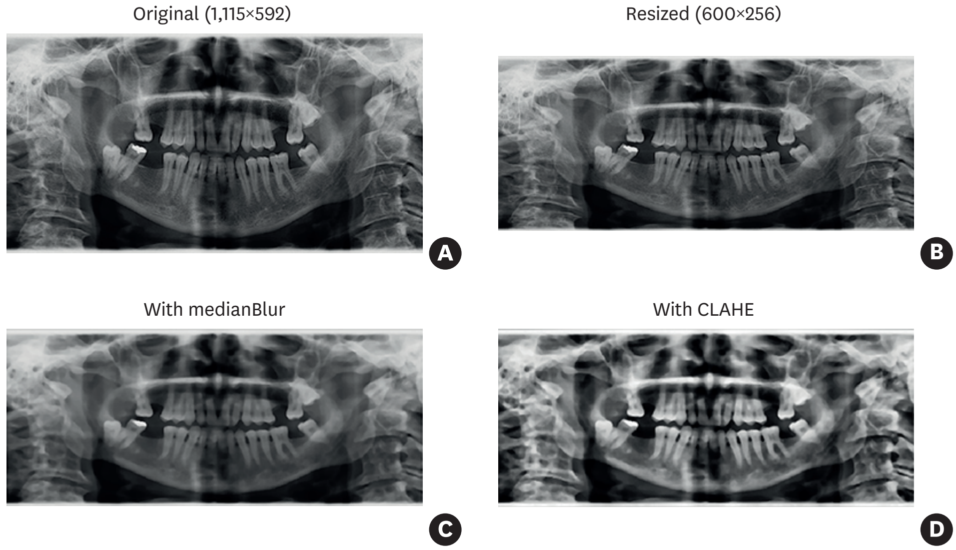

Figure 2. One of the orthopantomographs obtained with the digital panoramic machine and its preprocessing for a deep convolutional neural network.

supervised algorithms in the Python programming language. The algorithms used were k-nearest neighbors (kNN), artificial neural network (ANN), tree, support vector machine (SVM), random forest, naive Bayes, and logistic regression. Machine learning algorithms have hyperparameters with a major impact on model success. A preliminary study was conducted to determine the optimum parameters of these algorithms. Accordingly, the number of neighbors was selected as 5 and the Euclidean distance metric was used in the kNN. As the ANN parameters, the number of iterations was 100, the number of neurons in the hidden layer was 20, the alpha value was 0.005, the activation function was ReLU, and the optimization algorithm was L-BFGS-B. For the tree algorithm, the minimum number of instances in leaves was 1 and the maximum tree depth was selected as 20. For the parameters of the SVM algorithm, the iteration number was determined as 100, the cost value was 6, the epsilon value was 0.1, the tolerance value was 0.005, and the kernel function was determined as a second-degree polynomial. For the random forest algorithm, 20 was determined as the number of trees. For the logistic regression algorithm, the c-value was set as 1. The naive Bayes model, in contrast, has no parameters due to the inherent structure of its algorithm. In order to obtain more accurate/successful results, the training and test sets were separated using leave-one-out cross-validation.

A scoring system of the features was used to determine the weights of attributes. For this, calculations such as the information gain ratio, Gini ratio, and chi-square statistic were used. The performance values were examined by ordering the attributes according to the highest information gain ratio and trying the pairs of values from the highest to the lowest.

Determination of periodontitis stage and grade with image processing

In this second stage of the experimental analysis in this study, an attempt was made to perform classification (staging and grading) with deep learning algorithms only on photographs by using panoramic images from 144 patients. Transfer learning was used in the deep learning phase, with the DenseNet121, EfficientNetB0, InceptionV3, ResNet50, and VGG16 networks. In order to train the deep learning algorithms, the 2021.1.2 version of the PyCharm Editor, which uses the Python language, and libraries such as TensorFlow, Keras, and Sklearn were used. Ten-fold cross-validation was used on the photographs obtained from panoramic radiographs due to the large image sizes and the long training period.

In the study, the effect of preprocessing on success was also examined using the medianBlur and CLAHE methods. In the preprocessing stage, all the images were first resized to 600×256. Noise reduction was then applied with medianBlur with a 5×5 kernel size, followed by the application of CLAHE, which provides contrast amplification close to a certain pixel value, with a grid size of 8×8 and a clipping limit of 2.0.

In addition, all images were normalized to values between 0 and 1. For transfer learning, the original weights of the relevant networks were not trained, only the weights of the output layer. In this way, deep learning algorithms were also used to create features. In addition to deep learning algorithms, hybrid models were also created with various machine learning algorithms added to the output layer of the relevant networks. Adam optimization was used to train the CNN models with 25 iterations and 0.005 learning rate, while for hybrid models, the linear kernel function is used for SVM, and the previously mentioned parameters were used for the tree and random forest algorithms.

A saliency map was used to measure and visualize the spatial support from a particular class in images. A Grad-CAM based saliency map was used in this study.

Performance evaluation metrics

In machine learning, different evaluation metrics are applied according to the xlink:type of problem. Accuracy, precision, recall, the F1-score, the receiver operating characteristic (ROC) curve, and the area under curve (AUC) are used for classification tasks, whereas in regression tasks, metrics such as the mean squared error, root mean squared error, mean absolute error, mean absolute percentage error, and coefficient of determination are used. In segmentation tasks, in addition to classification metrics, the dice score and intersection over union can provide an effective evaluation. As this study was based on a classification task, the evaluation criteria of accuracy, precision, recall, the F1-score, the ROC curve, and the AUC were used to evaluate the classification performance of the models. A confusion matrix was used to calculate these values. The confusion matrix has true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values. The equations for accuracy, recall (TP rate [TPR]), precision, and the F1-score, which are performance evaluation metrics, are given in Equation (1) through Equation (4), respectively. The recall (TPR) and FP rate (FPR) were used in the calculation of AUC, and the equation for the FPR is given in Equation (5).

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

$$\mathrm{TPR} = \frac{TP}{TP + FN}$$

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

$$\mathrm{F1-Score} = \frac{2 * Precision * Recall}{Precision + Recall}$$

$$\mathrm{FPR} = \frac{FP}{TN + FP}$$

https://doi.org/10.5051/jpis.2201060053

JPIS Journal of Periodontal & Implant Science

For successful classification results, the information gain ratio was used to determine which features had a greater impact on the classification. If needed, the existing images were also passed through some preprocessing steps to increase the classification success or to create the final images for the classification.

RESULTS

The age of the population varied between 25 and 82 years (mean age, 62.23±2.82 years), and 117 of the 144 patients were women. The 144 periodontitis patients in our study included 1 with stage I periodontitis, 3 with stage II, 90 with stage III, and 50 with stage IV. Thirty-four patients had diabetes. There were 11 patients who smoked <10 cigarettes/day and 2 patients who smoked ≥10 cigarettes/day.

Table 1. Information gain ratios of the attributes used for periodontitis staging

Attribute	Gain ratio	Gini	χ^2
Fewer than 20 teeth	0.850	0.395	44.591
Percentage of bone loss/age	0.105	0.018	3.222
Radiographic bone loss	0.097	0.010	6.744
Age	0.094	0.058	15.298
Diabetes	0.059	0.021	6.369
Smoking	0.049	0.014	4.124
Tooth loss	0.043	0.015	5.323
Vertical bone loss ≥3 mm	0.039	0.002	3.683
Furcation problem of grade 2–3	0.035	0.004	2.946
Interdental CAL	0.029	0.007	5.090

CAL: clinical attachment loss.

Table 2. CA values of models for the staging and grading of periodontitis with clinical and radiographic attributes

Characteristic	Model	AUC	CA	F1	Precision	Recall
Staging	kNN	0.760	0.646	0.625	0.616	0.646
	ANN	0.952	0.910	0.908	0.907	0.910
	Tree	0.947	**0.972**	0.969	0.967	0.972
	SVM	0.955	0.944	0.942	0.941	0.944
	Random Forest	0.975	0.965	0.962	0.959	0.965
	Naive Bayes	0.914	0.486	0.648	0.972	0.486
	Logistic regression	0.962	0.944	0.931	0.919	0.944
Grading (n=10)	kNN	0.930	0.868	0.849	0.834	0.868
	ANN	0.979	0.972	0.972	0.972	0.972
	Tree	0.928	0.965	0.965	0.966	0.965
	SVM	0.975	0.965	0.965	0.965	0.965
	Random Forest	0.996	**0.979**	0.979	0.979	0.979
	Naive Bayes	0.966	0.896	0.898	0.906	0.896
	Logistic regression	0.989	0.910	0.887	0.866	0.910
Grading (n=4)	kNN	0.989	**0.986**	0.986	0.987	0.986
	ANN	0.991	0.972	0.972	0.972	0.972
	Tree	0.928	0.965	0.965	0.966	0.965
	SVM	0.990	0.979	0.979	0.979	0.979
	Random Forest	0.989	**0.986**	0.986	0.987	0.986
	Naive Bayes	0.989	0.979	0.979	0.979	0.979
	Logistic regression	0.990	0.938	0.914	0.892	0.938

Bold values indicate the highest CA values for each characteristic.
kNN: k-nearest neighbor, ANN: artificial neural network, SVM: support vector machine, AUC: area under the curve, CA: classification accuracy.

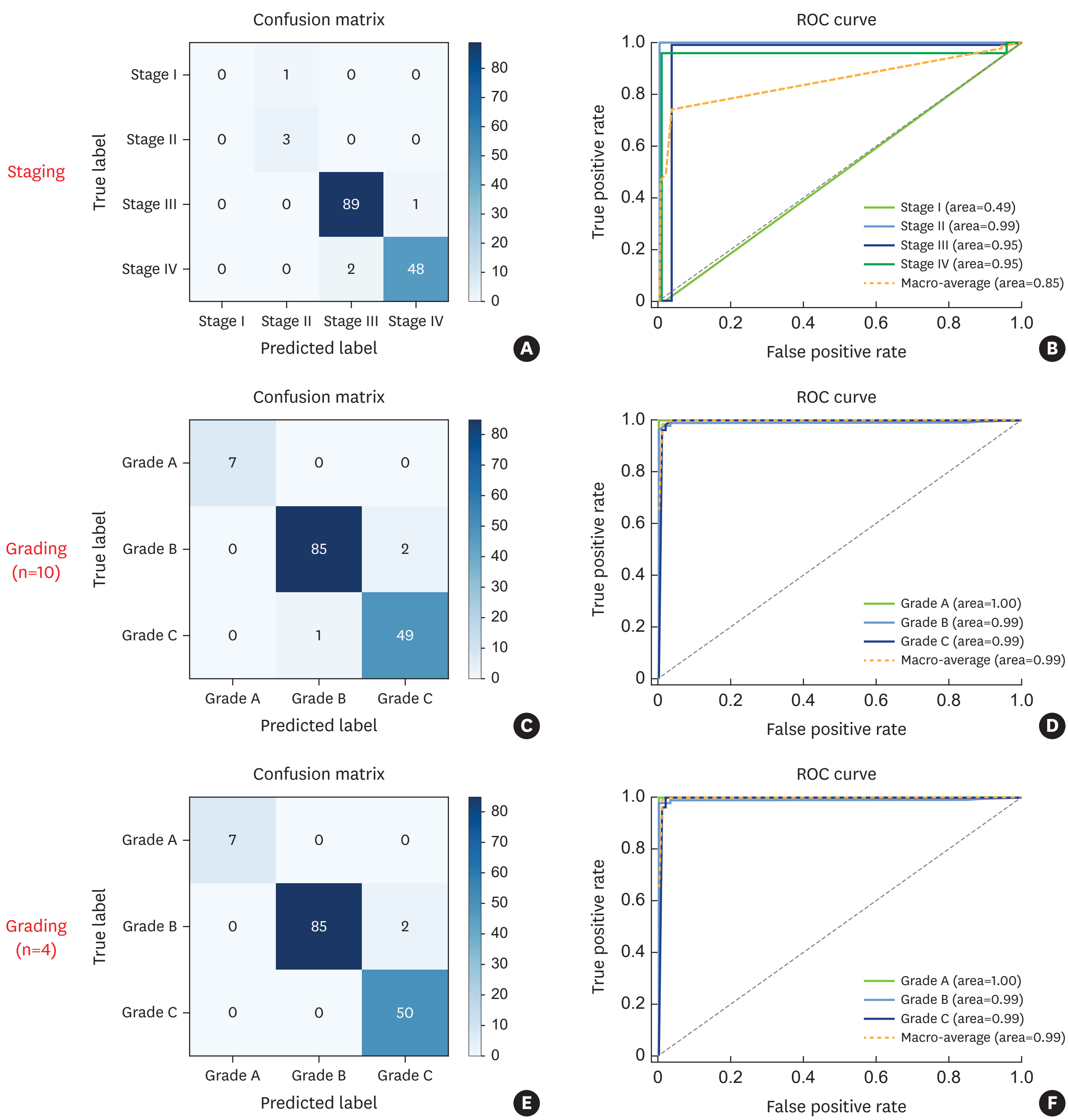

Figure 3. The tree algorithm for staging: (A) confusion matrix, (B) ROC curve. The Random Forest algorithm for grading (n=10): (C) confusion matrix, (D) ROC curve. The Random Forest algorithm for grading (n=4): (E) confusion matrix, (F) ROC curve.
ROC: receiver operating characteristic.

Determination of periodontitis stages and grades with clinical data using machine learning algorithms

Determination of the stage of periodontitis

The attributes in the dataset for determining the stage of periodontitis, as shown in **Table 1**, were used. The method providing the best CA (0.972) was the tree algorithm, and all scores are given in **Table 2**. The confusion matrix for this algorithm and the ROC curves are given in **Figure 3A and B**.

Determination of attribute weights and selection of attributes for periodontitis staging

The attribute weights and information gain ratios for periodontitis staging are given in **Table 1**. According to the highest information gain ratio for the random forest algorithm, which provided the second-best value with 10 attributes, the classification success rates did not change when the first 4 attributes (fewer than 20 teeth, percentage of bone loss/age, radiographic bone loss and age) were used. These first 4 attributes achieved the same success in the staging of periodontitis.

Determination of the grade of periodontitis

Using the attributes in **Table 3**, and it was determined that the method giving the best CA (0.979) was the random forest algorithm for 10 attributes, while the kNN and random forest yielded a CA of 0.986 for the first 4 attributes (**Table 2**). The confusion matrix and ROC curve of the random forest algorithm for 10 attributes are given in **Figure 3C and D**.

Determination of attribute weights and attribute selection for periodontitis grading

The attribute weights and information gain ratios for periodontitis grading are given in **Table 3**. When only these first 4 attributes (percentage of bone loss/age, diabetes, radiographic bone loss and smoking) were used, the CA for periodontitis grading increased to 0.986 for the random forest algorithm. The confusion matrix and ROC curves for this algorithm are given in **Figure 3E and F**.

Determination of periodontitis stage and grade with image processing

Determination of periodontitis stage

The DenseNet121, EfficientNetB0, InceptionV3, ResNet50, and VGG16 networks were used to determine the stage of periodontitis through image processing. Each of these networks was evaluated by creating hybrid models with the machine learning algorithms (tree, SVM, and random forest), and the CA results for each network are presented in **Table 4** using the highest values of their hybrid models, according to their preprocessed and unprocessed states. In the study, these models were preferred because they gave much better results than any of the other network models and hybrid methods with different parameters, or because they are more popular in the literature. Results containing all combinations of classical and hybrid models are available in the **Supplementary Table 1**.

The EfficientNetB0 + SVM hybrid model provided a staging CA value of 0.861 for images that were not preprocessed, and the confusion matrix and ROC curve for this algorithm are shown

Table 3. Information gain ratios of the attributes used for periodontitis grading

Attribute	Gain ratio	Gini	χ^2
Percentage of bone loss/age	0.558	0.205	13.422
Diabetes	0.491	0.175	48.880
Radiographic bone loss	0.261	0.077	14.328
Smoking	0.179	0.047	30.665
Vertical bone loss ≥3 mm	0.100	0.020	9.935
Furcation problem of grade 2–3	0.066	0.009	5.550
Interdental CAL	0.057	0.031	13.968
Tooth loss	0.041	0.011	5.590
Age	0.037	0.027	3.489
Fewer than 20 teeth	0.007	0.003	0.420

CAL: clinical attachment loss.

https://doi.org/10.5051/jpis.2201060053

JPIS Journal of Periodontal & Implant Science

Table 4. Highest CA values of the DenseNet121, EfficientNetB0, InceptionV3, ResNet50, VGG16, and hybrid models in staging periodontitis with image processing

Characteristic	Model	Pre-processing	AUC	CA	F1	Precision	Recall
Staging	DenseNet121 + SVM	+	0.761	0.854	0.841	0.830	0.854
		−	0.707	0.799	0.786	0.775	0.799
	EfficientNetB0 + SVM	+	0.796	0.814	0.800	0.789	0.813
		−	0.774	0.861	0.849	0.837	0.861
	InceptionV3 + Random Forest	+	0.792	0.806	0.789	0.785	0.806
		−	0.708	0.728	0.713	0.703	0.729
	ResNet50 + SVM	+	0.799	**0.882**	0.872	0.864	0.882
		−	0.742	0.860	0.852	0.843	0.861
	VGG16	+	0.574	0.798	0.784	0.776	0.799
		−	0.610	0.827	0.811	0.806	0.826

Bold value indicates the highest CA value.
CA: classification accuracy, SVM: support vector machine, AUC: area under the curve.

in **Figure 4A and B**. The ResNet50 + SVM model, which is a hybrid algorithm, gave the best periodontitis stage CA value, with a result of 0.882 for preprocessed images. The confusion matrix and ROC curve for this algorithm are shown in **Figure 4C and D**.

The hybrid ResNet50 + SVM architecture, which gave the best CA for staging periodontitis, and the saliency maps of some correctly classified images are shown in **Figure 5**. In the saliency maps, the red color highlights the class-related activation region predicted by the classifier.

Determination of the grade of periodontitis

DenseNet121, EfficientNetB0, InceptionV3, ResNet50, and VGG16 networks, which are deep learning models, were used to determine the periodontitis grade after image processing. Each of these networks was also evaluated by creating hybrid models with the machine learning algorithms (tree, SVM, and random forest). The CA results were compared between the hybrid models and each network itself, according to their preprocessed and unprocessed states, and an attempt was made to determine the best model.

VGG16 + SVM, which is a hybrid algorithm, had the highest CA for periodontitis grade using preprocessed images (0.645), and the confusion matrix and ROC curve for this algorithm are shown in **Figure 4E and F**.

DISCUSSION

The aim of this study was to facilitate the staging and grading of periodontitis according to the new classification [12] with clinical and radiographic attributes and images. This was achieved by developing machine learning algorithms and a deep CNN-based decision system. With the developed algorithms, the hypothesis that machine learning can facilitate the usage of the current classification system with clinical data and image processing, especially for inexperienced clinicians, was generally confirmed.

Periodontal diseases have been classified in many ways from past to present, and finally, the current Classification of Periodontal and Peri-implant Diseases and Conditions was prepared in the World Workshop jointly organized by the American Academy of Periodontology and the European Federation of Periodontology in 2017, and it was published and put into practice in 2018 [12]. However, clinicians have had difficulties adopting and applying this classification in daily practice, and many clinicians have complained about difficulties in

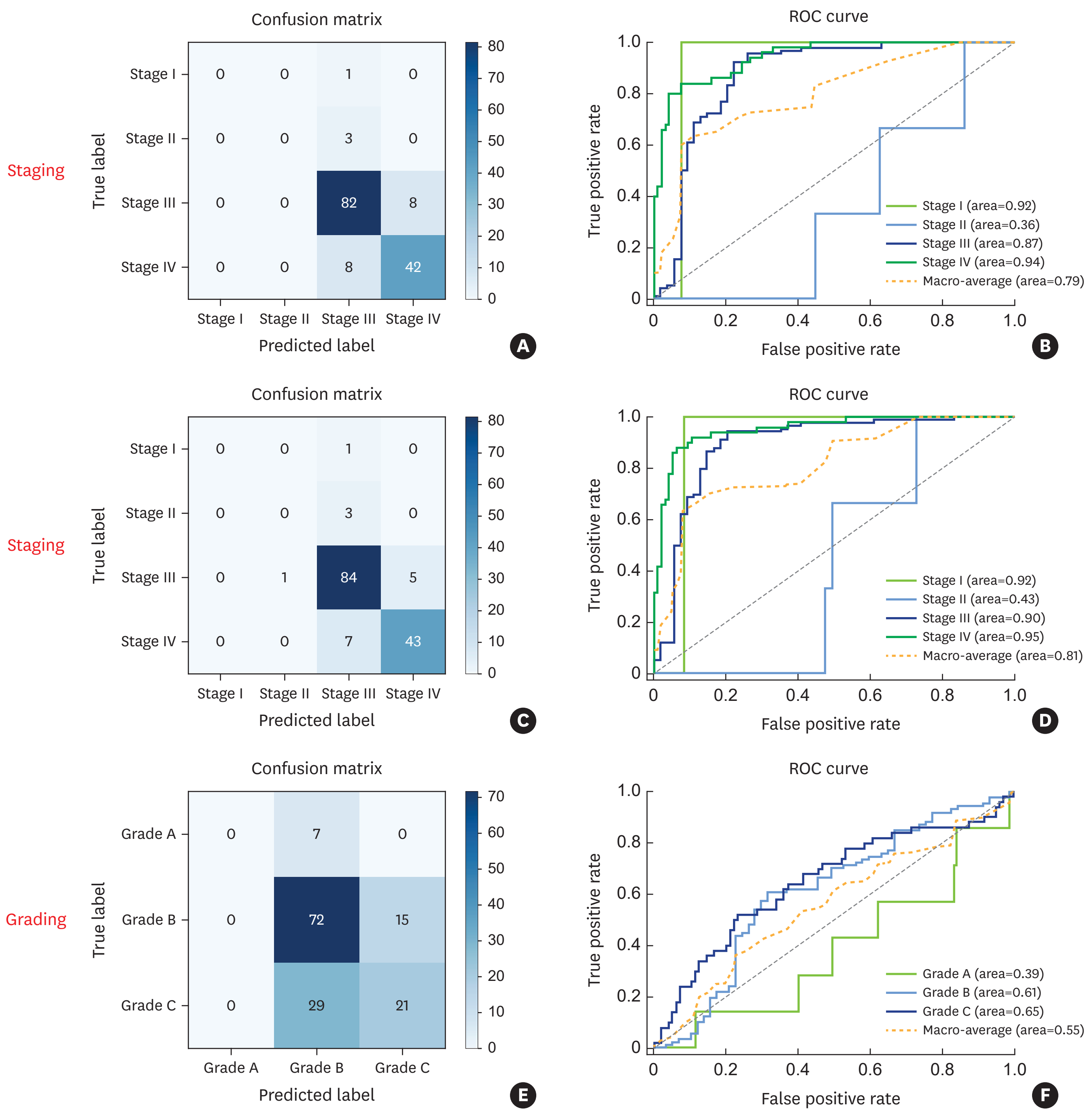

Figure 4. The EfficientNetB0 + SVM hybrid model for staging: (A) confusion matrix, (B) ROC curve. The ResNet50 + SVM hybrid model for staging: (C) confusion matrix, (D) ROC curve. The VGG16 + SVM hybrid model for grading: (E) confusion matrix, (F) ROC curve.
ROC: receiver operating characteristic.

determining the stage and grade of periodontitis because of the presence of many clinical and radiographic factors that need to be considered in the current classification and in periodontal screening studies [3]. To overcome these problems, simple and rapid decision flowcharts have been developed and proposed not only to facilitate the performance of fast and accurate periodontitis staging and grading, but also to minimize confusion and inconsistent diagnoses [13].

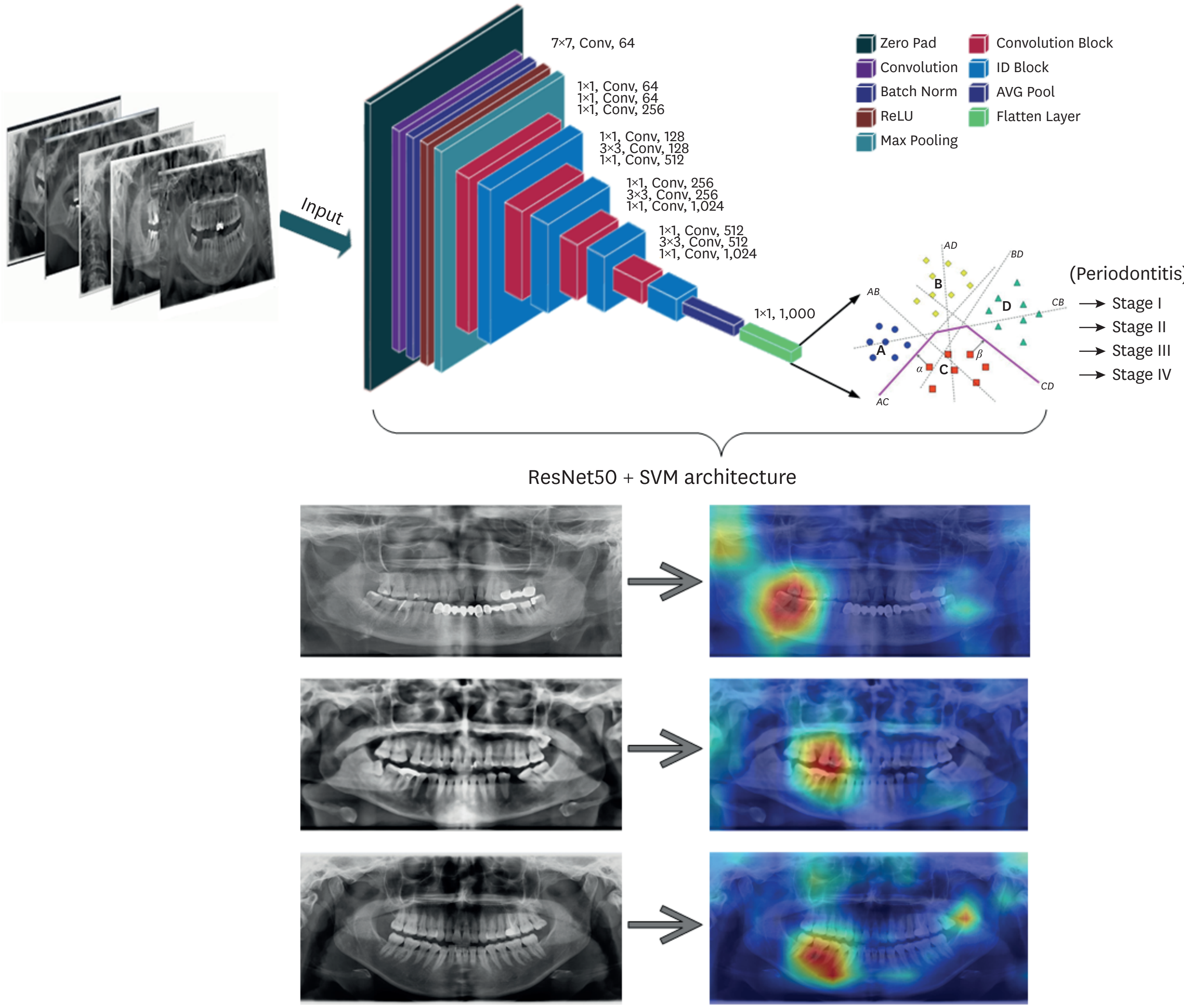

Figure 5. Hybrid ResNet50 + SVM architecture giving the best periodontitis staging classification accuracy and saliency maps of some correctly classified images.
SVM: support vector machine.

In our study, machine learning algorithms were applied to the attributes used in the staging and grading of periodontitis in the new classification. In the algorithms the aim was to use the least number of attributes while keeping the periodontitis staging and grading accuracy rate the highest. A high CA (97.2%) was obtained with the tree algorithm by using 10 attributes to determine the stage of periodontitis. In addition, in the random forest algorithm, which gave the second-best value, when 4 attributes (presence of fewer than 20 teeth, percentage of bone loss/age, radiographic bone loss, age) were used, as the attributes with the highest information gain ratios, a high value for classification success (96.5%) was achieved. Thus, our goal of using the fewest attributes while keeping the highest possible accuracy rate was realized. When making a diagnosis, clinicians can achieve a high level of classification success by using these 4 attributes instead of 10 attributes. There may be differences in the number of these attributes and the CA values of different attributes if

https://doi.org/10.5051/jpis.2201060053

the proposed algorithm is developed with more data, if new models are developed, and if different attributes are selected after hyperparameter optimization.

The random forest algorithm with 10 attributes provided the best CA value (97.9%) when used to determine the grade of periodontitis. Then, the CA value was increased using the random forest and kNN algorithms for 4 attributes (percentage of bone loss/age, diabetes, radiographic bone loss, smoking) evaluated as part of grading. A CA value of 98.6% was reached. Identifying clinical and radiological attributes and transferring them to the algorithm is a more challenging process than just examining images; therefore, this part of the study has considerable importance for real-time applications as the number of attributes is reduced.

In the second part of our study, the CA values for periodontitis staging and grading were examined in algorithm-based models using images. The hybrid algorithm models DenseNet121 + SVM, EfficientNetB0 + SVM InceptionV3 + Random Forest, ResNet50 + SVM, and VGG16 + SVM had staging CA values for the preprocessed images of 85.4%, 81.4%, 80.6%, 88.2% and 81.3%, respectively. For the unprocessed images, the DenseNet121 + SVM, EfficientNetB0 + SVM InceptionV3 + SVM, ResNet50 + SVM hybrid algorithm models and VGG16 classified the data with the following staging CA values: 79.9%, 86.1%, 79.8%, 86%, and 82.7%, respectively. Preprocessing increased the success rate, and the ResNet50 + SVM algorithm came to the fore with the highest staging CA value. The original version of the VGG16 model had high success for unprocessed images. In general, hybrid methods using the SVM algorithm and image preprocessing increased the success rate. The ResNet50 + SVM hybrid algorithm offered the best solution for the problem in preprocessed images, with a CA value of 88.2%.

According to the findings from the first stage of the study, the attribute "presence of fewer than 20 teeth" in the staging process was effective for the differentiation of stage IV, the attribute "percentage of bone loss/age" was effective for the differentiation of stage I, the attribute "radiographic bone loss" was effective for the differentiation of stage II, and the attribute "age" was effective for the differentiation of stages III and IV. In the grading process, the attribute "percentage of bone loss/age" was effective for the differentiation of grade A, the attribute "diabetes" was effective for the differentiation of grade B, and the attribute "radiographic bone loss" was effective for the differentiation of grade A. The fact that more than half of the study population was 65 years or older (as a dataset of individuals presenting to a geriatric clinic was used) may have made the effect of the age attribute more visible. Effective attributes and CA rates may vary in further studies with more individuals in each stage and grade group.

In general, the periodontitis grading CA values showed low success when compared to the staging CA values in deep CNN algorithms. The VGG16 + SVM hybrid algorithm model gave the highest grading CA value (64.5%) for preprocessed images. Preprocessing the images and adding the SVM algorithm to the hybrid models increased the success rate of the model, as was the case for staging.

In our study, panoramic radiographs were used in deep CNN algorithms with and without preprocessing to determine the stage and grade of periodontitis. The developed algorithms did not require any marking of the images, and it is thought that this aspect of the method could be particularly helpful for dentists in diagnosing the stage and grade of periodontitis on panoramic radiographs.

JPIS Journal of Periodontal & Implant Science

There are several studies in the literature about the application of AI in periodontal diagnoses, such as the evaluation of CA only from panoramic radiographs [14], the identification of teeth without periodontal support and the prediction of tooth extraction using periapical radiographs [5], and the identification of teeth without periodontal support with fast CNN [15]. Studies have evaluated the success of classification regarding gingivitis and localized or generalized aggressive periodontitis according to the previous classification [16,17], the use of clinical immunological parameters and radiographic bone loss data to classify cases with chronic and aggressive periodontitis [18], and the classification of gingival enlargement, gingivitis, and periodontitis using histological photographs [19]. However, the classifications used, as well as the clinical periodontal data and radiographs evaluated in those studies, differ from those in our study. Furthermore, the deep CNN models and success rates have also varied considerably.

In the study of Chang et al. [8], the periodontal bone level, cemento-enamel junction (CEJ), teeth, and implants were marked on panoramic radiographs, and the stage of periodontitis was determined with this marking in the axis of the teeth (segmentation) using deep convolutional hybrid models developed with a modified automatic method; that study is the most similar to ours, using the current classification [10]. After segmentation on panoramic radiographs only, without clinical data/attributes, only staging was performed on teeth with a deep CNN and VGG19 hybrid algorithm model [8]. In our study, staging and grading were performed with algorithm-based models without marking on panoramic radiographs, only with or without preprocessing. Furthermore, the staging and grading in our study were performed using different algorithm models with the clinical and radiographic attributes utilized in the current classification.

Since this was a retrospective archive study, it is subject to some limitations. A previously recorded dataset was used in our study, and these records and periodontal diagnoses were therefore not made according to the current classification; instead, they were made in accordance with the 1999 classification [9]. Furthermore, the cause of tooth loss could not be determined. Therefore, the existing tooth loss could not be considered periodontal, and if there was an indication for extraction of the teeth due to periodontitis in the dentition, the number of missing teeth with periodontal causes was added to the dataset. The presence of bite collapse in patients with severe destruction could not be detected only from panoramic images. In addition, a sufficient number could not be obtained in the dataset used to determine the classification success for stage I and stage II periodontitis with the algorithm.

The fact that this study did not determine the consistency of the algorithm with specialists' diagnoses in real clinical conditions can be considered as another limitation. Such an approach might provide more descriptive and clear findings. However, the absence or low number of stages I and II periodontitis patients and the other limitations listed herein mean that such an evaluation should be postponed until we fully overcome the shortcomings of the algorithm in its current form and further develop the algorithm.

To determine the grade of periodontitis using clinical and radiographic attributes, the HbA1c values of patients with diabetes were also used. Since the actual HbA1c values of the patients could not be obtained, 2 different values were accepted in patients with diabetes. All patients with diabetes were considered to have a value of HbA1c <7% (grade B), or the HbA1c values of all the participants with diabetes were considered to be ≥7% (grade C), but this was not a defining attribute for classification, as all patients with diabetes had the same default value. Therefore, the inability to analyze the actual HbA1c values of patients with diabetes is another

https://doi.org/10.5051/jpis.2201060053

JPIS Journal of Periodontal & Implant Science

limitation of our study. With prospective studies planned, the current HbA1c values of the patients will help elucidate the role of this variable.

In the image evaluation, 3 points (the root tip, alveolar crest, CEJ) were marked in the interproximal area of the tooth with the highest CAL. When the image was not clear or the CEJ could not be identified due to the presence of a filling or caries/prosthetic restoration in the related tooth, the second most severe area was marked. This may have caused a lower or higher stage or grade to be determined. The markings mentioned above were not used, and the images were included in the algorithm with or without preprocessing. Although this may seem like a limitation, the high CA rates obtained by directly applying the algorithm to the images without using any marking method will help inexperienced clinicians to determine the stage and grade in a practical way.

A comparison between the findings of the limited number of studies in the literature and the findings of our study would be challenging, since each algorithm is unique and the attributes and classifications used in diagnosis are also different. However, in this study, machine learning models trained using a certain amount of data provided acceptable and satisfactory results for determining the staging and grading of periodontitis according to the current classification. Prospective studies are planned to optimize the algorithm and expand them to include the stages with insufficient data.

In the algorithms evaluated in our study, staging and grading success was found to be higher with clinical and radiographic attributes, and the success obtained with the algorithms when only images were used was not as high as when clinical and radiographic data were utilized. The classification success obtained by the best-performing algorithm identified herein will increase further with prospective studies and the inclusion of new image processing algorithms and segmentation.

SUPPLEMENTARY MATERIALS

Supplementary Table 1

Evaluation scores of various classic and hybrid models with different parameters in staging periodontitis with image processing

Click here to view

Supplementary Video 1

Flow of processes performed in this study to determine the staging and grading of periodontitis.

Click here to view

REFERENCES

1. Tonetti MS, Jepsen S, Jin L, Otomo-Corgel J. Impact of the global burden of periodontal diseases on health, nutrition and wellbeing of mankind: a call for global action. J Clin Periodontol 2017;44:456-62. PUBMED | CROSSREF

2. Tonetti MS, Greenwell H, Kornman KS. Staging and grading of periodontitis: framework and proposal of a new classification and case definition. J Clin Periodontol 2018;45 Suppl 20:S149-61.
PUBMED | CROSSREF
3. Sanz M, Tonetti M. Clinical decision tree for staging and grading. J Clin Periodontol 2019;46:S398-405.
4. Chang SW, Abdul-Kareem S, Merican AF, Zain RB. Oral cancer prognosis based on clinicopathologic and genomic markers using a hybrid of feature selection and machine learning methods. BMC Bioinformatics 2013;14:170.
PUBMED | CROSSREF
5. Lee JH, Kim DH, Jeong SN, Choi SH. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. J Periodontal Implant Sci 2018;48:114-23.
PUBMED | CROSSREF
6. Wang CW, Huang CT, Lee JH, Li CH, Chang SW, Siao MJ, et al. A benchmark for comparison of dental radiography analysis algorithms. Med Image Anal 2016;31:63-76.
PUBMED | CROSSREF
7. Rabunal JR, Dorado J. Artificial neural networks in real-life applications. Hershey: IGI Global; 2005.
8. Chang HJ, Lee SJ, Yong TH, Shin NY, Jang BG, Kim JE, et al. Deep learning hybrid method to automatically diagnose periodontal bone loss and stage periodontitis. Sci Rep 2020;10:7531.
PUBMED | CROSSREF
9. Armitage GC. Development of a classification system for periodontal diseases and conditions. Ann Periodontol 1999;4:1-6.
PUBMED | CROSSREF
10. Caton JG, Armitage G, Berglundh T, Chapple IL, Jepsen S, Kornman KS, et al. A new classification scheme for periodontal and peri-implant diseases and conditions - Introduction and key changes from the 1999 classification. J Clin Periodontol 2018;45 Suppl 20:S1-8.
PUBMED | CROSSREF
11. Tonetti MS, Greenwell H, Kornman KS. Staging and grading of periodontitis: Framework and proposal of a new classification and case definition. J Periodontol 2018;89 Suppl 1:S159-72.
PUBMED | CROSSREF
12. Papapanou PN, Sanz M, Buduneli N, Dietrich T, Feres M, Fine DH, et al. Periodontitis: consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. J Clin Periodontol 2018;45 Suppl 20:S162-70.
PUBMED | CROSSREF
13. Sutthiboonyapan P, Wang HL, Charatkulangkun O. Flowcharts for easy periodontal diagnosis based on the 2018 new periodontal classification. Clin Adv Periodontics 2020;10:155-60.
PUBMED | CROSSREF
14. Krois J, Ekert T, Meinhold L, Golla T, Kharbot B, Wittemeier A, et al. Deep learning for the radiographic detection of periodontal bone loss. Sci Rep 2019;9:8495.
PUBMED | CROSSREF
15. Thanathornwong B, Suebnukarn S. Automatic detection of periodontal compromised teeth in digital panoramic radiographs using faster regional convolutional neural networks. Imaging Sci Dent 2020;50:169-74.
PUBMED | CROSSREF
16. Ozden FO, Özgönenel O, Özden B, Aydogdu A. Diagnosis of periodontal diseases using different classification algorithms: a preliminary study. Niger J Clin Pract 2015;18:416-21.
PUBMED | CROSSREF
17. Farhadian M, Shokouhi P, Torkzaban P. A decision support system based on support vector machine for diagnosis of periodontal disease. BMC Res Notes 2020;13:337.
PUBMED | CROSSREF
18. Youssif AA, Gawish AS, Moussa ME. Automated periodontal diseases classification system. Int J Adv Comput Sci Appl 2012;3:40-8.
CROSSREF
19. Papantonopoulos G, Takahashi K, Bountis T, Loos BG. Artificial neural networks for the diagnosis of aggressive periodontitis trained by immunologic parameters. PLoS One 2014;9:e89757.
PUBMED | CROSSREF