International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
Article Received: *25 November2023* ***Revised:*** *12 December 2023* ***Accepted:*** *30 January 2024*

Deep Learning Methods for Tooth Detection and Classification in Various Dental Image Datasets: A Taxonomy and Future Directions

Samah AbuSalim
Department of Computer and Information Sciences
Universiti Teknologi PETRONAS
Seri Iskandar, Malaysia
samah_21000332@utp.edu.my

Nordin Zakaria
Department of Computer and Information Sciences
Universiti Teknologi PETRONAS
Seri Iskandar, Malaysia
nordinzakaria@utp.edu.my

Abdul Saboor
Department of Computer and Information Sciences
Universiti Teknologi PETRONAS
Seri Iskandar, Malaysia
abdul_19001745@utp.edu.my

Aarish Maqsood
Department of Agricultural Engineering
Bahauddin Zakariya University
Multan, Pakistan
aarishmaqsood@gmail.com

Said Jadid Abdulkadir
Department of Computer and Information Sciences
Universiti Teknologi PETRONAS
Seri Iskandar, Malaysia
saidjadid.a@utp.edu.my

Norehan Mokhtar
Dental Simulation and Virtual Learning Research Excellence Consortium, Department of Dental Science, Advanced Medical and Dental Institute
Universiti Sains Malaysia
Penang, Malaysia
norehanmokhtar@usm.my

Abstract—Deep learning approaches have made significant advancements in recent years, generating considerable interest in using them for medical image analysis. In dentistry, the precision of tooth detection and classification serves as the cornerstone of dental practice as it can identify the presence of dental abnormalizes at an early stage. This paper presents an exploration of the potential of deep learning methods for tooth detection and classification across a variety of dental imaging datasets including radiographs, cone-beam computed tomography (CBCT) scans, and photograph images. Convolutional Neural Networks (CNNs) have emerged as one of the most widely used and effective deep learning methods in the field of dental disease diagnosis and medical image analysis. The study aims to conceptualize how these models can effectively learn intricate tooth features, despite having variations in tooth morphology, image quality, and imaging techniques. It highlights the increasing role of deep learning in diagnosing dental diseases and emphasizes the importance of accurate tooth classification for effective treatment planning. The study reviews existing research in deep learning-based tooth classification, discusses challenges including dataset scarcity and model interpretability, and suggests future directions.

Keywords; deep learning; dental images; tooth detection; tooth classification; dental disease diagnosis

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
***Article Received:** 25 November2023 **Revised:** 12 December 2023 **Accepted:** 30 January 2024*

I. INTRODUCTION *(HEADING 1)*

According to statistics from the World Health Organization (WHO) [1], oral and dental disorders are the most prevalent noncommunicable health issues that impact people's well-being, resulting in discomfort, pain, and in some cases, even death. As a result, early detection plays a pivotal role in the effective treatment of oral diseases [2]. Tooth detection and classification tasks are a fundamental step and essential for dental abnormality diagnosis [3], tracking a patient's dental history [4], and reducing the potential for human error [5]. Dental images play an important part in clinical dental diagnosis because they properly represent tooth shape and distribution. They are quite useful in assisting dentists in developing precise and effective dental treatment programs.

In dental clinical practice, two primary types of images are commonly employed for diagnostic purposes: (i) extraoral CBCT images that offer a three-dimensional examination of the underlying teeth [6], and (ii) two-dimensional images, which include intraoral periapical, bitewing, panoramic, and photography images. These dental images illustrate several teeth simultaneously, such as peri-apical radiographs featuring 2-3 teeth, bitewings displaying 4-10 teeth, panoramic images revealing as many as 32 teeth, and intraoral photography showing varying numbers of teeth based on the level of magnification [7]. As a result, the classification of teeth plays a pivotal role in numerous research projects within oral medicine and forms the foundation for dental diagnosis and treatment driven by advanced deep learning methods.

Deep learning (DL) techniques have demonstrated remarkable effectiveness in various medical diagnostic tasks [8] [9]. DL models have been used in dentistry throughout the past few years to improve the diagnosis of dental diseases including dental cavities [10], gingivitis [11], and periodontitis bone loss [12]. Although these models perform well at identifying various dental diseases, they are unable to detect and identify teeth to provide specific diagnostic information on individual teeth, which restricts their practical utility [13]. Given the common practice of utilizing multiple types of dental imaging in dental procedures, developing a DL model capable of accurately identifying tooth numbers across diverse dental images becomes crucial. For instance, various DL models were proposed for automatic teeth segmentation [14] [15], and tooth detection and classification [16] [17]. However, these models were primarily designed for panoramic, periapical, and bitewing radiographs, which possess distinct characteristics compared to other types of dental images like CBCT and intra-oral photography images. Consequently, certain studies have been undertaken to address the specific challenges of detecting and classifying teeth in CBCT [18] [19] and intra-oral photo-graph images [20]. It's worth mentioning that each DL model has been designed and proven its reliability, usability, and applicability based on the image features which are different between dental images.

Among the numerous categories of DL models, Convolutional Neural Networks (CNNs) are widely applied in various dental applications [2]. Dentistry leverages CNNs for tasks like tooth detection, classification, disease diagnosis from dental images, and treatment planning. These models are valued for their ability to analyze dental images effectively and provide valuable insights for diagnosis and treatment decisions. Fig.1 illustrates the general steps taken in tooth detection: obtaining the image, pre-processing, feature extraction, and training and validation DL model for object or image classification.

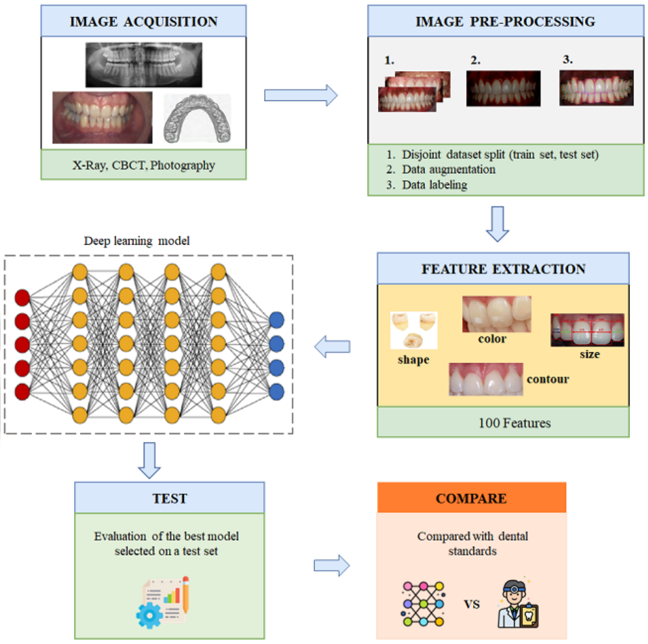


Figure 1. The process of tooth detection and classification.

Various CNN models such as AlexNet, You Only Look Once (YOLO), Faster Re-gion-based Convolutional Neural Network (Faster R-CNN), and Single Shot MultiBox Detector (SSD) have been utilized for tooth detection and classification. These models are trained using a diverse set of dental imaging datasets to acquire knowledge of distinguishing features and patterns for different types of teeth. However, the complexity of annotating dental images, the presence of occlusions and overlapping teeth, and the demanding computational resources required for training deep CNN models pose significant challenges [21]. Therefore, the deep learning-based teeth detection and classification models are covered in-depth in this study. This study focuses on the presenting of a thorough analysis of traditional DL methods for tooth detection and classification. The following research questions were established for the current study:

- What are the main characteristics of dental images that contribute to the complexities in the task of tooth detection?
- What are the main deep-learning methods for tooth detection and classification?

The study is divided into the following sections. Section 2 emphasizes the importance of accurate tooth detection and classification. An overview of the characteristics of dental image modalities is described in Section 3. Section 4 reviews literature related to deep learning models-based tooth detection and classification and establishes the need for the presented study. Additionally, the limitation of each study was mentioned. Section 5 discussed the challenges of deep learning models in tooth classification. Section 6 introduced various directions for future work in this field. Finally, section 7 concludes this research.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
***Article Received:** 25 November2023 **Revised:** 12 December 2023 **Accepted:** 30 January 2024*

II. THE IMPORTANCE OF TOOTH DETECTION AND CLASSIFICATION

The detection and classification of teeth is a critical first step in improving diagnostic accuracy, treatment planning efficiency, and overall patient care. This section demonstrates the advantages of teeth classification in a variety of domains.

- Accurate Diagnosis: tooth classification aids in distinguishing various tooth types and their specific position within the mouth cavity. This data helps dentists discover potential dental concerns including tooth decay or gum disease as well as detect the existence of extra or missing teeth.
- Orthodontics: the perfect alignment and location of each tooth are critical when planning orthodontic treatment. This lets orthodontists see the existing tooth arrangement, evaluate any misalignments or malocclusions, and calculate the best target alignment. This information also directs the selection of orthodontic tools (such as braces or clear aligners) and tooth movement planning, ensuring that teeth are relocated in a way that improves both functionality and looks.
- Dental Morphology Studies: researchers can investigate differences in tooth size, shape, and arrangement among various groups of people. This knowledge contributes to a better understanding of the genetic, environmental, and developmental factors that determine dental morphology, as well as insights into dental evolution and adaptation.
- Interactive Learning Tools: to construct interactive learning tools, tooth categorization data can be linked to educational platforms and software. These applications enable dental students and professionals to virtually investigate and manipulate segmented dental models, thereby improving their grasp of dental anatomy and pathology.

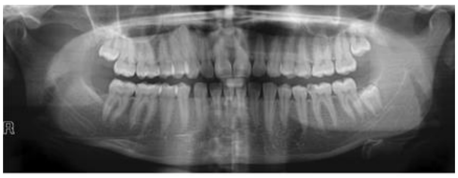
III. OVERVIEW OF THE CHARACTERISTICS OF DENTAL IMAGE MODALITIES

Dental image datasets serve as specialized tools for capturing images of oral and dental structures, facilitating diagnosis, treatment planning, and monitoring. These datasets encompass a range of options, such as intraoral radiography, panoramic radiography, CBCT, and intraoral photography.

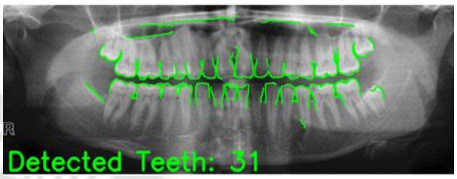
A. X-Ray Dental Image

X-ray images are widely used in dental treatments, such as panoramic radiographs (PR), periapical radiographs (PA), and bitewing radiographs (BW). PR gives a thorough view of the entire mouth since it includes all of the teeth in the upper and lower jaws in one image [22]. PA focuses on a specific area of the mouth, displaying a tooth from its crown to its root, along with the surrounding alveolar bone [23]. Conversely, BW is typically utilized for caries diagnosis and assessing bone levels, revealing only a partial view of the tooth and bone structures in the image [24].

These images offer immediate availability and involve relatively low levels of radiation exposure [25] [26]. They illustrate the detailed structures of teeth, including their crowns and roots, as well as information regarding tooth count, position, and any variations in crown and root morphology. Fig.2 shows a sample of the original image taken from the panoramic dental x-ray dataset [27] and the visualization of teeth detection.



(a)



(b)

Figure 2. Teeth detection result using X-ray image: (a) X-Ray original image [27] ; (b) Teeth detection result.

However, these images are known for having several drawbacks. They often suffer from issues like image overlaps, ghost images, and other artifacts, which have the potential to introduce errors during image processing [28]. Moreover, they suffer from noise and low contrast due to radiation limitations, often resulting in poorly defined or indistinguishable boundaries [29].

B. CBCT Dental Image

CBCT is an advanced imaging technique used in dentistry to create three-dimensional representations of dental structures, including teeth, jaws, and surrounding tissues [18]. The 3D information enables orthodontists to assess and treat complex malocclusions with improved accuracy and efficiency [30].

However, the analysis and interpretation of CBCT images present significant challenges and often require a considerable amount of time. The automatic classification of teeth is complicated due to limited datasets, the intricate appearance of the human skull, the adjacent teeth's edges are not clearly defined, and the mandible bone and teeth's intensity levels are comparable [31]. Additionally, dental CBCT scans tend to exhibit more noise and metal artifacts compared to conventional CT scans [32].

C. Intra-Oral Photograph Dental Image

Due to restricted access to dental care services and the significant financial burden of healthcare, regular dental visits are often unfeasible for many individuals. As a solution, intra-oral dental images can be used as a low-cost data acquisition method to help facilitate deep learning-based clinical studies, disease diagnostic procedures, and planning further treatment [33]. This kind of image offers several advantages, such as the lack of specialized equipment needed for data collecting, and cheap computational costs for image processing. Fig. 3 shows a sample of images taken from the ODSI-DB dataset [34] and the visualization of teeth detection.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
***Article Received:** 25 November2023 **Revised:** 12 December 2023 **Accepted:** 30 January 2024*

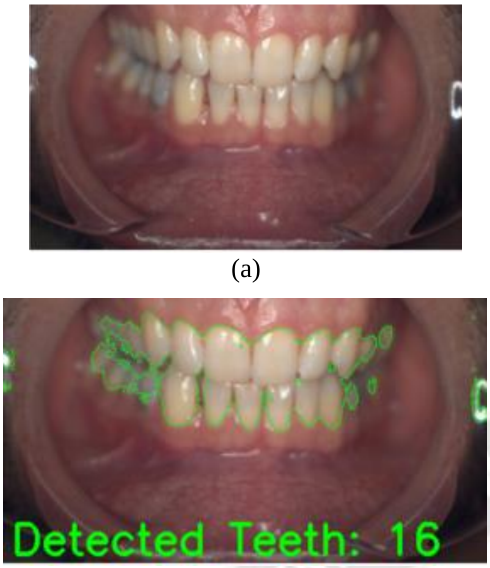


Figure 3. Teeth detection result using Intra-Oral image: (a) Intra-Oral original image [27]; (b) Teeth detection result.

IV. RELATED WORK

Extensive research has been carried out to investigate the application of DL-based object detection models for tooth classification using dental images. This section introduces an overview of DL model-based tooth classification in various dental imaging datasets.

A. *X-Ray Dental Image*

Oktay [36] introduced a CNN-based approach to detect teeth in dental panoramic X-ray images. The collected dataset included 100 dental panoramic images. The modified version of the AlexNet architecture was employed for tooth detection, utilizing multi-class classification. The approach was modified by incorporating the sliding window technique into the AlexNet architecture. The method achieved an accuracy exceeding 90%. Unfortunately, some inaccuracies were observed in tooth detection, particularly in neighboring teeth and tooth gums. It can be concluded that accurately outlining the exact contours of individual teeth using the sliding window approach becomes challenging. Keerthana et al. [37] developed an automated system aimed at classifying teeth into four categories: Incisor, Canine, Premolars, and Molars. The dataset used for training and evaluation consisted of 50 X-ray images for each tooth type. They used an algorithm called the vertical projection profile to determine the type of tooth root. Shape analysis of the photos was then used to classify the teeth. The proposed system achieved an overall accuracy of 92.54% in tooth classification. However, there were notable instances of misclassification. These misclassifications were attributed to the model's limited ability to effectively extract features from images with high levels of noise.

The Faster R-CNN model was used by Chen et al. [5] for tooth classification. Their approach involved several stages. First, they developed a filtering technique to remove overlapping bounding boxes generated by Faster R-CNN, ensuring that each box corresponded to a distinct tooth. Next, they implemented a neural network model to identify any missing teeth in the images. Finally, they introduced a rule-based module that utilized a tooth numbering system to align the labels of detected tooth boxes and correct any outcomes that don't follow some basic logic criteria. Their dataset comprises 1250 periapical images. The precision of tooth type numbering achieved in their experiments ranged from 71.5% to 91.7%. While Faster R-CNN effectively located tooth positions with a high Intersection over Union (IOU) value, it exhibited sensitivity to objects with missing features, such as broken teeth. Similarly, Tuzoff et al. [38] used CNNs for tooth detection and numbering in panoramic radiographs. They employed a dataset consisting of 1352 periapical images. Their proposed system was structured in two stages: first, they utilized the Faster R-CNN model to detect the presence of teeth in the radiographs, and then they employed a VGG-16 network to assign specific numbers to the maxillary and mandibular teeth within each image. The proposed system obtained a sensitivity and precision of almost 0.99. However, the model exhibited sensitivity to sudden changes in tooth features, which occasionally affected its classification accuracy, particularly in cases where teeth were adjacent to missing teeth. Furthermore, Mahdi et al. [39] utilized the Faster R-CNN based on ResNets to implement a teeth detection model. The study involved a dataset of 1,000 panoramic radiographs. In their research, the authors introduced a candidate optimization algorithm that considers both the positional patterns of detected boxes and the confidence scores. By combining the optimization algorithm with DL models, the proposed method achieves mAPs over 0.97 for both ResNet-50 and ResNet-101 models. However, to ensure a positive influence on the overall recognition performance, it is crucial to carefully select the optimization parameters according to the dataset.

Certain researchers have integrated the Inception v2 architecture within the Faster R-CNN. The inclusion of Inception v2, known for its capacity to extract features across multiple scales, can enhance the ability of Faster R-CNN to capture and represent intricate patterns and features in images. For example, in a study proposed by Bilgir et al. [40] Inception v2 Faster R-CNN was used to detect teeth in panoramic radiographs. The dataset employed in this research included 2482 images. The proposed model consisted of 22 deep layers, allowing it to extract features at various scales by using filters of different sizes within the same layer. The sensitivity, accuracy, and F-measure for the proposed model were 0.9559, 0.9652, and 0.9606 respectively. However, further research is needed to assess the model's performance in accurately estimating each type of tooth using more extensive datasets. Additionally, Görürgöz et al. [41] aimed to evaluate the effectiveness of a Faster R-CNN for tooth detection and numbering using periapical images. For their study, they collected a dataset of 1686 images. To prepare the data for analysis, they employed a pre-trained model (GoogLeNet Inception v3 CNN) and utilized transfer learning techniques during the dataset training process. A jaw classification model, region detection models, and a final algorithm that incorporates these models form the three main components of the proposed approach. The proposed model achieved a precision of 0.7812, an F1 score of 0.8720, and a sensitivity of 0.9867 in tooth detection and numbering. However, their results were not compared to state-of-the-art models, which could potentially introduce bias in the interpretation of the results. Another study proposed by Yasa et al. [42] to develop an automated model for tooth classification, utilizing the Faster R-CNN model. Their dataset consisted of 1125 bitewing radiographs from dental patients. The pre-

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
***Article Received:** 25 November2023 **Revised:** 12 December 2023 **Accepted:** 30 January 2024*

processing of the training data involved employing a pre-trained Google Net Inception v2 network through a technique called transfer learning. To enhance the model's accuracy, training was conducted using separate models for the right and left sides of the teeth, and two additional models were trained to specifically identify the right and left teeth. The evaluation of the system's performance measured using metrics like the F1 score, precision, and sensitivity, yielded high values of 0.9515, 0.9293, and 0.9748, respectively. However, the model had limitations in accurately recognizing the exact contours of the teeth.

Few studies have used the YOLO model for this task, such as Kaya et al. [43] conducted a study to evaluate the performance of YOLOv4 in automatically detecting and numbering teeth in pediatric panoramic radiographs. Their dataset included 4545 images. Importantly, they used the original structure of YOLOv4 without making any modifications. According to the findings, the model effectively detects and classifies both primary and permanent teeth, demonstrating mean Average Precision (mAP), F1-score, and mean Average Recall (mAR) of 92.22%, 0.91%, and 94.44 respectively.

Some researchers have conducted comparative studies between YOLO and Faster R-CNN models to evaluate their tooth detection and classification performance. For instance, Yilmaz DDS et al. [44] utilized both Faster R-CNN and YOLOv4 models using a dataset comprising 1200 panoramic radiograph images. The researchers employed the original configurations of the Faster R-CNN and YOLOv4 models to predict tooth types and then compared the performance of these models. The results indicate that YOLOv4 was faster in processing dental images. It yielded mAP of 99.90%, F1 score of 99.54%, and recall of 99.18%. In contrast, the Faster R-CNN method achieved slightly lower metrics with mAP of 93.67%, F1 score of 92.21%, and recall of 90.79%. The YOLOv4 method demonstrated superiority in working with small datasets, making it a more effective choice compared to other models. However, accurately detecting and localizing small or partially occluded teeth poses a challenge for YOLOv4. Additionally, Celik [45] evaluated the performance of two different detectors, Faster RCNN and YOLOv3 for identifying impacted third molar teeth in panoramic radiography. Their dataset consists of 440 panoramic radiographs. They also investigated the effects of combining the Faster RCNN detector with three other backbone designs, namely Res-Net50, AlexNet, and VGG16. The results indicated that when paired with the Res-Net50 backbone, the Faster RCNN detector achieved a mAP of 0.91. Notably, the YOLOv3 detector outperformed all others regarding detection efficiency, achieving a mAP of 0.96, and a recall of 0.93. However, it's important to note two limitations of the study. First, the dataset used was relatively limited. Second, the study focused specifically on mandibular third molars due to their higher prevalence, which may not cover all scenarios in dental image analysis.

TABLE I. A SUMMARY OF DEEP-LEARNING-BASED TOOTH CLASSIFICATION STUDIES USING X-RAY DENTAL IMAGES

Authors	Model	Results	Limitations
Oktay [36]	AlexNet	accuracy= 90%	sliding window approach is sensitive to missing objects
Keerthana et al. [37]	Vertical projection profile algorithm	accuracy= 92.5%	model's limited ability to extract features from noisy data
Chen et al. [5]	Faster R-CNN	precision ranged from 71.5% to 91.7%	model is sensitive to objects with missing features such as broken teeth
Tuzoff et al. [38]	Faster R-CNN	sensitivity= 0.9941, precision= 0.9945	model does not handle partial occlusion well
Mahdi et al. [39]	Faster R-CNN	(ResNet-50): mAPs= 0.974 (ResNet-101): mAPs= 0.981	model's high accuracy comes at a cost of prior knowledge of dataset features
Bilgir et al. [40]	Faster R-CNN	sensitivity= 0.9559 precision= 0.9652	small dataset targeting only four types of teeth
Görürgöz et al. [41]	Faster R-CNN	F1-score= 0.87 precision= 0.78 sensitivity= 0.98	ground truth was not verified by medical experts
Yasa et al. [42]	Faster R-CNN	F1 score= 0.95 precision= 0.92 sensitivity= 0.97	the model has a limited ability to detect edges
Kaya et al. [43]	YOLOv4	mAP= 92.22% mAR= 94.44% F1-score= 0.91	the study focused on pediatric dental images
Yilmaz DDS et al. [44]	YOLOv4, Faster R-CNN	(YOLOv4): precision=99.90% (Faster R-CNN): precision=93.67%	predetermined anchor box size limits the model's ability to recognize occluded teeth

Table 1 summarizes that CNN models, like Faster R-CNN, and YOLO have been utilized for tooth detection and classification. These models are trained on diverse datasets of dental images to acquire knowledge of distinguishing features and patterns for different types of teeth. However, teeth may be occluded or exhibit shape, size, and position variations presenting challenges in accurate detection and classification. Additionally, annotating large dental datasets with tooth labels is a time-consuming task requiring expert knowledge. To overcome these issues, various solutions have been proposed in previous studies. For instance, augmentation techniques can be applied to increase the number of training data, mitigating class imbalance and enhancing model generalization [46]. Utilizing Transfer Learning with pre-trained models from large general image datasets can fine-tune the models on dental images, utilizing learned features and reducing the need for extensive dental-specific labeled data [41]. Ensemble Models, involving the combination of multiple models or predictions from different models, offer an approach to tackle occlusions and variations, resulting in improved overall accuracy [47].

B. CBCT Dental Image

Miki et al. [46] proposed deep CNNs for classifying different tooth types in CBCT images. Their dataset consisted of 52 CBCT scans and adopted the AlexNet architecture for their neural network. To improve performance and prevent overfitting, they applied data augmentation techniques. They manually isolated individual teeth by cropping them from 2D slices extracted from CBCT images, using these cropped tooth images as input data for the network's tooth type classification. The study reported an average classification accuracy of 88.8% with augmented training data, but it had limitations: a relatively small evaluation dataset and an approach that treated slice images independently, potentially overlooking contextual information in adjacent slices. Additionally, Du et al. [48]

introduced a teeth-detection approach designed for processing CBCT images, utilizing a dataset of 25 dental CBCT scans. Their method involved several key steps including combining YOLOv3 and multi-level teeth detection techniques within Region of Interest (ROI) regions to identify bounding boxes for each tooth. The results indicated that the proposed method significantly reduced training and prediction times by 80% and 62%, respectively, compared to Faster R-CNN. Additionally, the proposed method achieved a high Object Inclusion Ratio (OIR) metric of 96.27%, whereas Faster R-CNN attained an OIR of 91.40%. However, there were some situations where the identified bounding boxes were very close to the actual tooth but not entirely accurate. Another study proposed by Jang et al. [19] presented a fully automated method for identifying and segmenting individual teeth in 3D using dental CBCT images. Their approach tackled the challenge of distinguishing a single tooth from neighboring teeth and the surrounding alveolar bone in CBCT scans by employing a deep learning-based hierarchical multi-step model. To begin with, the model automatically generates panoramic images of both the upper and lower jaws in the initial stage. These 2D panoramic images are then used to identify 2D individual teeth and capture specific ROIs that represent both loose and tight areas within the 3D individual teeth. The experimental findings highlight the efficacy of the proposed approach, with an F1-score of 93.35% achieved for tooth identification. Furthermore, Younis et al. [49] aim to enhance the clarity of CBCT radiographs through a series of pre-processing stages to facilitate a more accurate and efficient tooth identification process. They employed contourlet transformation for feature extraction from individual CBCT images. The outcomes of this feature extraction were then subjected to a novel hybrid particle swarm optimization (PSO) technique referred to as the "contourlet PSO" algorithm (CPSO). This newly developed CPSO algorithm operates at a higher speed and achieves greater precision than conventional PSO methods. Upon applying the proposed CPSO algorithm to 100 CBCT radiographs, a high detection ratio of 98% was attained.

TABLE II. A SUMMARY OF DEEP-LEARNING-BASED TOOTH CLASSIFICATION STUDIES USING CBCT DENTAL IMAGES

Authors	Model	Results	Limitations
Miki et al. [46]	CNN	accuracy= 88.8%	very small dataset was employed and sliced images may lead to missing contextual information
Du et al. [48]	YOLOv3	OIR= 96.27%	model shows sensitivity in identifying the center of the tooth resulting in accurate detection
Jang et al. [19]	Hierarchical multi-step DL model	F1-score= 93.35	model faces difficulties in distinguishing between first pre-molars and wisdom teeth
Younis et al. [49]	Contourlet PSO	detection ratio= 98%	model struggles to accurately differentiate between cases of teeth loss

Table 2 summarizes that CBCT provides extensive three-dimensional volumetric information encompassing all oral tissues, including teeth [18]. It offers a comprehensive visualization of the 3D anatomical structure, facilitating the assessment of patients with malocclusion, tooth shape, and positional distribution [30]. However, the automatic classification of teeth in these images is complicated due to limited datasets, the intricate appearance of the human skull, indistinct boundaries between neighboring teeth, and similar intensity values between teeth and the mandible bone [31]. To deal with these issues, several methods have been suggested in earlier studies. For instance, a hierarchical multi-level deep learning model as the hierarchical representation facilitates the capture of contextual information associated with tooth structure, neighboring teeth, and the overall configuration of the dental arch [50].

V. THE CHALLENGES OF TOOTH DETECTION AND CLASSIFICATION TASK

This section presents the open challenges in the reviewed studies for tooth classification based on dental image datasets and deep learning models.

A. Challenges of Dental Image Datasets for Teeth Detection and Classification Task

- Digital radiograph images are extensively used in the diagnosis of dental diseases and tooth detection and classification. However, radiographs frequently exhibit image overlaps, ghost images, and other artifacts, which can introduce errors in image processing [28].
- Digital radiograph images suffer from noise and low contrast due to radiation limitations, often resulting in poorly defined or indistinguishable tooth boundaries [29]. Additionally, lighting inconsistencies may manifest as variations in brightness and contrast across the image. These fluctuations in lighting can pose challenges when it comes to accurately detecting and categorizing teeth within the image [29].
- CBCT provides extensive three-dimensional volumetric information encompassing all oral tissues, including teeth [18]. However, the automatic classification of individual teeth in CBCT images is complicated due to limited datasets, the intricate appearance of the human skull, indistinct boundaries between neighboring teeth, and similar intensity values between teeth and the mandible bone [31].
- CBCT scans tend to exhibit more noise and metal artifacts compared to conventional Computed Tomography (CT) scans [32]. This is due to the unique challenges posed by the presence of metal objects in the mouth and the specific way the CBCT technology works. These artifacts can potentially make the interpretation of dental CBCT scans more challenging for deep-learning models.
- Intra-oral dental images can be used as a low-cost data acquisition method to help facilitate deep learning-based clinical studies, disease diagnostic procedures, and planning further treatment [33]. However, the identification and detection of individual teeth in these images present certain challenges, such as partial occlusion, overlapping teeth, and variations in illumination [21]. Additionally, a significant issue is the limited availability of comprehensive intra-oral image datasets.

- Due to the tooth's uneven texture or color distribution, classifying teeth in intra-oral photography imaging can be challenging [51]. Additionally, there are significant variations in surface appearance with the same type of tooth among individuals despite some shared morphological markers for identifying tooth type.

B. Challenges of Deep Learning Methods for Teeth Detection and Classification Task

- Deep learning models typically require large amounts of labeled data for training. However, in the medical domain, acquiring annotated medical images, especially for rare conditions or specific populations, can be challenging and time-consuming [52].
- Although the deep CNN method has made considerable progress in classifying medical images, current deep learning techniques tend to concentrate more on regions with small differences in object size and are therefore less able to adjust to object variations and misaligned teeth [53].
- For tooth classification using deep learning, an interpretable model would provide insights into which specific aspects of a dental image led the model to classify a tooth in a particular way. For instance, the model might focus on certain patterns in tooth shape, color, or texture that align with established dental diagnostic criteria [54]. Interpretability could involve techniques such as attention maps, saliency maps, and gradient-based methods to highlight the regions of the image that were most influential in the classification decision [55].
- Training deep learning models involves optimizing a large number of parameters in complex neural network architectures through a process called gradient descent. This optimization process requires performing numerous calculations on vast amounts of data, and as a result, it demands significant computational resources [56].
- CNNs are sensitive to the location of features in an image. If a significant portion of a tooth's features is missing due to a broken tooth, the CNN might struggle to recognize the tooth correctly. This is because the broken tooth's missing features can disrupt the pattern the model has learned [5].
- In the context of image analysis, "contextual information" refers to the surrounding details and elements in an image that provide additional clues about the objects of interest. For example, when classifying a tooth, the con-text might include neighboring teeth, gum tissue, and the overall arrangement of the teeth in the mouth. When a tooth is partially occluded or obscured by other teeth or structures, the model might not have access to enough contextual information to make an accurate classification [38].
- Severe malocclusion can significantly alter the expected arrangement of teeth in the oral cavity. When a tooth detection model is trained on images with normal alignment, it might struggle to accurately identify teeth in cases of extreme misalignment. The model might not have encountered such variations during training and therefore might not recognize the irregular tooth arrangement [48].

VI. FUTURE DIRECTIONS OF DEEP LEARNING IN TOOTH CLASSIFICATION

The future of deep learning in tooth classification is likely to involve advancements in multiple aspects, driven by ongoing research, technological innovations, and the increasing integration of AI in healthcare. Here are some potential future directions for deep learning in tooth classification:

- Improved Data Collection and Annotation. Efforts should continue to focus on gathering diverse and well-annotated dental image datasets. This includes collecting images from various populations, age groups, and dental conditions. High-quality annotations will be essential for training accurate and robust models.
- Proposition of a deep learning model-based tooth classification in Intra-Oral photography images. Researchers have dedicated considerable efforts to conducting extensive research on tooth classification in various dental images. However, classifying individual tooth types in intra-oral dental images remains a formidable challenge due to the presence of complex and diverse structures within these images [35].
- Hierarchical-based approach for tooth classification. According to Chen et al. [57], a hierarchical classification structure implies a greater ability for tooth type recognition when considering the logical hierarchy of dental categories. Therefore, the use of hierarchical classification will increase the probability that an entire object or part of an object is contained in a particular hierarchy scale.
- Develop a model that can accurately classify malocclusion types in dental images. By leveraging various views of the same image, using pre-learned features, and focusing on relevant regions, the model can better handle misalignments and occlusions caused by malocclusion, leading to more accurate and reliable classifications.

VII. CONCLUSION

This paper significantly contributes to the field of dentistry by providing a comprehensive overview of the integration of deep learning techniques into dental image analysis. The paper starts by giving a thorough description of how deep learning approaches have been incorporated into dentistry. It covers the latest advancements and research in the domain, offering valuable insights for both practitioners and researchers. The study underlines the critical importance of accurate tooth detection and classification, as it constitutes a fundamental preliminary step in numerous dental analyses and treatment procedures. Additionally, this paper has discussed various CNN models for tooth detection and classification with each model having its strengths and weaknesses. The thorough selection of the categorization model is crucial for attaining the best results. It has been identified that the Faster R-CNN model has demonstrated promising results, particularly in the context of dental X-ray images. However, it also highlights the challenges posed by occluded teeth and variations in shape, size, and position, motivating further research into addressing these issues.

Acknowledgment

The authors would like to express their sincere gratitude and appreciation to Universiti Teknologi PETRONAS for their generous provision of resources and materials, which were instrumental in the successful completion of this research work.

References

[1] Peres MA, Macpherson LMD, Weyant RJ, Daly B, Venturelli R, Mathur MR, et al. Oral diseases: a global public health challenge. The Lancet. 2019 Jul;394(10194):249–60.

[2] AbuSalim S, Zakaria N, Islam MR, Kumar G, Mokhtar N, Abdulkadir SJ. Analysis of Deep Learning Techniques for Dental Informatics: A Systematic Literature Review. Healthcare. 2022 Oct;10(10):1892.

[3] Lee CT, Kabir T, Nelson J, Sheng S, Meng HW, Van Dyke TE, et al. Use of the deep learning approach to measure alveolar bone level. J Clin Periodontol. 2022 Mar;49(3):260–9.

[4] Dianat O, Gupta S, Price JB, Mostoufi B. Guided Endodontic Access in a Maxillary Molar Using a Dynamic Navigation System. J Endod. 2021 Apr 1;47(4):658–62.

[5] Chen H, Zhang K, Lyu P, Li H, Zhang L, Wu J, et al. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. Sci Rep. 2019 Mar 7;9(1):3840.

[6] Fan W, Zhang J, Wang N, Li J, Hu L. The Application of Deep Learning on CBCT in Dentistry. Diagnostics. 2023 Jan;13(12):2056.

[7] Krois J, Schneider L, Schwendicke F. Impact of Image Context on Deep Learning for Classification of Teeth on Radiographs. J Clin Med. 2021 Jan;10(8):1635.

[8] Biabani S a. A, Tayyib NA. A Review on the Use of Machine Learning Against the Covid-19 Pandemic. Eng Technol Appl Sci Res. 2022 Feb 12;12(1):8039–44.

[9] Hamdani M, Youcefi M, Rabehi A, Nail B, Douara A. Design and Implementation of a Medical TeleMonitoring System based on IoT. Eng Technol Appl Sci Res. 2022 Aug 7;12(4):8949–53.

[10] Bhattacharjee N. Automated Dental Cavity Detection System Using Deep Learning and Explainable AI. AMIA Summits Transl Sci Proc. 2022 May 23;2022:140–8.

[11] Li W, Liang Y, Zhang X, Liu C, He L, Miao L, et al. A deep learning approach to automatic gingivitis screening based on classification and localization in RGB photos. Sci Rep. 2021 Aug 19;11(1):16831.

[12] Ryu J, Lee DM, Jung YH, Kwon O, Park S, Hwang J, et al. Automated Detection of Periodontal Bone Loss Using Deep Learning and Panoramic Radiographs: A Convolutional Neural Network Approach. Appl Sci. 2023 Jan;13(9):5261.

[13] Kabir T, Lee CT, Chen L, Jiang X, Shams S. A comprehensive artificial intelligence framework for dental diagnosis and charting. BMC Oral Health. 2022 Nov 9;22(1):480.

[14] Nader R, Smorodin A, De La Fournière N, Amouriq Y, Autrusseau F. Automatic teeth segmentation on panoramic X-rays using deep neural networks. In: 2022 26th International Conference on Pattern Recognition (ICPR). 2022. p. 4299–305.

[15] Rubiu G, Bologna M, Cellina M, Cè M, Sala D, Pagani R, et al. Teeth Segmentation in Panoramic Dental X-ray Using Mask Regional Convolutional Neural Network. Appl Sci. 2023 Jan;13(13):7947.

[16] Chung M, Lee J, Park S, Lee M, Lee CE, Lee J, et al. Individual Tooth Detection and Identification from Dental Panoramic X-Ray Images via Point-wise Localization and Distance Regularization. Artif Intell Med. 2021 Jan;111:101996.

[17] Muresan MP, Barbura AR, Nedevschi S. Teeth Detection and Dental Problem Classification in Panoramic X-Ray Images using Deep Learning and Image Processing Techniques. In: 2020 IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP). 2020. p. 457–63.

[18] Cui Z, Li C, Wang W. ToothNet: Automatic Tooth Instance Segmentation and Identification From Cone Beam CT Images. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. p. 6361–70.

[19] Jang TJ, Kim KC, Cho HC, Seo JK. A Fully Automated Method for 3D Individual Tooth Identification and Segmentation in Dental CBCT. IEEE Trans Pattern Anal Mach Intell. 2022 Oct;44(10):6562–8.

[20] AbuSalim S, Zakaria N, Mostafa SA, Hooi YK, Mokhtar N, Abdulkadir SJ. Multi-Granularity Tooth Analysis via Faster Region-Convolutional Neural Networks for Effective Tooth Detection and Classification. Int J Adv Comput Sci Appl [Internet]. 2023 [cited 2023 Jul 30];14(6). Available from: http://thesai.org/Publications/ViewPaper?Volume=14&Issue=6&Code=IJACSA&SerialNo=78

[21] AbuSalim S, Zakaria N, Mokhtar N, Mostafa SA, Abdulkadir SJ. Data Augmentation on Intra-Oral Images Using Image Manipulation Techniques. In: 2022 International Conference on Digital Transformation and Intelligence (ICDI). 2022. p. 117–20.

[22] Al Kheraif AA, Wahba AA, Fouad H. Detection of dental diseases from radiographic 2d dental image using hybrid graph-cut technique and convolutional neural network. Measurement. 2019 Nov 1;146:333–42.

[23] Meusburger T, Wülk A, Kessler A, Heck K, Hickel R, Dujic H, et al. The Detection of Dental Pathologies on Periapical Radiographs—Results from a Reliability Study. J Clin Med. 2023 Jan;12(6):2224.

[24] Lee S, Oh S il, Jo J, Kang S, Shin Y, Park J won. Deep learning for early dental caries detection in bitewing radiographs. Sci Rep. 2021 Aug 19;11(1):16807.

[25] Geetha V, Aprameya KS, Hinduja DM. Dental caries diagnosis in digital radiographs using back-propagation neural network. Health Inf Sci Syst. 2020 Jan 3;8(1):8.

[26] Kumar T, Ponnusamy R. Robust Medical X-Ray Image Classification by Deep Learning with Multi-Versus Optimizer. Eng Technol Appl Sci Res. 2023 Aug 9;13(4):111406–11.

[27] Abdi A, Kasaei S. Panoramic Dental X-rays With Segmented Mandibles. 2020 Jul 2 [cited 2023 Sep 30];2. Available from: https://data.mendeley.com/datasets/hxt48yk462/2

[28] Verma D, Puri S, Prabhu S, Smriti K. Anomaly detection in panoramic dental x-rays using a hybrid Deep Learning and Machine Learning approach. In: 2020 IEEE REGION 10 CONFERENCE (TENCON). 2020. p. 263–8.

[29] Fariza A, Arifin AZ, Astuti ER. Automatic Tooth and Background Segmentation in Dental X-ray Using U-Net Convolution Network. In: 2020 6th International Conference on Science in Information Technology (ICSITech). 2020. p. 144–9.

[30] Tian S, Dai N, Zhang B, Yuan F, Yu Q, Cheng X. Automatic Classification and Segmentation of Teeth on 3D Dental Model Using Hierarchical Deep Learning Networks. IEEE Access. 2019;7:84817–28.

[31] Kwak GH, Kwak EJ, Song JM, Park HR, Jung YH, Cho BH, et al. Automatic mandibular canal detection using a deep convolutional neural network. Sci Rep. 2020 Mar 31;10(1):5711.

[32] Qiu B, Wel H van der, Kraeima J, Glas HH, Guo J, Borra RJH, et al. Robust and Accurate Mandible Segmentation on Dental CBCT Scans Affected by Metal Artifacts Using a Prior Shape Model. J Pers Med [Internet]. 2021 May [cited 2023 Jul 3];11(5). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8147374/

[33] Fadhillah ED, Bramastagiri PC, Sigit R, Sukaridhoto S, Brahmanta A, Dewantara BSB. Smart Odontogram: Dental Diagnosis of Patients Using Deep Learning. In: 2021 International Electronics Symposium (IES). 2021. p. 532–7.

[34] Hyttinen J, Fält P, Jäsberg H, Kullaa A, Hauta-Kasari M. Oral and Dental Spectral Image Database—ODSI-DB. Appl Sci. 2020 Jan;10(20):7246.

[35] Ragodos R, Wang T, Padilla C, Hecht JT, Poletta FA, Orioli IM, et al. Dental anomaly detection using intraoral photos via deep learning. Sci Rep. 2022 Jul 8;12(1):11577.

[36] Betul Oktay A. Tooth detection with Convolutional Neural Networks. In: 2017 Medical Technologies National Congress (TIPTEKNO). 2017. p. 1–4.

[37] Keerthana KM, Rajeshwari B, Keerthi S, Menon HP. Classification of tooth type from dental X-ray image using projection profile analysis. In: 2017 International Conference on Signal Processing and Communication (ICSPC). 2017. p. 394–8.

[38] Tuzoff DV, Tuzova LN, Bornstein MM, Krasnov AS, Kharchenko MA, Nikolenko SI, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. Dentomaxillofacial Radiol. 2019 Mar 2;48(4):20180051.

[39] Mahdi FP, Motoki K, Kobashi S. Optimization technique combined with deep learning method for teeth recognition in dental panoramic radiographs. Sci Rep. 2020 Nov 6;10(1):19261.

[40] Bilgir E, Bayrakdar İŞ, Çelik Ö, Orhan K, Akkoca F, Sağlam H, et al. An artificial ıntelligence approach to automatic tooth detection and numbering in panoramic radiographs. BMC Med Imaging. 2021 Aug 13;21:124.

[41] Görürgöz C, Orhan K, Bayrakdar IS, Çelik Ö, Bilgir E, Odabaş A, et al. Performance of a convolutional neural network algorithm for tooth detection and numbering on periapical radiographs. Dentomaxillofacial Radiol. 2022 Mar 1;51(3):20210246.

[42] Yasa Y, Çelik Ö, Bayrakdar IS, Pekince A, Orhan K, Akarsu S, et al. An artificial intelligence proposal to automatic teeth detection and numbering in dental bite-wing radiographs. Acta Odontol Scand. 2021 May;79(4):275–81.

[43] Kaya E, Gunec HG, Gokyay SS, Kutal S, Gulum S, Ates HF. Proposing a CNN Method for Primary and Permanent Tooth Detection and Enumeration on Pediatric Dental Radiographs. J Clin Pediatr Dent. 2022 Jul 1;46(4):293–8.

[44] Yilmaz S, Tasyurek M, Amuk M, Celik M, Canger EM. "Developing Deep Learning Methods for Classification of Teeth in Dental Panoramic Radiography." Oral Surg Oral Med Oral Pathol Oral Radiol [Internet]. 2023 Mar 30 [cited 2023 Jun 10]; Available from: https://www.sciencedirect.com/science/article/pii/S2212440323001165

[45] Celik ME. Deep Learning Based Detection Tool for Impacted Mandibular Third Molar Teeth. Diagnostics. 2022 Apr;12(4):942.

[46] Miki Y, Muramatsu C, Hayashi T, Zhou X, Hara T, Katsumata A, et al. Classification of teeth in cone-beam CT using deep convolutional neural network. Comput Biol Med. 2017 Jan 1;80:24–9.

[47] Kim C, Kim D, Jeong H, Yoon SJ, Youm S. Automatic Tooth Detection and Numbering Using a Combination of a CNN and Heuristic Algorithm. Appl Sci. 2020 Jan;10(16):5624.

[48] Du M, Wu X, Ye Y, Fang S, Zhang H, Chen M. A Combined Approach for Accurate and Accelerated Teeth Detection on Cone Beam CT Images. Diagnostics. 2022 Jul 10;12(7):1679.

[49] Younis HA, Hammadi DS, Younis AN. Identify tooth cone beam computed tomography based on contourlet particle swarm optimization. IAES Int J Artif Intell IJ-AI. 2022 Mar 1;11(1):397–404.

[50] Muramatsu C, Morishita T, Takahashi R, Hayashi T, Nishiyama W, Ariji Y, et al. Tooth detection and classification on panoramic radiographs for automatic dental chart filing: improved classification by multi-sized input data. Oral Radiol. 2021 Jan;37(1):13–9.

[51] Wu C, Bradley D, Garrido P, Zollhöfer M, Theobalt C, Gross M, et al. Model-based teeth reconstruction. ACM Trans Graph. 2016 Dec 5;35(6):220:1-220:13.

[52] Luca AR, Ursuleanu TF, Gheorghe L, Grigorovici R, Iancu S, Hlusneac M, et al. Impact of quality, type and volume of data used by deep learning models in the analysis of medical images. Inform Med Unlocked. 2022 Jan 1;29:100911.

[53] Xie X, Pan X, Zhang W, An J. A context hierarchical integrated network for medical image segmentation. Comput Electr Eng. 2022 Jul 1;101:108029.

[54] Huang C, Wang J, Wang S, Zhang Y. A review of deep learning in dentistry. Neurocomputing. 2023 Oct 14;554:126629.

[55] Reyes M, Meier R, Pereira S, Silva CA, Dahlweid FM, Tengg-Kobligk HV, et al. On the Interpretability of Artificial Intelligence in Radiology: Challenges and Opportunities. Radiol Artif Intell. 2020 May 1;2(3):e190043.

[56] Habib G, Qureshi S. Optimization and acceleration of convolutional neural networks: A survey. J King Saud Univ - Comput Inf Sci. 2022 Jul 1;34(7):4244–68.

[57] Chen Q, Huang J, Salehi HS, Zhu H, Lian L, Lai X, et al. Hierarchical CNN-based occlusal surface morphology analysis for classifying posterior tooth type using augmented images from 3D dental surface models. Comput Methods Programs Biomed. 2021 Sep 1;208:106295.