

Combining Image Processing and Artificial Intelligence for Dental Image Analysis: Trends, Challenges, and Applications



M. B. H. Moran, M. D. B. Faria, L. F. Bastos, G. A. Giraldi, and A. Conci

1 Introduction

In the last decades, digital image examinations have been introduced in dental practice, and nowadays, they constitute a prevalent tool employed in the diagnosis of oral diseases. Digital sensors have shorter radiation exposure time than analog radiographs. Moreover, digital radiography provides high-quality images. Its uses have been increasing in clinical practice and scientific researches, facilitating the application of computer methods to process and analyze examinations.

The most common examinations in dental practice are the intraoral (periapical, bitewing, and occlusal) and extraoral (especially the panoramic) radiographs and the cone-beam computed tomography. Each one of these types of imaging focuses

M. B. H. Moran (✉)

Instituto de Computação, Universidade Federal Fluminense, Niterói, Brazil

Departamento de Diagnóstico e Cirurgia, Faculdade de Odontologia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil

e-mail: mhernandez@id.uff.br

M. D. B. Faria

Departamento de Diagnóstico e Cirurgia, Faculdade de Odontologia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil

Faculdade de Odontologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

L. F. Bastos

Departamento de Diagnóstico e Cirurgia, Faculdade de Odontologia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil

G. A. Giraldi

Laboratório Nacional de Computação Científica, Petrópolis, Brazil

A. Conci

Instituto de Computação, Universidade Federal Fluminense, Niterói, Brazil

on different anatomical structures and is used for different purposes. Their acquisition processes also differ one from another. One of the topics discussed in this chapter is the principles of the image formation and acquisition processes behind those techniques.

According to Abdalla-Aslan [1], computer methods, especially those that include artificial intelligence (AI), can be used to improve the accuracy and consistency of diagnosis. Various AI solutions for oral radiology have emerged in the last few years. Several previous works present efforts to automate the identification and evaluation of oral diseases in image-based exams, which would reduce possible errors related to experts' subjectivity [1].

Although most works use traditional image processing methods, recently, machine learning algorithms and even convolutional neural networks (CNNs) have shown promising results for this problem. According to Schwendicke et al. [2], dental imaging presents an excellent potential for image processing solutions since diagnostic imaging is an essential part of dentistry. In other words, image evaluation already consists of an important step in the diagnosis of several oral diseases.

In European countries, most of the radiographs acquired consist of oral image examinations [2]. It is estimated that around 250–300 dental images are acquired per 1000 individuals [2]. AI-based solutions tend to be suitable for a wide range of image applications, including oral ones. This chapter also briefly presents their main principles.

The use of AI-based image techniques tends to increase the effectiveness of diagnosis and lower costs by eliminating routine tasks [3]. Consequently, in the last years, the number of works proposing solutions on the use of such techniques has been increasing.

The main areas of dentistry for which AI-based image processing techniques were applied are cariology, endodontics, periodontology, orthodontics, and forensic dentistry [2].

The most popular applications are segmentation, detection, classification, and their combinations [2]. The detection is related to carious lesions. The anatomical structures for classification include teeth, jaw bone, skeletal landmarks, and biofilm classification. Classification considers the endodontic treatment conditions and even their results. The detection and classification evaluate periodontal inflammation, bone loss, and facial features [2].

The tasks involved in those applications are localization and measurement of anatomic structures, diagnosis of osteoporosis, classification and segmentation of maxillofacial cysts or tumors, identification of alveolar bone resorption, classification of periapical lesions, diagnosis of multiple dental diseases, and classification of tooth types [4]. Other applications less explored are identification of root canals, diagnosis of the maxillary sinusitis, identification of inflamed gum, identification of dental plaque, detection of dental caries, and classification of the stages of the lower third molar [4]. This chapter also discusses some of the main application problems for AI-based image processing. To demonstrate the feasibility of using the presented techniques, three of the mentioned applications (identification of periodontal

diseases, detection of dental caries, and radiograph image enhancement) are selected for practical exemplification.

Although there is a great potential for the use of AI-based image processing techniques in dental imaging, there are several challenges to be overcome in this context. Among these are the lack of available data, the subjectivity of oral diseases' diagnosis, the lack of diagnostic standards, the complexity of some oral diseases, and the resistance from dentists to include computational tools in their routine. All these aspects are discussed in this chapter, as well.

2 An Overview on Digital Dental Imaging

This section discusses the principles of image formation and representation for oral radiographs and presents the most common imaging exams used in dentistry.

2.1 X-Ray Images

Medical radiographs consist of a type of biomedical data from particles' interactions in the electromagnetic x-ray spectrum, which consider very short wavelengths. The range that covers x-ray photon energies is delimited from 10 keV (1.6×10^{-15} J) to 100 keV, *i.e.*, wavelengths ranging from 0.124 to 0.0124 nm [5]. Basically, the devices used to obtain radiographic exams are constituted by an x-ray source and a detector, also called receptor (Fig. 1), which can be a film or a digital device.

The visual interpretation of radiographic exams is based on the radiodensity concept. When x-ray photons are irradiated from an x-ray source to an object (or biological tissue) composed of an impenetrable material, the amount of radiation

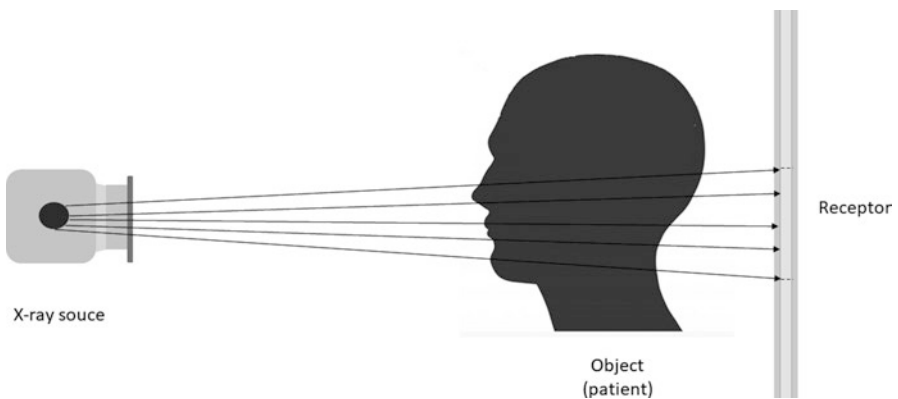


Fig. 1 Main configuration of the patient and devices in the acquisition of radiography

reaching the detector device (or film) is low since the object absorbs most of the x-ray photons. Consequently, its respective projection in the final based radiography is light. This property of materials of retaining x-rays and producing light is called radiopacity. On the other hand, when x-ray photons are irradiated in a material that promotes its propagation, the amount of radiation that reaches the detector device is high, resulting in a respective dark projection in the radiographic image. This property is called radiolucency [5, 6].

Attenuation is the main physical property used in image formation for most conventional x-ray machines and computed tomography scan systems. It is defined as the difference in the amount of x-ray energy emitted by the transmitter and the energy received by the receptor (digital sensor or film) after transposing an object (patient's biological tissue) during the examination.

For dental radiographs, such differences related to the attenuation values can result in four aspects: (1) coherent scattering (when incident photons scatter from outer electrons), (2) photoelectric absorption (when incident photons eject their inner electrons and fade away, releasing photons of a characteristic type), (3) Compton scattering when incident photons eject outer electrons, and (4) other possible scatterings. Attenuation values are commonly expressed in Hounsfield units (HU) that are based in the water's attenuation, so the resultant HU value for a specific tissue can be defined as:

$$\text{HU}_{\text{tissue}} = 1000 \times \frac{\mu_{\text{tissue}} - \mu_{\text{water}}}{\mu_{\text{water}}} \quad (1)$$

where μ_{water} and μ_{tissue} are the water's and tissue's attenuations.

After the exam execution, its results can be represented as images. The analog approach for that is achieved by using film imaging and its processing as in analogical photography. Image formation based on film imaging uses a sensitive layer that is modified by x-ray photons, through oxidation, proportionally to the amount of radiation exposition [5]. The film is then chemically processed to produce a gray-scale image, reflecting the x-ray opacity of each tissue in a continuous range (i.e., this analog approach is similar to the principles applied for the former analogical film imaging).

Digital radiographic application has been increased in the last years, allowing the use of several computer-based processing, like those presented here. Radiography acquired using the digital approaches presents differences from the analog-signal-based ones, starting from the way they are represented. In digital radiography, the measures obtained by the acquisition process are spatially distributed in a discrete way, represented in a digital file and used in a matrix-like structure, defined by its resolution that is the number of rows and columns of such a matrix [5, 6]. When the electronic receptor, used in digital devices, absorbs the x-ray photons that go through the object, it generated a small voltage for each pixel, which is proportional to the volume of photons received by each device position. After that, a process called analog-to-digital conversion (ADC) is performed. It consists of defining ranges for the voltage's values obtained in a way that the pixels whose values are in a defined

range are grouped together, and then the same digital value is assigned to a point [5, 6], forming the digital image. A visualization tool (as a computer monitor) reads these values and assigns a corresponding gray shade for each one to display the matrix as a grayscale image.

According to their physical properties, electronic receptors can be divided into solid-state detectors or photo stimulus phosphor detectors. The first uses solid semi-conducting materials to gather the charge generated by x-ray photons. The second uses photostimulable phosphor plates to absorb, store x-ray energy, and further release it as light after being stimulated by another light that presents an appropriate wavelength [6].

In order to obtain accurate measures and, consequently, higher-quality images, radiation doses should be adjusted. Nevertheless, a larger exposition to radiation poses a risk for the patient's health. For oral radiography, standard radiation ranges were proposed in order to obtain the highest quality preserving the patient's safety.

Usual radiography is an exam that consists of a planar projection of a 3D scene. This scene is composed of the patient's biological tissues and anatomical structures. The way the patient and the device are positioned changes with the exam focus and, consequently, the resultant projected image. Many configurations are defined to cover the different anatomic parts. The next section discusses the most common oral ones, including the positioning of the elements to achieve the images focusing on different structures according to the visualization objective.

2.2 Intraoral Radiographs

As suggested by the term, the acquisition process of intraoral radiographs involves positioning part of the device inside the patient's mouth. There are three main types of intraoral radiographs used in dental imaging: periapical views, bitewing views, and occlusal views. Figure 2 shows them.

To cover all the dental arches of a healthy adult patient, a set of 17 periapical views and 4 bitewing views is required for most cases [6]. Periapical views (Fig. 2(a)) cover the teeth's crowns, roots, and surrounding bones. Figure 3 shows

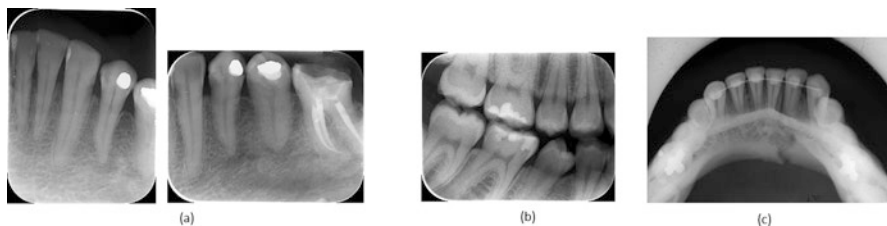


Fig. 2 Examples of (a) periapical projections, (b) bitewing projections, and (c) occlusal projections. Occlusal view by Coronation Dental Specialty Group under CC BY 3.0 via Wikimedia Commons

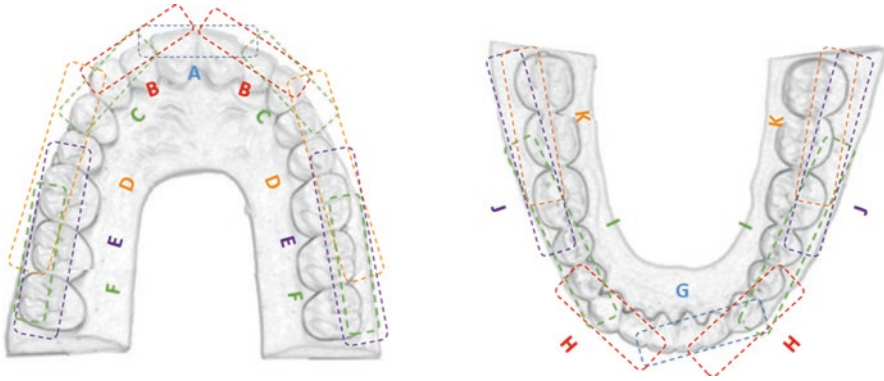


Fig. 3 Target teeth of periapical projections A to K

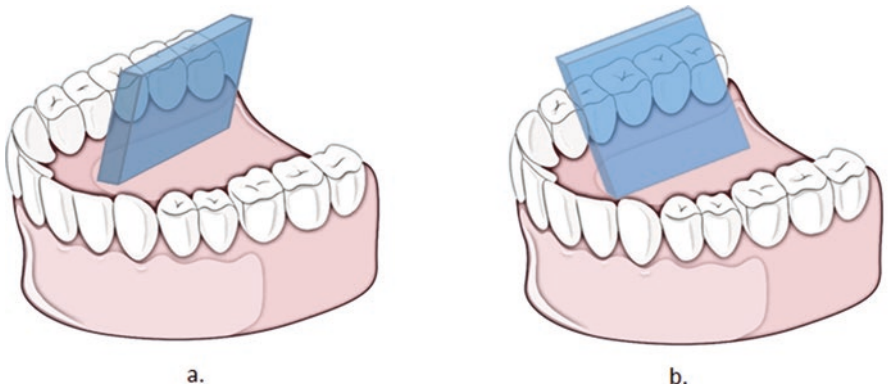


Fig. 4 Receptor positioning in the periapical radiography: (a) paralleling and (b) bisecting angle

the distribution of a complete set of periapical projections and the teeth they respectively cover. Projection A refers to the maxillary central incisors, projection B refers to the maxillary lateral incisors, projection C refers to the maxillary canines, projection D refers to the maxillary premolars, projection E refers to the maxillary molars, projection F refers to the maxillary distomolar, projection G refers to the mandibular centrolateral incisors, projection H refers to the mandibular canines, projection I refers to the mandibular premolars, projection J refers to the mandibular molars, and projection K refers to the mandibular distomolars.

Two projection techniques are mostly used in periapical radiograph acquisitions: paralleling and bisecting angle (Fig. 4).

The parallel periapical radiographs tend to result in an image with less distortion and are commonly recommended for digital imaging. This technique consists of positioning the x-ray receptor as parallel as possible to the dental arches inside the patient's mouth, so the projection is obtained orthogonal to the teeth and the receptor plane [6]. Considering a plane that approximates the surface of few consecutive

teeth (as the maxillary central incisors), the main idea of the parallel technique is positioning the receptor as a parallel plane of the mentioned teeth plan, so the x-ray hit them directly, in a perpendicular direction (Fig. 4(a)). This better reflects the teeth's true anatomical characteristics, reducing distortion in the acquisition.

The bisecting-angle technique is only used when the parallel technique cannot be applied due to large rigid sensors or the patient's anatomy. It is based on the geometric principle that states that two triangles are equal if they have two equal angles and share completely one of their sides. In this projection, the receptor is positioned as close as possible to the internal part of the dental arch, i.e., the lingual surface of the teeth. If the exam focuses on the mandibular teeth, the receptor must be held from the bottom by the palate. If the exam focuses on the maxillary teeth, the receptor must be held on top by the floor of the mouth (Fig. 4(b)). Holding instruments (or the patient's fingers) are used to promote the receptor's perfect positioning for both techniques and the x-ray emission direction is adjusted in a proper manner.

Bitewing views (Fig. 2(b)) are also called interproximal views. They cover the coronal portions of the maxillary, mandibular molars, and premolars in a single image. Four of these projections are acquired with the periapical views covering all arches: two for the premolars (Fig. 5a) and two for the molar teeth (Fig. 5b).

Occlusal projection is another possible intraoral radiograph. It covers a wide part of the dental arches. This exam is mostly used when the patient's mouth cannot hold the periapical receptors. As suggested by the term, the receptor is placed in the occlusion plane. In occlusal acquisitions, the receptor is located between the occlusal surfaces of the teeth. The most common occlusal views are anterior maxillary occlusal projection (Fig. 6a), cross-sectional maxillary occlusal projection (Fig. 6b), lateral maxillary occlusal projection (Fig. 6c), anterior mandibular occlusal projection (Fig. 6d), cross-sectional mandibular occlusal projection (Fig. 6e), and lateral mandibular occlusal projection (Fig. 6f).

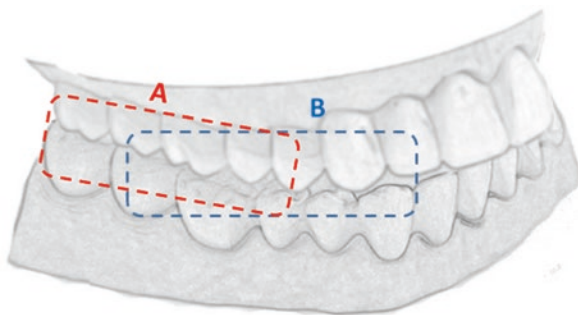


Fig. 5 Target teeth of the bitewing projections

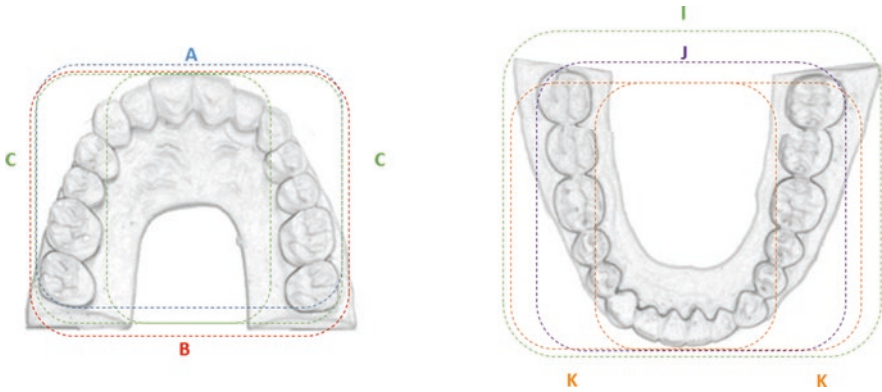


Fig. 6 Target teeth of the occlusal projections

Fig. 7 Example of panoramic radiography. Paronamic radiograph by Umanoide under CC via Unsplash



2.3 *Extraoral Radiographs*

Extraoral radiographs, as suggested by the term, consist of radiographic exams that do not include introducing part of the device into the patient's mouth. The most widely used extraoral image in dental practice is the panoramic view presented in Fig. 7. The panoramic view covers all of the maxillary and mandibular dental arches and a wide part of the face (Fig. 7). The quality of this kind of image can be considered lower than the intraoral radiographs since it promotes geometric distortions, moving shadows, and even the inclusion of overlapping effects due to the presence of other anatomical structures in the dental arch proximity, as the neck bones. It is mainly recommended for initial evaluations and for cases in which intraoral radiographs cannot be acquired [6].

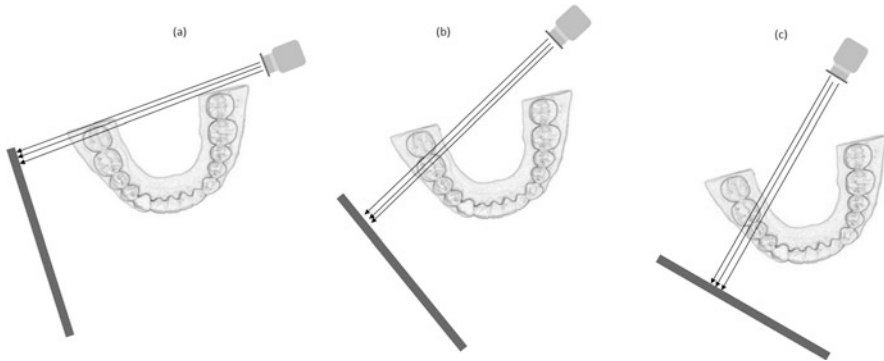


Fig. 8 Representation of the acquisition process in a panoramic view

For panoramic acquisition, the object of interest (mouth of the patient) is positioned in the plane (image layer) in a central point in relation to the x-ray source and receptor, which are on opposite sides. Then receptor and x-ray source move simultaneously. The panoramic image is formed dynamically, that is, its acquisition is made during the device movement, so each part of the image corresponds to a different position and time. To analyze this, due to such a dynamic capture, the receptor movement has to be considered, as well as the x-ray source's position and also the part of the mouth which is currently on focus.

Figure 8 shows the position of the device at three different times of the panoramic acquisition. Note that Fig. 8(a) corresponds to the first tooth of the acquisition process, so the part of the receptor that is directly receiving the x-rays emitted by the source corresponds to this part of the mouth. As the device continues, the acquisition process covers the rest of the dents and the receptor also moves to receive the x-ray of the corresponding part of the mouth (Fig. 8(b) and (c)). Note that the receptor moves close to the patient's teeth arches while the x-ray source moves behind the patient's neck. The receptor is intentionally positioned in this fashion because the structures close to the receptor are better projected in the resulting image. Due to the projection principles of these images, the structures that are close to the x-ray source are projected in the formed image in a way that it appears with magnification, resulting in deformed results and blur [6].

Extraoral radiographs also include several other projections. The most used ones are lateral skull projection (lateral cephalometric projection), submentovertebral (base) projection, Waters' projection, posteroanterior skull projection (posteroanterior cephalometric projection), reverse Towne projection (open mouth), and mandibular oblique lateral projections. These projections are mostly used in orthodontics and cephalometric landmark identification [6].

2.4 Computed Tomography and Cone-Beam Computed Tomography

Traditional computed tomography (CT), also called fan-beam tomography, acquires three-dimensional images by irradiating x-ray beams linearly. In order to cover a 3D object, the source and the receptor must surround a central point of the object plane, as shown in Fig. 9. When a rotation of 360° is completed, both source and receptor translate in the direction of the plane's normal axis, covering all the volume (object) [7]. Each rotation of the source results in a planar slice. As the source moves along the axial direction, new slices are obtained. After the end of the acquisition process, the slices are then computationally processed, creating a 3D digital volume. For a pair of consecutive slices, the values are interpolated to fill the region between them, resulting in a continuous-like representation. The number of slices depends mainly on the device's characteristics. Over time, CT acquisition process has improved, especially concerning the number of slices and the way that the x-ray source's rotation and axial displacement are performed. Nevertheless, the current CT devices still follow these principles.

In dental imaging, the most prevalent 3D image is the cone-beam computed tomography (CBCT). In the acquisition process of CBCT, 3D images are acquired in only one rotation of the x-ray source. The beam used in this process has a cone format, so in each step of the rotation, a complete 2D projection is acquired at once (Fig. 10). There is no need for axial displacements of the x-ray source [7]. So the number of slices depends mainly on the digital receptor discretization. For CBCT, the acquisition produces an entire discrete 3D volume, with voxels' sizes according to the receptor pixel size [7].

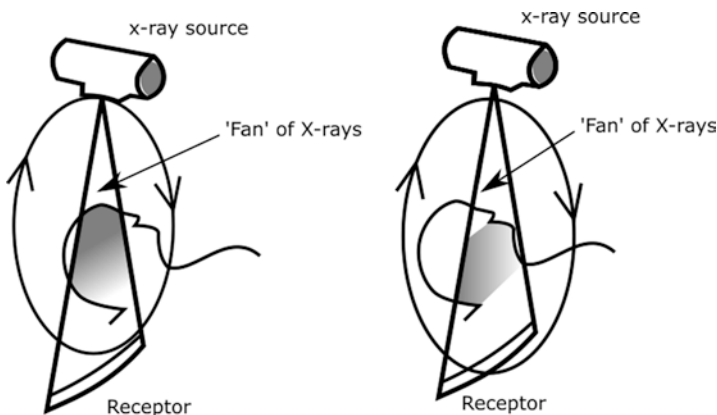
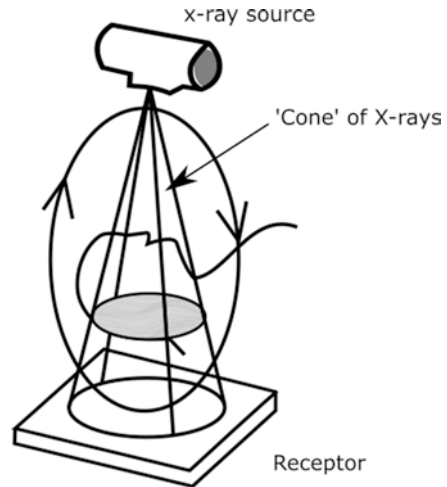


Fig. 9 Representation of the acquisition process of CT at two different positions

Fig. 10 Representation of the acquisition process of CBCT



3 Image Processing and Artificial Intelligence for Dental Image Analysis

This section describes the most common image processing tasks involving artificial intelligence techniques and their real application for dental imaging. In addition to image enhancement (which can be considered as an interesting application for dental imaging), it also discusses the most prevalent applications of the presented techniques for two dentistry sub-areas: periodontology and cariology. Some previous works in literature that present solutions for problems in the area are considered, as well. Moreover, practical examples applying AI and image processing for tasks involving classification, detection, and image enhancement are analyzed.

3.1 Artificial Intelligence Techniques for Image Processing

Digital image processing (IP) is a field of computer science and signal processing that studies digital signals presenting two-dimensional (2D) structures. Radiography is a type of biomedical imaging where the radiographic devices themselves apply some of these techniques after the images' acquisition and before their storage [6]. A wide range of available IP processing techniques can be employed to improve, analyze, and extract information from oral radiographs. Researchers can apply IP techniques as it better suits their objectives. Four techniques are mostly related to dental imaging applications: enhancement, segmentation, identification, and classification. Figure 11 illustrates them.

The enhancement task consists of processing the image to improve its quality concerning noise, resolution, edge definition, etc. The objective of the segmentation

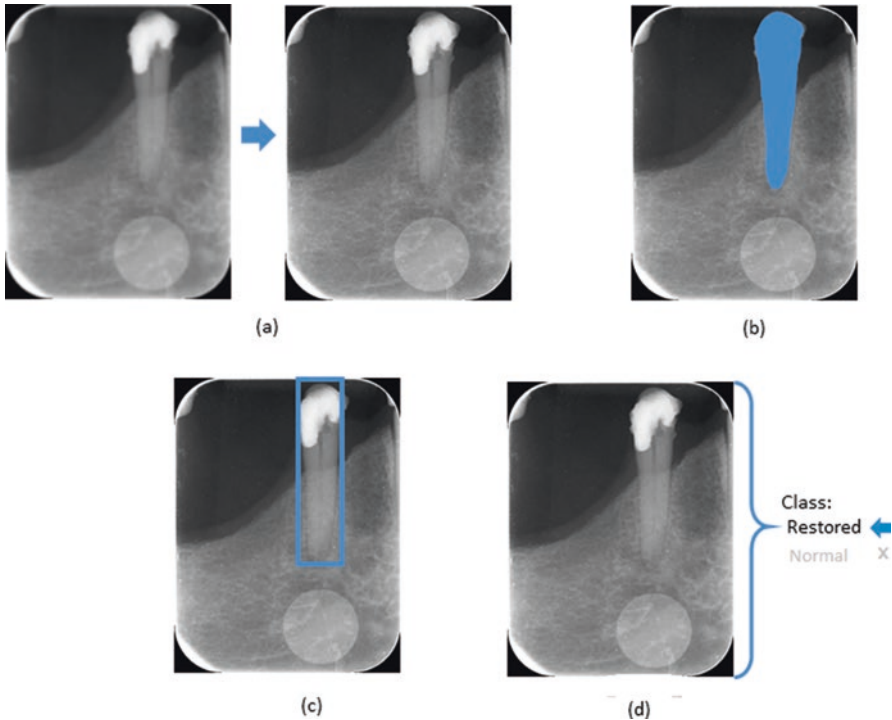


Fig. 11 Example of image processing tasks: (a) enhancement, (b) segmentation, (c) detection, and (d) classification

task is to identify an object in the image, determining its exact boundaries and isolating it from the rest of the image. Figure 11(b) shows an example of this task (tooth segmentation). The identification task focuses on determining the region of the image that encloses an object, for example, tooth detection (Fig. 11(c)). Finally, the classification task consists of analyzing the entire image and its visual patterns to associate it to a specific class, as in the example in Fig. 11(d), in which a tooth is classified as normal or restored.

In the last years, the use of artificial intelligence (AI) techniques as a support to traditional IP has been increasing, leading to results that demonstrate their potential. Convolutional neural networks (CNNs) are an essential part of AI being the basis of most AI-based algorithms for IP nowadays. CNNs consist of a specialized kind of intelligent algorithm for processing data that present a grid-like topology as digital images.

The IP convolution operation is the base of CNN algorithms. Convolution consists of transforming an input image using a kernel to achieve a feature map as output. More specifically, given two matrices, with the same numbers of elements, i.e., $n \times n$, one named kernel and the other being a part of the image to be convoluted, then convolution consists of multiplying correspondent position and adding them to

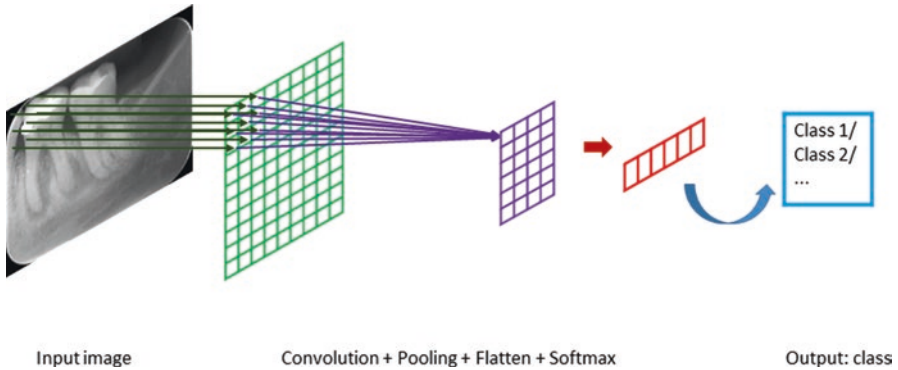


Fig. 12 Schematic representation of an input, a simple CNN, and an output

obtain a value that is used to compose the output for each position that the kernel can cover on the image. Figure 12 exemplifies the convolution and other operations that can be combined to compose a simple CNN to perform a classification task.

The output named feature map can be processed by another convolution or even other operations. Among the most used other types of operation are size reduction (or pooling), dimension reduction resulting in a 1D vector (or flatten), and mapping of the position to a class previously defined (or softmax).

Note that different CNN architectures can be achieved by combining convolutions with those operations in various manners, modifying the number of layers and even the way the operations are organized. Some architectures proved to be efficient for a wide range of applications receiving particular names as the ResNet and Inception. The most straightforward application of CNNs is for classification tasks.

Some detection and segmentation tasks can be modeled as extensions of the classification. For example, consider a 95×20 area of the input that includes a tooth to be detected as in Fig. 11(c). Note that if one divides this image into two sub-images of size 5×5 and classifies each sub-image to identify if they present a tooth or not, then the sub-image that encloses the tooth can be identified. Therefore, the region that is a union of each tooth sub-image can be considered as the tooth region resulting in a detection. However, in real cases, the object (tooth in this example) may not be so well positioned, so several different subdivisions must be tested to achieve the segmentation that covers only one complete object. This is the main idea behind most AI-based detection algorithms.

The segmentation task can also be considered as an extension of the classification task, but on a pixel scale, in the sense that each pixel is classified as belonging to the object's area or not (Fig. 11(b)). The main idea behind most AI-based segmentation algorithms is to analyze each pixel, considering their neighborhood, which is defined by a window size, to evaluate if it corresponds to the patterns that characterize the object. In other words, each pixel is classified as being part of the object or not considering a tiny sub-image as input, which is defined by a window that covers its neighborhood.

Actually, the real algorithms that perform these tasks are much more complex than this simple description presented here since they include several layers (operations), but this gives us the main idea of how they work.

Note that all these concepts can be extended to N-dimensional signals, including 3D data, so they can also be applied in tomographs, for example.

3.2 *Identify Periodontitis*

Periodontal disease (PD) is a consequence of interactions between bacterial biofilm and the host's immune response [6, 8], and differences in the degree of severity and impairment of this disease can be influenced by extrinsic factors, such as smoking, and intrinsic factors, such as diabetes mellitus [9]. PD can be divided into gingivitis and periodontitis [6, 10]. One of the consequences of tissue destruction due to periodontitis is bone loss. Radiographically and clinically, this loss can be observed as an increase in the distance between the enamel-cement junction to the alveolar crest.

The fact that this tissue destruction can be identified radiographically promotes the use of AI-based image processing techniques for this purpose. Recently, Lin et al. [11, 12] proposed the use of deep learning models for alveolar bone loss identification [11] and measurement [12]. The model proposed by Lee et al. [13] focuses specifically on the identification and severity assessment of premolars and molars periodontally compromised. Similarly, the studies of Carmody et al. [14] and Mol et al. [15] aim to classify periapical lesions according to their extent. A considerable part of the works that focus on identifying periodontitis/periapical diseases uses panoramic radiographs. For example, Ekert et al. [16] used convolutional neural networks (CNNs) to detect apical lesions on panoramic dental radiographs. The implemented network, a custom-made seven-layer deep neural network, achieved a sensitivity value of 0.65, a specificity value of 0.87, a positive predictive value of 0.49, and a negative predictive value of 0.93. Krois et al. [17] applied a seven-layer deep neural network to detect PBL on panoramic dental radiographs. The classification accuracy of the CNN was 0.81, and the sensitivity and specificity were 0.81 and 0.81, respectively.

Classifying Approximal Bone Loss in Periapical Radiographs Identification of periodontal diseases is a common application area for AI-based image processing, as exposed previously in this section. Intraoral radiography, especially periapical exams, is an important tool for identifying these anomalies, facilitating their diagnosis, treatment, and prognosis [18]. Next, this section demonstrates the use of some AI and image processing techniques to pre-process and classify interproximal regions in periapical examinations according to the presence of proximal bone loss. For that, a brief evaluation of the use of two CNNs architectures is performed, specifically ResNet and Inception networks, to demonstrate how different architectures can influence the quality of the final results.

This experiment used 1079 interproximal regions manually extracted from 467 different periapical radiographs. All images are in grayscale, in “jpeg” format. This experiment is focused on a classification task. Therefore, the region extraction was performed manually. The next section covers a detection task to automatically extract the regions of interest from oral radiographs using image processing techniques.

Firstly, an adaptive histogram equalization [19] was applied to the periapical images in order to increase their quality. The adaptive histogram equalization is an image processing technique used to improve contrast in images and enhance their details. It adjusts the image contrast by considering its most frequent tonalities. The process is similar to the original histogram equalization; however, it considers parts of the images rather than the entire image, allowing it to create different histograms and use them to calculate the equalization [19]. The main idea of this technique is to define a neighborhood window to be considered in the histogram of the transformation function for each pixel. In this experiment, after some initial testing, an 8×8 window was selected. As in the ordinary histogram equalization, the transformation function of the adaptive histogram equalization is proportional to the cumulative distribution function (CDF) of the pixel values in the neighborhood [19].

Experts marked the regions of interest (ROIs) in each exam. The ROIs cover the areas that can be affected by bone loss. These regions consist of interproximal (between two teeth) areas, limited superiorly by the enamel-cement junctions and inferiorly by the alveolar crests. To be used as input of a convolutional neural network for the proposed classification task, all these data must be labeled, i.e., for each case/image an associate class must be assigned to it by experts. This process is called data labeling and can be performed using several auxiliary tools. This example used the labeling tool named DataTurks (available at <https://dataturks.com/>). Two experts annotated the exams’ ROIs, using bounding boxes, denoting which of them present any bone loss and which do not. They are experienced dentists specialized in oral radiology. They annotated 1079 regions: 388 with no lesions and 691 with bone loss (no differences between experts’ annotations).

In order to prepare the data for the classification task, this data must be organized into three different sets: training, validation, and test sets. This process is called dataset split. The test dataset was formed by 52 samples of each class randomly selected. The remaining images underwent a data augmentation process based on horizontal and vertical flips. After that, the 639 remaining annotated regions with vertical bone loss provided 1278 images, using only horizontal flips. The remaining 336 images of healthy regions provided 1344 images, using both horizontal and vertical flips. In that way, the CNNs’ training and validation sets are formed by these 2622 images.

Finally, the actual classification task is performed. As mentioned, this example includes an evaluation of two different CNNs for the classification task in order to compare which is the most appropriate to the proposed problem. This experiment included two architectures that demonstrate good performance for a wide range of applications: ResNet and Inception architectures. The ResNet architecture used in this work has 50 layers in total. It is composed of several stacked blocks, called

residual units. Such units consist of two convolutional layers and two activation functions [20]. On the other hand, the Inception architecture is formed by blocks called Inception modules [21], consisting of a combination of convolutional layers with different kernel sizes and a pooling layer. This study used the official Keras ResNet and Inception implementations. The data processing performed in this work used the Python language and the scikit-image library. The parameters used for training the CNNs are outlined in Table 1.

The CNNs' training used the backpropagation algorithm and included 180 epochs. For each epoch, we checked the accuracy and loss values. Each epoch corresponds to one time in which CNN weights are updated considering all elements of the training dataset. The models used in this example were previously pre-trained using the ImageNet dataset [25] to obtain better initial weight values.

An important measure to be considered in the evaluation of CNNs in a classification task is test accuracy (proportion of cases properly classified by the considered model). Other measures are sensitivity (recall), specificity, precision (positive predictive value, PPV), and negative predictive value (NPV) [26]. In this example, such measures are based on:

- True negatives (TN) – regions correctly classified as healthy
- True positives (TP) – regions correctly classified as regions with bone loss
- False negatives (FN) – regions with bone loss incorrectly classified as healthy
- False positives (FP) – healthy regions incorrectly classified as regions with bone loss

In that way, the mentioned measures are defined as sensitivity = $\frac{TP}{FN + TP}$, specificity = $\frac{TN}{FP + TN}$, precision = $\frac{TP}{FP + TP}$, and negative predictive value = $\frac{TN}{FN + TN}$ [26].

Most evaluations also include the receiver operating characteristic (ROC) and the precision-recall (PR) curves [26]. In this example, all measures were calculated using Python and the scikit-learn library.

At the end of the training process, the Inception model presented an in-sample accuracy of 0.984 and a validation accuracy of 0.933. On the other hand, the ResNet model had an in-sample accuracy of 0.919 and a validation accuracy of 0.818. Concerning the evaluation based on the test set, the results are shown in the respective confusion matrices (Table 2). Note that the test accuracy (proportion of

Table 1 Hyperparameters used in CNNs' training

Parameter	ResNet	Inception
Optimizer	Momentum [22]	RMSprop [23]
Batchsize	32	32
Learning rate	0.01	0.01
Loss function	RMSE [24]	RMSE [24]

Table 2 Confusion matrices for the ResNet and Inception models

ResNet	$n = 104$		Predicted	
	Actual	Healthy	38	14
		Vertical bone loss	13	39
Inception	$n = 104$		Predicted	
	Actual	Healthy	37	15
		Vertical bone loss	4	48

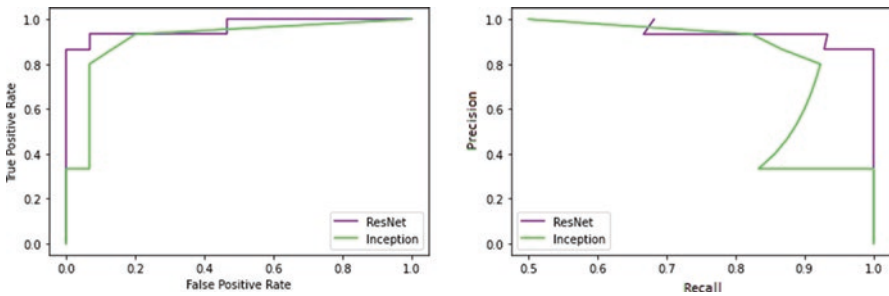


Fig. 13 ROC (left) and PR (right) curves for each model

Table 3 Test results

Measure	ResNet	Inception
PPV (precision)	0.736	0.762
Sensitivity (recall)	0.750	0.923
Specificity	0.731	0.711
NPV	0.745	0.902
AUC-ROC	0.864	0.860
AUC-PR curve	0.868	0.847

examples correctly classified) of the ResNet model was 0.740 and the Inception’s was 0.817. Table 3 summarizes the other test measures, and Fig. 13 shows the ROC and PR curves for each model.

Note that the Inception model had the best overall performance (Tables 2 and 3). The lower performance of the ResNet model consists of misclassifications almost equally distributed between both healthy and bone loss classes. On the other hand, the misclassifications for the Inception model are mainly healthy regions incorrectly classified as regions that present vertical bone loss. Finally, the good results of the considered CNNs are denoted by the ROC and PR curves.

3.3 *Detection of Dental Caries*

Dental caries is a multifactorial oral disease affected by sucrose consumption. It presents a high prevalence [27], and its prevention demands early detection and treatment. Its development depends on the presence of bacteria, especially mutant streptococci, that ferment carbohydrates, resulting in the demineralization of hard dental tissues [28–30]. The accumulation of such bacteria forms what is known as plaque (biofilm) [27]. It initially affects the tooth surface, and after severe demineralization or cavity formation, it can penetrate the hard tissues. Clinically, when it is visible, dental caries presents as a matte white spot (indicating ongoing activity) or an opaque or dark brownish spot (indicating past activity) [6, 31]. Demineralization may extend into the dentin, the enamel, or even the pulp and can destroy the entire tooth structure [30].

Most caries lesions are visible in periapical images. Approximal caries affects the interproximal area between two consecutive teeth. They are generally detected through image examinations, especially bitewing radiographic images, because the positions of such lesions prevent a clinical evaluation. In bitewing images, dental caries appears as a darker area due to their low x-ray absorption [6].

Several previous works have focused on identifying dental caries by examining images such as optical coherence tomography (OCT), periapical radiography, and bitewing images. Although initially, traditional image processing methods were applied in most works [32–35], machine learning algorithms have recently become a more common approach to visual problems, including dental images. Deep convolutional neural network (CNN) algorithms have been used for human oral tissue classification to provide early detection of dental caries [36]. A CNN model analyzes optical coherence tomography (OCT) images of different densities of oral tissues and determines variations related to the demineralization process. That suggests that variations in caries lesion may be identified in other image examinations as well, as previously mentioned.

Deep CNNs have also been applied to the detection and diagnosis of dental caries on periapical radiography images [37]. A pre-trained GoogLeNet Inception v3 model was used to process 3000 periapical radiographs. Three different models were created: a premolar version, a molar version, and a final version for both premolar and molars. These models achieved impressive accuracy results (89.0%, 88.0%, and 82.0%, respectively). Thus, considering the good performance of the presented method, the study showed the feasibility of using a deep CNN architecture to detect and diagnose dental caries.

Bitewing images have also previously been evaluated to identify dental caries stages and potential false diagnoses [38]. In that study, several texture features were extracted from the evaluated images via a gray-level co-occurrence matrix (GLCM). These feature values were processed by an algorithm that combines a logit-based artificial bee colony optimization algorithm with a backpropagation neural network to increase the classification accuracy. The proposed approach achieved an accuracy of 99.16%.

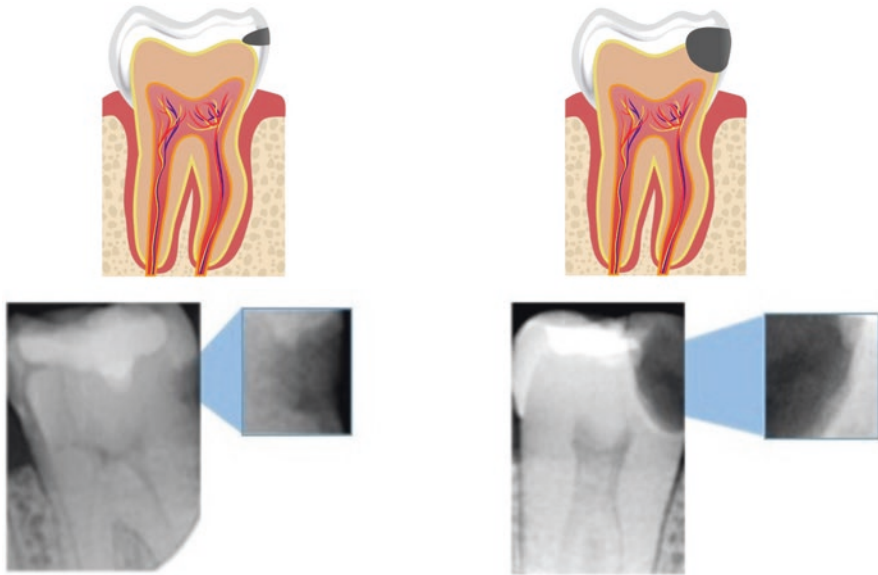


Fig. 14 Tooth stages considering the caries severity: (a) representation of tooth with an incipient lesion, (b) bitewing image with incipient lesion highlighted, (c) representation with an advanced lesion, and (d) real example of a bitewing exam with advanced caries

Approximal Dental Caries Detection and Classification in Bitewing Images As previously mentioned in this section, caries detection is a common application area for AI-based image processing. Next, this section demonstrates the use of some AI and image techniques to detect approximal caries in bitewing images and classify them according to their severity. Consider three different caries stages based on their lesion severity: normal (no lesion), incipient (superficial lesion affecting the enamel; Fig. 14a and b), and advanced (lesion affecting a considerable part of the tooth, expanding into the dentin and the pulp; Fig. 14c and d).

The first step to prepare the data for the CNN classification is to detect the teeth in the bitewing radiographs using image processing techniques. Each of the detected teeth was separated, creating individual tooth images. As previously mentioned, tooth detection is a task for which previous works applied deep neural networks, as YOLO and Fast RCNN. Nevertheless, in an ideal scenario, classic image processing techniques may also present a good performance, as demonstrated in this experiment. This experiment excludes cases of dental implants, crowding, and malocclusion. For these cases, deep learning solutions may present better results.

The teeth detection method based on classic image processing techniques has as the first step an equalization operation (Fig. 15) to enhance the details and differentiate between background and tooth areas more easily. This example uses the adaptive histogram equalization. As a result of the equalization process, teeth and background can be more easily differentiated in the images because their tonalities

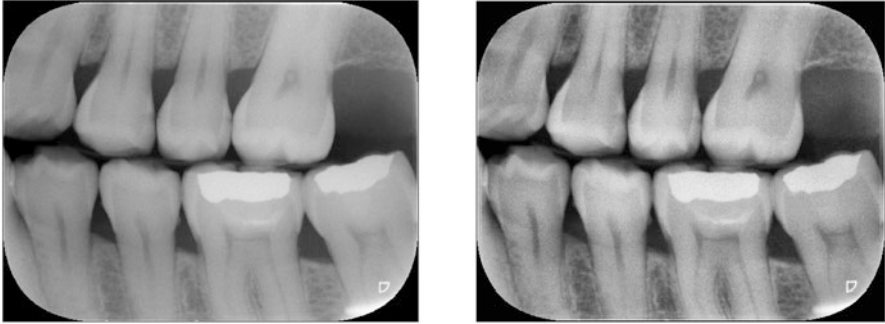


Fig. 15 Application of adaptive equalization: (a) original image and (b) equalized image



Fig. 16 Pre-processing using morphologic operations: (a) thresholded image, (b) eroded image, (c) open image, and (d) dilated image

differ more substantially. Thus, a threshold can be used to transform the original grayscale images into binary images where the background is black and the tooth area is white. This example used the Otsu threshold [39].

Observe that in the resulting binary images, sometimes the gum area is considered as background, and sometimes it is included in the white tooth area due to tonal similarities between the tooth and gum regions. These gum regions are removed using morphological operators [40]. The white areas related to teeth consist of large regions with few holes, while the white areas pertaining to gum are mostly small and irregular and can easily be removed using erosion and opening morphological operations applied consecutively [40]. Considering the thresholded image (Fig. 16a), the next step is to apply erosion using a 130×20 rectangle as a structuring element (Fig. 16b). This specific element was chosen after evaluating the gum areas' shapes. The use of smaller elements did not result in the correct elimination of the gum areas. Similarly, the use of larger elements resulted in considerable losses in the identified tooth regions. Furthermore, using a uniform, symmetrical square or circle element did not allow the separation of teeth that are close together.

Next, an opening operation was applied, using a circle with a radius of 20 pixels as the structural element. This operation results in the elimination of the remaining undesirable parts (Fig. 16c). Finally, dilation is applied using a circle with a radius of 15 pixels as a structuring element, which results in the inclusion of the tooth borders in the tooth areas (Fig. 16d).

After removing the gum areas, the binary images are composed of large white areas on a black background. Each area refers to a different tooth. New images of

each tooth are created based on the bounding boxes around these areas. Thus, the original image is repeatedly cropped, using the bounding boxes' limits to obtain individual images for each tooth.

A total of 480 different tooth images were extracted from the 112 bitewing radiographs by the described detection method. To be used as an input of a convolutional neural network for the proposed classification task, all data must be labeled, i.e., the lesion severity class must be assigned for each tooth. To obtain the labels for each of the 480 teeth, 2 experts used a labeling tool named DataTurks (available at <https://dataturks.com/>) to associate each detected tooth to 1 of the considered classes: healthy, incipient, or advanced. These experts are experienced dentists, and one is specialized in oral radiology. This labeling process pointed out that the set of 480 detected teeth included 305 normal teeth, 113 teeth that present incipient lesions, and 62 teeth that present advanced lesions. There was no discrepancy between their annotations, i.e., they pointed to the same classes for all cases.

The next step in the data preparation for the classification task is the dataset split. The data must be split into training and test sets used to train and evaluate the CNN model, respectively. Fifteen cases of each class are used as a test set, resulting in 45 teeth. The remaining 435 tooth images (divided into 290, 98, and 47 images for normal, incipient, and advanced classes, respectively) underwent a data augmentation process. The data augmentation process consists of creating variation in input images to increase the data volume, which was proved to be essential to achieve good results in deep learning models [41]. This example's data augmentation processes consist of applying rotate and flip operations to the tooth images, creating 1160, 1176, and 1128 sample images for healthy (normal), incipient, and advanced classes, respectively.

Due to the outstanding performances presented by Inception v3 models in prior medical image classification studies, this CNN architecture was chosen to be used in this experiment [42]. The parameters used for training the CNNs are outlined in Table 4.

The models used in this example were previously pre-trained using the ImageNet dataset [25] to achieve better initial weight values. The fine-tuning training process included 11,500 steps and 3 different values (0.1, 0.01, and 0.001) as the initial learning rate to evaluate which of these parameter values would be the most appropriate.

The final accuracy and loss values, considering the training and validation sets, achieved after completing the training process, pointed out that the best Inception model was the one with a learning rate of 0.001. Therefore, this CNN model must be evaluated using the test dataset. In addition to the test accuracy, CNN evaluation

Table 4 Hyperparameters used in CNN training

Parameter	Value
Optimizer	Momentum [22]
Batch size	16
Learning rate	0.1, 0.01, and 0.001

currently includes the following measures: sensitivity (recall), specificity, positive predictive value (PPV, or precision), negative predictive value (NPV), and the area under the curve (AUC) for the receiver operating characteristic (ROC) curve, to evaluate the model’s performance considering the data in the test dataset. The model’s evaluation for each class based on the test data resulted in the values shown in Table 5. The confusion matrices in Table 6 summarize the overall and the specific results for each class. Another essential measure considered in this evaluation is the ROC curve. Figure 17 shows the curves for each class.

Observe that there is some disparity in the performance considering the three different classes, which is perceptible in confusion matrices (Table 6), the main test results (Table 5), and the ROC curves (Fig. 17). Nevertheless, the results suggest the applicability of CNNs for the proposed task.

Table 5 Test results

Class	Precision	Recall	Specificity	NPV	AUC-ROC
Normal	0.818	0.600	0.933	0.823	0.643
Incipient	0.722	0.866	0.833	0.926	0.861
Advanced	0.687	0.733	0.833	0.862	0.810

Table 6 Confusion matrix

		Predicted		
		Normal	Incipient	Advanced
True	Normal	60% (9)	13% (2)	27% (4)
	Incipient	7% (1)	86% (13)	7% (1)
	Advanced	7% (1)	20% (3)	73% (11)

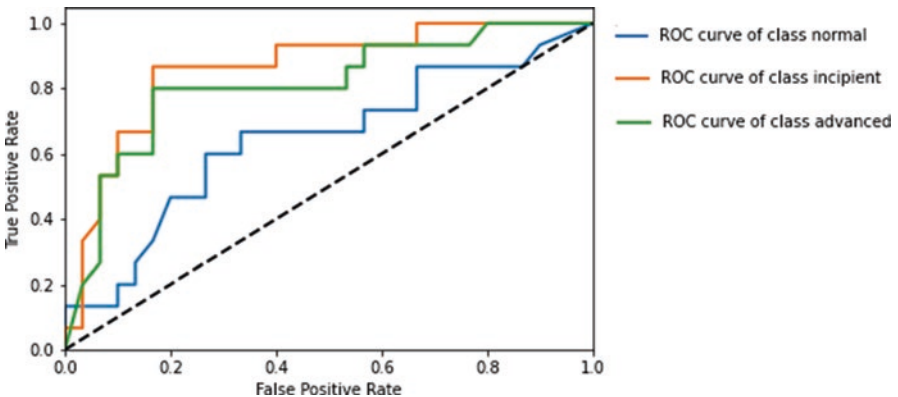


Fig. 17 ROC curves of each class for the model

3.4 *Image Enhancement*

The limitations in radiographic acquisition devices can result in low-resolution images, compromising the diagnosis process [43, 44]. Traditional image processing techniques, as interpolation methods, can be used to increase images' resolution. However, their results can be improved by AI-based methods.

Radiographic image enhancement also includes noise removal and image reconstruction, i.e., recovering missing parts of the image. Actually, for oral radiographs, these tasks are closely related since the missing data is comprehended as noise in this context. Moreover, noise removal demands reconstruction to replace this noisy data with the actual data. AI-based solutions are also popular for these tasks. As discussed in the section Challenge Issues, radiographs' acquisition and image formation processes can lead to a wide range of artifacts and noise. According to Schulze et al. [45], artifacts and noise in oral radiographs include blur, scatter artifacts, extinction artifacts (missing value), beam hardening artifacts, exponential edge gradient effects, aliasing artifacts, ring artifacts, and motion and misalignment artifacts. Another critical artifact that significantly affects the image quality is the metal artifact. Image processing techniques can aid the reduction of some of these artifacts, especially those that include AI-based techniques. There are a significant number of works in literature focused on artifact removal in dental imaging. Among them are the works of Wang et al. and Chang et al. [46, 47] that propose using neural networks for ring artifact removal in CBCT images. Xie et al. [48] present an algorithm based on convolutional neural networks to reduce scatter artifacts in CBCT. Zhang et al. [49] developed a convolutional neural network-based framework to reduce the effects of metal artifacts.

Increasing the Quality of Digital Periapical Radiographs Using SRCNN This section's example demonstrates the application of a widely known deep learning algorithm, called super-resolution convolution neural network (SRCNN) [50], to obtain high-resolution periapical images from low-resolution ones, reaching a magnitude improvement of 4×. Its results are compared with other super-resolution solutions based on more traditional image processing techniques, which are the nearest, bilinear, bicubic, and Lanczos interpolations.

SRCNN is a widely used deep learning-based super-resolution method. Dong [50] initially proposed it in 2016. In its pre-processing, the original low-resolution image is rescaled to its final size by applying the bicubic interpolation. Such a rescaled image is the input of the network that manipulated it in three main steps: patch extraction and representation, nonlinear mapping, and reconstruction. In the first step, patches are extracted from the bicubic rescaled image. Such patches are represented as high-dimensional vectors. In the second step, these high-dimensional vectors are mapped into other vectors, in a nonlinear way. In the third step, it aggregates the high-resolution patch-wise representations to obtain the output (high-resolution image). Figure 18 shows a representation of the steps that compose the SRCNN.

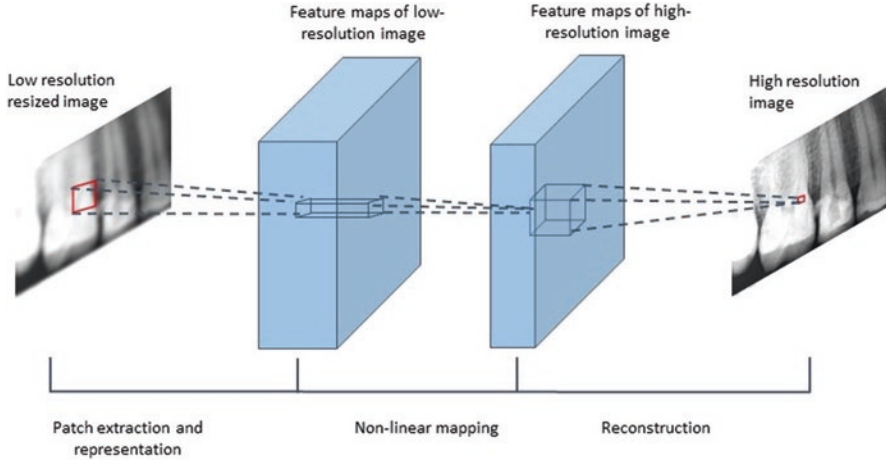


Fig. 18 SRCNN representation

Table 7 Evaluative measures for each considered method

Method	MSE	PSNR	SSIM
Nearest interpolation	101.04 (± 47.35)	28.79 (± 3.08)	0.92 (± 0.02)
Bilinear interpolation	62.06 (± 29.89)	31.16 (± 3.48)	0.94 (± 0.03)
Bicubic interpolation	53.39 (± 26.01)	31.87 (± 3.63)	0.94 (± 0.02)
Lanczos interpolation	48.67 (± 24.07)	32.33 (± 3.75)	0.95 (± 0.02)
SRCNN	33.69 (± 14.17)	33.53 (± 2.86)	0.98 (± 0.01)

The training process of the SRCNN model included 10,000 epochs and used the Adam optimizer and a learning rate of 3×10^{-4} . The dataset used in the training process of the SRCNN model is composed of 228 different periapical radiographs.

After the training process, the obtained model must be evaluated considering the test set that refers to a new set of images not used for training. The test set is formed by 100 selected periapical radiographs, from the 120 that compose the original dataset provided by Rad et al. [51]. Such radiographs were collected in the Dental Clinic of the Universiti Teknologi Malaysia (UTM) Health Center using a Sirona device. All images are in grayscale, in “jpeg” format, with dimensions of 748×512 . Also, a padding operation is applied to the images of both training and testing sets, in order to obtain squared images, with the same number of rows and columns, in order to facilitate the SRCNN processing.

The analysis of the results included three metrics to evaluate the similarity between the images achieved by the considered methods and the ground-truth high-resolution images: mean square error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity index measure (SSIM). The results for all considered methods using the test dataset are presented in Table 7.

Note that the values presented in Table 7 demonstrate the SRCNN model’s superiority, which outperformed all other methods for all considered measures. This

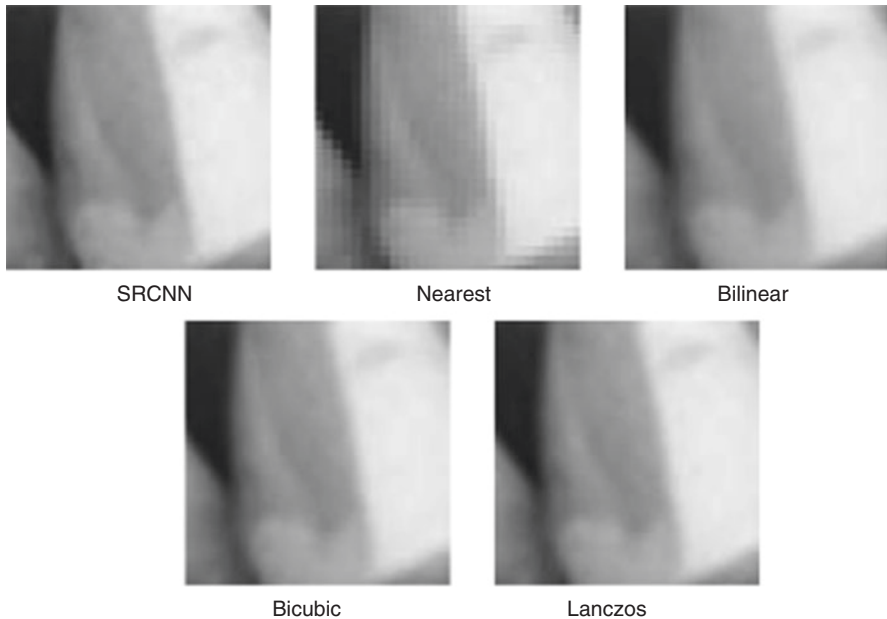


Fig. 19 Detail of images generated by the considered methods

quality increase can also be observed by visually analyzing the images generated by each method (Fig. 19). Note the aliasing effects of the nearest interpolation's image and the blur effects of the bilinear's, bicubic's, and Lanczos' images. On the other hand, the SRCNN model led to less noise and more detailed edges.

3.5 Other Applications

There is a wide range of dental imaging applications for which the techniques covered in this chapter can be applied, as mentioned in the Introduction of this chapter. Both Schwendicke et al. [2] and Hung et al. [4] pointed out the localization of cephalometric landmarks is a very popular application. Cephalometric landmark localization in dental practice is performed manually by experts or supported by computerized tools, mostly in a semi-automatic way. As an attempt to automatize this process, several AI-based image processing solutions have been proposed in the last years. As discussed in the Challenge Issues section of this chapter, there is no consensus on the number of landmarks to be used in dental practice, so the works in literature that cover this topic vary by considering from 10 to 43 landmarks. Half of the works analyzed by Hung et al. [4] presented results considered by the authors as promising, with accuracy values ranging from 35% to 84.70%. Nevertheless, the

quality of the results of the solutions available in literature still does not suit clinical requirements.

In that context, Arik [52] used shape-based CNN models to recognize landmarks' appearance patterns, defining probabilistic estimations for landmark locations. Song [53] proposed a two-step automatic method to detect cephalometric landmarks, which consists of (1) extracting image patches for each landmark and (2) – detecting the associated landmark in each patch using a ResNet model. The network directly outputs the coordinates of the landmarks.

Another popular application in this context is the detection of osteoporosis and low bone mineral density (BMD). Both of these conditions can be identified in radiographs due to their radiodensity-related aspects. Recent works for these applications achieved around 95% for accuracy, sensitivity, and specificity, suggesting that their inclusion in real-world dental practice is close. A great amount of these works defined features to be used as input in classifiers [54–57].

Diagnosis and segmentation of maxillofacial cysts and tumors using the mentioned tools are also commonly assessed in literature. The work presented by Abdolali et al. [58] considers the symmetry of oral anatomy to identify areas referent to cysts. Mikulka et al. and Nurtanio et al. [59, 60] proposed semi-automatic solutions using AI-based image processing techniques to detect, segment, and classify lesions of this type. More recently, Lee et al. [61] proposed using a GoogLeNet Inception v3 model to detect and classify odontogenic keratocysts, dentigerous cysts, and periapical cysts in CBCT, achieving an AUC of 0.914, a sensitivity of 96.1%, and a specificity of 77.1%. Kwon et al. [62] developed a CNN model inspired by the YOLOv3 architecture to detect and classify odontogenic cysts and tumors, which present 88.9% for sensitivity, 97.2% for specificity, 95.6% for accuracy, and 0.94 for AUC.

Other application areas in dental imaging are detection, segmentation, and classification of other anatomical structures, including teeth, jaw bone, and root canals; biofilm classification; diagnosis of multiple dental diseases; classification of tooth types; identification of inflamed gum; identification of dental plaque; and classification of the lower third molar stages.

3.6 Conclusions and Challenge Issues

Although there are several clinical decision support systems that have been developed in the last years, few of them are actually used in clinical settings and previous studies denote a low clinical acceptance of them [63–65], even considering that there is a consensus about their improvement in care and promotion on experts' efficiency [66]. That is, the great potential of the computational tools for dental image analysis based on image processing and artificial intelligence techniques still faces a significant amount of resistance from dentists and oral radiologists. In part, this may happen due to their novelty aspect: Research works employing CNNs in dentistry started in 2015 [2, 3]. This also could be related to the fact that the great

majority of AI-based solutions do not consider the dentist comprehension factor, working basically as a *black box*, which affects their reliability hugely from the users' point of view.

However, even the use of computer-aided image examinations is not always well received by dental experts, who tend to demand second opinions for these evaluations, if used, since they believe it leads to an inconclusive diagnosis [67]. It is observed that the low quality of the images is one of the reasons that hugely present influences in such resistance and difficult development of user-friendly tools, which can aid in computer-aided diagnostic popularization and more uses [67]. Consequently, for oral diseases, manual clinical evaluation is still the gold standard in diagnosis.

Other critical aspects related to the difficulties of the development of AI systems are the subjectivity in expert conclusions and the lack of standards for some oral disease diagnosis. The perception of caries severity may vary among experts; for instance, it is common that there is no agreement about the amounts of the teeth that must be compromised in order to consider a lesion as incipient or not. As observed by Dave [67], there is experts' judgment disagreement even for defining more concrete points when a patient presents an anatomic abnormality. This hugely affects the development of public databases with a ground truth that must be used for computational based solutions in the dental imaging area since there is a hidden feeling related to fear of peers' opinion about the correctness of the report made (*i.e.* based on the definition of the diagnosis), so the experts' annotations are considered by the physicians as a risk to them because it may not be considered correct by other experts and consequently it is almost impossible to find one or promote construction of databases. Moreover, this tends to restrict the applications that can be considered in the development of computer-based solutions, and this aspect could be reduced when there are widely accepted diagnosis standards as in the case of BIRADS degree for breast researches [68].

Truthfully, this lack of public available data is the main and critical challenge issue to be considered, faced, and resolved in dental imaging applications. Very few open datasets are available, with the "ISBI 2015 Grand Challenge in Dental X-ray Image Analysis" being the most popular one. Most works in literature use private datasets from their associated institutions, which can lead to bias since different institutions tend to target different populations [2, 3] and so their databases are more representative and trustworthy. For example, public emergency hospitals tend to attend more to vulnerable patients with more severe lesions and often more neglected oral health. A dataset acquired in private institutions may have a higher number of healthy patients, preventing demographically correct representations and promoting the construction of non-generic solutions due to sub-representation aspects.

Oral radiographs are also more susceptible to present artifacts since dental prostheses and implants are substantially more prevalent than other body regions. Artifacts greatly affect dental radiographic images preventing a quality diagnosis and influencing the signal patterns used in detection algorithms. Moreover, oral radiographs are also influenced by same phenomena that affect radiographs in general, as acquisition problems and noise resultant from limitations in the image

formation. The radio densities of some oral structures are difficult to detect in several oral diseases [69]. For example, bitewing images present a low sensitivity for both proximal and occlusive surfaces, and oral radiographs, in general, have a poor performance for detecting noncavity lesions.

Finally, oral diseases are heterogeneous and hard to model computationally (even impracticable sometimes) [2, 3, 6, 27, 70], restricting the application problems in which the proposed methods can be applied [32, 36, 43, 56, 58, 69].

References

1. R. Abdalla-Aslan, T. Yeshua, D. Kabla, I. Leichter, C. Nadler, An artificial intelligence system using machine-learning for automatic detection and classification of dental restorations in panoramic radiography: Automated detection and classification of panoramic dental restoration. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **130**(5), 593–602 (2020)
2. F. Schwendicke, T. Golla, M. Dreher, J. Krois, Convolutional neural networks for dental image diagnostics: A scoping review. *J. Dent.* **91**, 103226 (2019)
3. F. Schwendicke, W. Samek, J. Krois, Artificial intelligence in dentistry: Chances and challenges. *J. Dent. Res.* **99**(7), 769–774 (2020)
4. K. Hung, C. Montalvao, R. Tanaka, T. Kawai, M.M. Bornstein, The use and performance of artificial intelligence applications in dental and maxillofacial radiology: A systematic review. *Dentomaxillofacial Radiol.* **49**, 20190107 (2020)
5. K. Najarian, R. Splinter, *Biomedical Signal and Image Processing* (Taylor and Francis, 2012)
6. S.C. White, M.J. Pharoah, *Oral Radiology: Principles and Interpretation* (Elsevier Health Sciences, 2014)
7. P. Sukovic, Cone beam computed tomography in craniofacial imaging. *Orthodontics Craniofacial Res.* **6**, 31–36 (2003)
8. M.P. Bartold, 2006 periodontal tissues in health and disease: Introduction. *Periodontol.* **40**, 7–10 (2000)
9. R. Farina, A. Simonelli, A. Rizzi, L. Trombelli, Effect of smoking status on pocket probing depth and bleeding on probing following non-surgical periodontal therapy. *Minerva Stomatol.* **59**, 1–12 (2010)
10. K.S. Kornman, Mapping the pathogenesis of periodontitis: A new look. *J. Periodontol.* **8**, 1560–1568 (2008)
11. P.L. Lin, P.W. Huang, P.Y. Huang, H.C. Hsu, Alveolar bone-loss area localization in periodontitis radiographs based on threshold segmentation with a hybrid feature fused of intensity and the H-value of fractional Brownian motion model. *Comput. Methods Prog. Biomed.* **121**, 117–126 (2015)
12. P.L. Lin, P.Y. Huang, P.W. Huang, Automatic methods for alveolar bone loss degree measurement in periodontitis periapical radiographs. *Comput. Methods Prog. Biomed.* **148**, 1–11 (2017)
13. J.H. Lee, D.H. Kim, S.N. Jeong, S.H. Choi, Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. *J. Periodontal Implant Sci.* **48**, 114–123 (2018)
14. D.P. Carmody, S.P. McGrath, S.M. Dunn, P.F. van der Stelt, E. Schouten, Machine classification of dental images with visual search. *Acad. Radiol.* **8**, 1239–1246 (2001)
15. A. Mol, P.F. van der Stelt, Application of computer-aided image interpretation to the diagnosis of periapical bone lesions. *Dentomaxillofac. Radiol.* **21**, 190–194 (1992)
16. T. Ekert, J. Krois, L. Meinhold, K. Elhennawy, R. Emara, T. Golla, F. Schwendicke, Deep learning for the radiographic detection of apical lesions. *J. Endod.* **45**, 917–922 (2019)

17. J. Krois, T. Ekert, L. Meinhold, T. Golla, B. Kharbot, A. Wittemeier, F. Schwendicke, Deep learning for the radiographic detection of periodontal bone loss. *Sci. Rep.* **9**, 1–6 (2019)
18. M.K. Jeffcoat, I.C. Wang, M.S. Reddy, Radiographic diagnosis in periodontics. *Periodontol.* **7**, 54–68 (1995)
19. S.M. Pizer, E.P. Amburn, J.D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B.H. Romeny, J.B. Zimmerman, K. Zuiderveld, Adaptive histogram equalization and its variations. *Comput. Vis. Graph. Image Process* **39**, 355–368 (1987)
20. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2016), pp. 770–778. <https://doi.org/10.1109/CVPR.2016.90>
21. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, E. Dumitru, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2015), pp. 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>
22. I. Sutskever, J. Martens, G. Dahl, G. Hinton, On the importance of initialization and momentum in deep learning, in *International Conference on Machine Learning*. <http://dl.acm.org/citation.cfm?id=3042817.3043064>, (2013), pp. 1139–1147
23. Y. Bengio, Rmsprop and equilibrated adaptive learning rates for nonconvex optimization. *CoRR abs/1502.04390* (2015)
24. T. Chai, R.R. Draxler, Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **7**, 1247–1250 (2014)
25. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, Imagenet large scale visual recognition challenge. *Int. J. Comput. Vision* **115**, 211–252 (2015)
26. D. Powers, Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation. *J. Mach. Learn. Technol.* **2**, 37–63 (2007)
27. C. Anderson, M. Curzon, C. Van Loveren, C. Tatsi, M. Duggal, Sucrose and dental caries: A review of the evidence. *Obes. Rev.* **10**, 41–54 (2009)
28. M. Balakrishnan, R.S. Simmonds, J.R. Tagg, Dental caries is a preventable infectious disease. *Aust. Dent. J.* **45**, 235–245 (2000)
29. P.W. Caufield, A.L. Griffen, Dental caries: An infectious and transmissible disease. *Pediatr. Clin. N. Am.* **47**, 1001–1019 (2000)
30. W.J. Loesche, Role of streptococcus mutans in human dental decay. *Microbiol. Rev.* **50**, 353 (1986)
31. P.D. Marsh, Microbiology of dental plaque biofilms and their role in oral health and caries. *Dent. Clin.* **54**, 441–454 (2010)
32. A. Wenzel, H. Hintze, Perception of image quality in direct digital radiography after application of various image treatment filters for detectability of dental disease. *Dentomaxillofacial Radiol.* **22**, 131–134 (1993)
33. Z. Akarslan, M. Akdevelioglu, K. Gungor, H. Erten, A comparison of the diagnostic accuracy of bitewing, periapical, unfiltered and filtered digital panoramic images for approximal caries detection in posterior teeth. *Dentomaxillofacial Radiol.* **37**, 458–463 (2008)
34. S.B. Dove, W. McDavid, A comparison of conventional intra-oral radiography and computer imaging techniques for the detection of proximal surface dental caries. *Dentomaxillofacial Radiol.* **21**, 127–134 (1992)
35. M.Z. Booshehry, A. Davari, F.E. Ardakani, M.R.R. Nejad, Efficacy of application of pseudo-color filters in the detection of interproximal caries. *J. Dent. Res. Dent. Clin. Dent. Prospects* **4**, 79 (2010)
36. N. Karimian, H.S. Salehi, M. Mahdian, H. Alnajjar, et al., A Deep learning classifier with optical coherence tomography images for early dental caries detection, in *Proceedings of SPIE 10473: Lasers in Dentistry XXIV*, (2018), pp. 10473041–10473048
37. J.H. Lee, D.H. Kim, S.N. Jeong, S.H. Choi, Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J. Dent.* **77**, 106–111 (2018)

38. M. Sornam, M. Prabhakaran, Logit-based artificial bee colony optimization (Ib-abc) approach for dental caries classification using a back propagation neural network, in *Integrated Intelligent Computing, Communication and Security*, (2019), pp. 79–91
39. X. Xu, S. Xu, L. Jin, E. Song, Characteristic analysis of otsu threshold and its applications. *Pattern Recogn. Lett.* **32**, 956–961 (2011)
40. J. Serra, Morphological filtering: An overview. *Signal Process.* **38**, 3–11 (1994)
41. J. Schmidhuber, Deep learning in neural networks: An overview. *Neural Netw.* **61**, 85–117 (2015)
42. G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, C.I. Sanchez, A survey on deep learning in medical image analysis. *Med. Image Anal.* **42**, 60–88 (2017)
43. S. Kositbowornchai, M. Basiw, Y. Promwang, H. Moragorn, N. Sooksuntisakoonchai, Accuracy of diagnosing occlusal caries using enhanced digital images. *Dentomaxillofacial Radiol.* **33**, 236–240 (2004)
44. J. De Moraes, C.E. Sakakura, L.C.M. Loffredo, G. Scaf, Accuracy of zoomed digital image in the detection of periodontal bone defect: In vitro study. *Dentomaxillofacial Radiol.* **35**, 139–142 (2006)
45. R. Schulze, U. Heil, D. GroB, D.D. Bruellmann, E. Dranischnikow, U. Schwanecke, E. Schoemer, Artefacts in CBCT: A review. *Dentomaxillofacial Radiol.* **40**, 265–273 (2011)
46. Z. Wang, L. Jianwu, E. Mogendi, Removing ring artifacts in CBCT images via generative adversarial networks with unidirectional relative total variation loss. *Neural Comput. & Applic.* **31**, 5147–5158 (2019)
47. S. Chang, X. Chen, J. Duan, X. Mou, A CNN based hybrid ring artifact reduction algorithm for CT images. *IEEE Trans. Radiat. Plasma Med. Sci.* **5**(2), 253–260 (2020)
48. S. Xie, C. Yang, Z. Zhang, H. Li, Scatter artifacts removal using learning-based method for CBCT in IGRT system. *IEEE Access* **6**, 78031–78037 (2018)
49. Y. Zhang, Y. Hengyong, Convolutional neural network based metal artifact reduction in x-ray computed tomography. *IEEE Trans. Med. Imaging* **37**, 1370–1381 (2018)
50. C. Dong, C.C. Loy, K. He, X. Tang, Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**, 295–307 (2015)
51. A.E. Rad, M.S.M. Rahim, A. Rehman, R. Saba, Digital dental X-ray database for caries screening. *3D. Research* **7**, 18 (2016)
52. S.O. Arik, B.U. Ibragimov, L. Xing, Fully automated quantitative cephalometry using convolutional neural networks. *J. Med. Imaging* **4**, 014501 (2017)
53. Y. Song, X. Qiao, Y. Iwamoto, Y.W. Chen, Automatic cephalometric landmark detection on X-ray images using a deep-learning method. *Appl. Sci.* **10**, 2547 (2020)
54. J.J. Hwang, J.-H. Lee, S.-S. Han, Y.H. Kim, H.-G. Jeong, Y.J. Choi, Strut analysis for osteoporosis detection model using dental panoramic radiography. *Dentomaxillofac. Radiol.* **46**, 20170006 (2017)
55. M.S. Kavitha, A. Asano, A. Taguchi, M.-S. Heo, The combination of a histogram-based clustering algorithm and support vector machine for the diagnosis of osteoporosis. *Imaging Sci. Dent.* **43**, 153–161 (2013)
56. M.S. Kavitha, A. Asano, A. Taguchi, T. Kurita, M. Sanada, Diagnosis of osteoporosis from dental panoramic radiographs using the support vector machine method in a computer-aided system. *BMC Med. Imaging* **12**, 1 (2012)
57. M.S. Kavitha, P. Ganesh Kumar, S.-Y. Park, K.-H. Huh, M.-S. Heo, T. Kurita, Automatic detection of osteoporosis based on hybrid genetic swarm fuzzy classifier approaches. *Dentomaxillofac. Radiol.* **45**, 20160076 (2016)
58. F. Abdolali, R.A. Zoroofi, Y. Otake, Y. Sato, Automatic segmentation of maxillofacial cysts in cone beam CT images. *Comput. Biol. Med.* **72**, 108–119 (2016)
59. J. Mikulka, E. Gescheidtova, M. Kabrda, V. Perina, Classification of jaw bone cysts and necrosis via the processing of Orthopantomograms. *Radioengineering* **22**, 114–122 (2013)

60. I. Nurtanio, E.R. Astuti, I. Ketut Eddy Pumama, M. Hariadi, M.H. Purnomo, Classifying cyst and tumor lesion using support vector machine based on dental panoramic images texture features. *IAENG Int. J. Comput. Sci.* **40**, 29–37 (2013)
61. J.H. Lee, D.H. Kim, S.N. Jeong, Diagnosis of cystic lesions using panoramic and cone beam computed tomographic images based on deep learning neural network. *Oral Dis.* **26**, 152–158 (2020)
62. O. Kwon, T.H. Yong, S.R. Kang, J.E. Kim, Automatic diagnosis for cysts and tumors of both jaws on panoramic radiographs using a deep convolution neural network. *Dentomaxillofacial Radiol.* **49**, 20200185 (2020)
63. B. Kaplan, Evaluating informatics applications—Clinical decision support systems literature review. *Int. J. Med. Inf.* **64**, 15–37 (2001)
64. J.G. Anderson, J.J. Stephen, *Use and Impact of Computers in Clinical Medicine* (Springer Verlag, New York, 1987)
65. B. Kaplan, Evaluating informatics applications—Some alternative approaches: Theory, social interactionism, and call for methodological pluralism. *Int. J. Med. Inf.* **64**, 39–56 (2001)
66. E.A. Balas, S.A. Boren, Clinical trials and information interventions, in *Clinical Decision Support Systems: Theory and Practice*, (Springer, New York, 1999), pp. 199–216
67. M. Dave, K. Horner, Challenges in X-ray diagnosis: A review of referrals for specialist opinion. *Br. Dent. J.* **222**, 431–437 (2017)
68. A. Conci, A. Sanchez, P. Liatsis, H. Usuki, Signal processing techniques for detection of breast diseases. *Signal Process.* **93**, 2783–2784 (2013)
69. A. Goren, Current and future trends in dental radiography. *Oral Health Dent. Manage.* **14**, 41 (2015)
70. A. Yaji, S. Prasad, A. Pai, Artificial intelligence in dento-maxillofacial radiology. *Acta Sci. Dent. Sci.* **3**, 116–121 (2019)