Multi Oral Disease Classification from Panoramic Radiograph using Transfer Learning and XGBoost

Priyanka Jaiswal[1], Dr. Vijay Katkar[2], Dr. S.G. Bhirud[3]
Professor[3]
Department of CE & IT, Veermata Jijabai Technical Institute, Mumbai-400019, India[1, 3]
Department of IT, YCCE, Nagpur[1]
Associate Professor, Department of Computer Engineering, Marwadi University, Rajkot, India[2]

Abstract*—The subject of oral healthcare is a crucial research field with significant technological development. This research examines the field of oral health care known as dentistry, a branch of medicine concerned with the anatomy, development, and disorders of the teeth. Good oral health is essential for speaking, smiling, testing, touching, digesting food, swallowing, and many other aspects, such as expressing a variety of emotions through facial expressions. Comfort in doing all these activities contributes to a person's self-confidence. For diagnosing multiple oral diseases at a time panoramic radiograph is used. Oral healthcare experts are important to appropriately detect and classify disorders. This automated approach was developed to eliminate the overhead of experts and the time required for diagnosis. This research is based on a self-created dataset of 500 images representing six distinct diseases in 46 possible combinations. Tooth wear, periapical, periodontitis, tooth decay, missing tooth, and impacted tooth are all examples of diseases. This system is developed using the concept of transfer learning with the use of a different pre-trained network such as "ResNet50V2", "ResNet101V2", "MobileNetV3Large", "MobileNetV3Small", "MobileNet", "EfficientNetB0", "EfficientNetB1", and "EfficientNetB2" with XGBoost and to get the final prediction The images in the dataset were divided into 80% training and 20% images for testing. To assess the performance of this system, various measuring metrics are used. Experiments revealed that the proposed model detected Tooth wear, periapical, periodontitis, tooth decay, missing tooth, and impacted tooth with an accuracy of 91.8%, 92.2%, 92.4%, 93.2%, 91.6%, and 90.8%, respectively.

***Keywords*—Panoramic radiograph; dentistry; deep learning; ensemble classifier; multi-disease classification and prediction; oral diseases; weighted ensemble module; XGBoost**

I. Introduction

According to the WHO survey, the population of seniors aged 65 and more has surpassed 150 million, and the population of patients with chronic conditions has surpassed 400 million. Due to the increase in periodontal diseases in the 35-44 age group, the percentage of infection is higher in the 65-74 age group. Additionally, proper oral health is essential for maintaining a healthy heart and digestive system. Aging and mouth problems have a long history of interdependence; a lack of attention to oral health might predispose older individuals to other health conditions, such as malnutrition, heart-related issues, Digestion related problems, etc. [2] This has increased the demand for automated medical and healthcare solutions that are secure and of superior quality. Which must be useful for the early detection and classification of the present disorders of an individual. To perform diagnosis of present tooth-related diseases in the mouth, Currently, dental radiography plays a crucial role in diagnosing the condition of oral health. Due to its accurate confirmation of clinical findings, dental panoramic radiography (DPR) images are attracting a rising amount of attention in the diagnostic process. Typically, three types of dental X-ray radiographs are used for diagnosis in dental imaging. Dental radiographs consist of bitewing, periapical, and panoramic types. Bitewing focuses mostly on the coronal half of teeth, periapical encompasses the apical region of teeth, and panoramic X-ray scans consider the entire upper and lower jaw, including teeth and supporting bones. For disease localization diagnosis, periapical radiography is suggested. But when the disease is widespread, panoramic X-rays are used to obtain a thorough perspective for diagnosis and therapy planning. They are easily standardizable and reproducible.

Fig. 1. Different types of dental radiography

Due to the unspecified quality of x-ray images, varying tooth sizes, and other factors, numerous challenges are associated with automatically separating teeth from dental radiograph images. Fig. 1 depicts the three main types of x-ray images of teeth: bitewing, periapical, and panoramic teeth. The dentist manually clarifies radiographs by identifying each tooth and the corresponding issue. But if the x-ray radiography is unclear, it can lead to misinterpretation during analysis [1]. Conventionally, a dentist uses dental radiographs and a clinical evaluation of the patient to make a diagnosis, based on the available infrastructure and expertise. These numerous methods encourage researchers to use and build new machine learning, deep learning, and dental image comprehension techniques more effectively [2][3]. This approach will help to classify multiple oral diseases automatically.

Image processing is currently the most prominent field in digital healthcare applications; it is also utilized well in the medical environment to analyze illness more precisely [4]. As a result of these developments, medical image processing has become prevalent in both medical and health sectors in analyzing diseases with precise sensitivity and measurement [5]. Similarly, medical image processing is applied to dental information to identify and classify disease types and severity [6]. To identify dental problems, all factors must be considered [7]; this is the predominant trend in the detection of dental diseases. The kind and severity of the condition differed between patients [8] due to inconsistencies in the data. Several mathematical models have been created to predict, identify, assess, and monitor the early oral health of teeth [9]. However, the issues are not addressed because of the complexity of the images [10]. Furthermore, dental images differ depending on the patient's genes and body condition. Predicting dental disease is thus a difficult task. Furthermore, a mathematical model has carried out the different cavity analysis processes practically. The cavity severity analysis procedure has varied depending on the patient's body and tooth conditions [11]. Finally, the current tooth diseases and their causes were identified. Nonetheless, those model features are insufficient for real-time data. Because actual time tooth X-ray data contains whiter spaces, and some tooth X-rays have tooth clips [12] disease identification becomes more difficult. Furthermore, those types of images have required more additional features to identify the affected tooth [13]. As a result, those approaches failed to identify the diseased tooth. Furthermore, a standard filtering process was used to remove the noisy features in the trained data, which increased the cost of the resource [29]. Furthermore, when the dataset images were complex, the filtering results were poor. In other domain concepts, deep learning is used to efficiently estimate remote heart rates and to classify breast lesions in mammography using a conventional neural network, which helped us understand the running complexity of different machine intelligence techniques [14]. These issues prompted the current study to develop optimized deep networks with filtering parameters to obtain more precise disease classification results [17].

The current article has descended into a new solution for maximizing disease classification by utilizing the different pre-trained deep learning models. The proposed study has a step-down in the multi-disease classification model, which is defined as multiple parameters in the proposed framework being upgraded to classify multi-disease features. The following section goes over the specifics and importance of detecting tooth wear, tooth decay(cavity), impacted tooth, missing tooth, periapical, and periodontitis at an early stage. The image of a patient suffering from tooth wear is shown in Fig. 2.

A. The Necessity of Detecting Tooth Wear in Dentistry

Teeth have specific anatomy consisting of several cusps and fissures to serve the function of chewing food like a mortar and pestle where the approximating cusps work as a pestle and fissures work as mortar crushing the food coming in between. If these anatomical landmarks are worn-out due to wear, this crushing mechanism gets disturbed and food slips from between the teeth causing a decrease in chewing efficiency ultimately leading to nutritional deficiency and decreased quality of life. Hence teeth wear if diagnosed early can be treated conservatively and a patient can be counseled to prevent further damage. Dental wear has several types which include dental attrition, erosion, abrasion, and abreaction. Dental attrition, abrasion, and erosion are commonly found in the Indian population. All these have specific patterns and causative factors following which they can be diagnosed.

Fig. 2. Tooth wear (erosive)

All these wear patterns can cause irreparable harm to a tooth's structure without creating any noticeable pain in the early stages. When wear and tear reached a severe stage, treatment can be quite costly for the patient. It is possible to avoid getting them by getting checked for them early and making some changes to one's way of living. [15]

B. The Necessity of Detecting Periodontitis

The following section describes the details of gingivitis, periodontitis, and the process of diagnosing these diseases using radiographs and clinical examination. Gingivitis is an inflammatory disease limited to gums adjoining the teeth in which the primary signs are redness and bleeding from the gums along with their swelling. Whereas periodontitis is when the inflammation of the gums extends to the adjoining tooth-supporting bone it is derived as periodontitis. Primary signs of periodontitis are loss of tooth-supporting bone and tooth mobility. Methods are different to diagnose periodontitis and Gingivitis as gingivitis can be diagnosed with the help of clinical examination, observing for the signs and symptoms of the disease. Like red, swollen, and bleeding gums. Periodontitis can be diagnosed with the help of clinical and radiographic examinations. Clinically Doctors can measure the loss of supporting bone with the help of a periodontal probe measuring the sulcus depth, and radio graphically the amount of horizontal and vertical bone loss can be identified to gauge the severity and prognosis of periodontitis. The reason for not considering Gingivitis in this implementation is as it is an early stage that can be diagnosed only on clinical examination and here in our research, we had considered OPGs as a sturdy base which is the most preferred diagnostic tool to diagnose periodontitis. By considering all mentioned parameters early diagnosis is needed to prevent the further loss of teeth-supporting tissues and plan for the repair. Fig. 3 represents an image with periapical and periodontitis.

Fig. 3. Periodontitis and periapical

C. The Necessity of Detecting Tooth Caries (Dental Cavity)

Dental caries is a disease that can be diagnosed easily on clinical and radiographic examination. Treatment of decayed teeth depends upon the extent of decay which can be viewed only radiographically. Orthopantomograms are used to diagnose caries between the teeth and over the teeth and Intraoral periapical radiographs and bitewing radiographs are the specific radiographs to detect the extent of the decay. A representation of a tooth with a cavity is shown in Fig. 4.

Fig. 4. Tooth caries/dental cavity

D. The Necessity of Detecting Periapical

periapical infections can only be diagnosed radiographically as clinical symptoms often appear in the delayed stage. In OPG the extent and the size of the periapical lesion can be exactly gauged. It is shown in Fig. 2 and Fig. 7.

E. The Necessity of Identification of Missing Teeth

Often one or more missing teeth do make complex changes in patients' quality of life. One missing tooth can lead to decay and periodontal disease to seven teeth as the teeth approximating the edentulous space and opposing arch teeth drift from their natural definite position. These all changes can be diagnosed only through an OPG. Fig. 5 shows missing teeth. it is observed that two teeth are missing and the last is affected by decay.

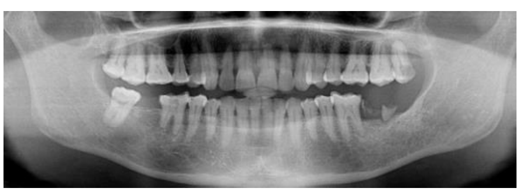

Fig. 5. Missing tooth

F. The Necessity of Identification of Impacted Tooth

Impacted wisdom teeth or any other tooth can’t be seen clinically. OPG gives the exact two-dimensional view of the impacted tooth. A representation of an Impacted tooth is shown in Fig. 6

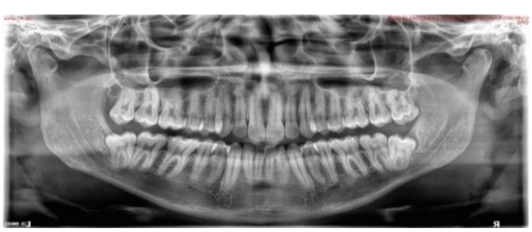

Fig. 6. Impacted tooth

Fig. 7. Wear, decay periapical, and impact

This research work is conceived as a collaborative effort between technical expertise and a dentistry stream in an area of significant interest to the medical community. Literature review reveals that extensive research on tooth extraction or diagnostic purposes used in cybercrimes such as estimation of age, identification of a person based on tooth anatomy, etc, but regrettably, there has been very little work combining the oral health care and engineering fields for continuous engineering and clinical problem-solving. This research began with an emphasis on the various stages of oral illnesses [16]. Such as tooth wear, periapical, periodontitis, tooth decay, tooth loss, and tooth fractures, impacted tooth.

The remaining sections are organized as follows. Section II offers a concise assessment of pertinent literature. It discusses the current state of the art in the use of image processing, machine learning, and deep learning techniques in the domain of dentistry. Understanding different types of radiographs and their importance in evaluating models is discussed using various methods and networks. Section III describes the data collection, radiograph selection, and disease identification processes. The importance of various diseases that can be identified using panoramic radiographs has already been discussed in the introduction section. This section focuses on dataset labeling as well as expert validation of labels. Section IV elaborates on the concept of transfer learning with XGBoost and ensemble learning concerning medical image processing. The proposed methodology is illustrated in Section V, which is followed by the results and discussion and future scope sections.

II. RELATED WORKS

A comparative analysis of the deep learning model for dental segmentation in panoramic radiographs is reported in [18]. The comparison is made between U-Net, DCU-Net, Double U-Net, and NanoNet. On the 1500-image dataset provided by Silva and Olivia. There is also a teeth mask available for this. Their system is designed to identify 32 teeth, fillings, braces, and dental implants. Data enhancement is achieved through random rotation and horizontal inversion. Using a dataset with data augmentation and segmentation without data augmentation based on the segmentation model, the experiment results are shown. The outcome is then compared to the current state of the art.

In this biomedical study [19], the U-Net approach to apical lesion segmentation on panoramic radiographs is discussed. The objective of this study was to extract apical lesions from 470 dental panoramic radiographs. This dataset included 380 images for training, 43 for validation, and 47 for testing. 1140 images were generated after augmentation was applied at the second stage of implementation. For augmentation, horizontal and vertical flips are used. Cropping (pre-processing) is used after augmentation to divide images into four parts: upper right, upper left, lower right, and lower left with size=1000x530. They then multiplied 1140 by 4 for the training group. This dataset is not available to the public.

For apical lesion segmentation from panoramic radiographs, [20] adopts a deep learning technique. There were a total of 1691 images utilized for segmentation. For the training set, all radiographs were manually tagged in red using a polygon labeling tool to construct lesions' contours. The name of the software is "deep stack." The image's original resolution of 1976x976 was reduced to 960x480. Various augmentation techniques, including flip, blur, shift, scale, rotation, sharpness, emboss, contrast, brightness, grid distortion, and elastic transform, are utilized to increase the dataset via online augmentation. Utilizing a pre-trained U-Net CNN for pre-processing and training.

In [21]: Work on a multimodal panoramic x-ray dataset for diagnostic system benchmarking was presented. There are a total of one thousand photos about the labeling of anomalies and teeth. This study was carried out on five levels: anatomical location, peripheral characteristics, radiodensity, surrounding structure effects, and anomaly category. Using eye tracking and a think-aloud methodology, this is the first-time radiologist skill has been captured. This study includes a publicly accessible dataset as well as benchmark performance analyses for a variety of cutting-edge devices. In addition, picture enhancement and image segmentation are carried out. The maxillomandibular region of interest is delineated using x-ray images, gaze maps, and multimodal audio text files. A thousand radiographs and a ground truth mask make up the dataset. 15% is utilized for validation, whereas 85% is used for training. Adam was trained on GPUs with 32 GB of RAM that was NVIDIA Tesla V100 devices. However, in terms of image quality, CLAHE outperforms well with AME Score=24.32 and Log AME=0.3702. Ground truth images were used for segmentation, and the results were compared using UNET, UNET++, PSPNet, DeepLabV3, and DeepLabV3+. VGG19, ResNet18, and ResNet 50 were used to train each of the given networks. The module learning rate is 1x10-3, the weight decay is 1x10-4, the epoch is 150, and the batch size is 16. Pixel accuracy, intersection over union, and dice coefficient are all evaluated. To achieve the desired result, CNN for periodontal bone loss and Squeeze Net for implant fixtures were used. To detect and classify periodontal bone loss, a hybrid framework with Mask RCNN and ResNet101 was used. For intraoral radiographs covering caries, a back propagation neural network is used.

Work on dental caries is being conducted by [c1concepts.]'s (cavity). Utilizing a multi-input deep convolutional neural network ensemble known as MI DCNNE, they were able to obtain accurate results and overcome negative perceptions of panoramic radiography. This system is based on computer-based intelligent vision systems supported by techniques for machine learning and image processing. This new approach accepts both raw periapical images and an improved version of them as inputs. In total, 340 photos were considered for this study. The model has a 99.13% accuracy rate when diagnosing dental caries from periapical pictures.

The primary goal of this work [25] is to extract the area of interest (ROI). Because of background noise, the original image may contain useless information. We occasionally come across images that are of poor quality. Because of this, separation is critical. U-Net was used in three phases to accomplish this: one feature encoder, corresponding decoder, and multipath feature extractor.

The author of [26] proposed an automatic teeth recognition model based on residual network-based faster R-CNN techniques. This model is implemented in two stages. The results obtained in step one, i.e., using R-CNN, are refined further using candidate optimization techniques. These improvements improve the results by about 1%. Ten-fold cross-validation is used in this system to test the model's feasibility and robustness. This model is used to describe the various tooth categories in a panoramic radiograph.

The implementation in [27] is based on the use of u-net architecture to perform panoramic radiograph segmentation. They used CNN to achieve accurate panorama image segmentation. With a 94% accuracy and a 94% dice score. The performance of the u-net is compared with various segmentation techniques in this work. This fuzzy c-means (FCM) is developed using a hybrid of FCM and a neutrophilic approach to segment the jaw and lesion present in the jaw using a panoramic radiogram. This paper investigates the comparison of global thresholding, fuzzy c-means, watershed, canny edge detection algorithms, and the U-NET model. When compared to the other techniques, n-net performed the best. Ivison lab provided the dataset for this study.

In [28], researchers offered an innovative approach to segmenting and classifying data. They have covered 15 various disorders including healthy teeth, missing teeth, dental restoration implants, fixed prostheses, mobile prostheses, etc. To depict all 15 diseases, a total of 2,000 images are made. All these images were taken at separate dental clinics. A convolutional neural network (CNN) was trained to perform semantic segmentation on images from the dataset, and many techniques were employed to enhance the annotated data. Images are segmented and then binarized with multiple thresholds. The detection of teeth is a two-stage algorithmic process. When discrepancies are uncovered during semantic segmentation, refinement techniques are employed to rectify the situation. In this study, the Nobel approach to teeth detection and dental problem classification using deep learning and image processing was introduced.

The paper [22] compares ten different segmentation methods for use with dental imaging. The offered solutions were evaluated and classified using five different metrics: accuracy, specificity, precision, recall, and F1-score. Because of the shattered bones in the buccal cavity, not one of the ten segmentation methods tried was successful in isolating the teeth.

With the help of a mask region-based convolution neural network, the authors of [23] suggested a method for instance segmentation of teeth in panoramic photos. In the wake of Resnet-101-driven feature extraction, a feature pyramid network (FPN) is built with predetermined anchors and extracted regions of interest.

Each region's proposal network consists of the FPN and its anchors (RPN). The regions of interest are then repositioned so that they are all the same size. Additionally, the bounding box coordinates are used to classify each feature as either a tooth or a background and to pinpoint its exact location. Finally, a bounding box is formed around the tooth once it has been segmented.

DeNTNet, described in [24], is a deep neural transfer network trained to identify panoramic dental radiographic evidence of periodontal bone loss (PBL). During detection, several convolutional neural networks are trained. A segmentation network is then trained to extract teeth from the ROI, and a second network is trained to predict areas of periodontal bone loss. Using the encoder portion of the lesion segmentation network as a pre-trained model, a classification network is constructed to predict the presence of PBL in each tooth.

Table I provides a complete overview of the various datasets available, the work done in dental image processing, and the disorders linked with it. It has been observed that work is being done in various disciplines of dentistry employing artificial intelligence techniques. To begin, a private dataset is produced and not made public. Another observation is that most systems are designed to segment teeth. Furthermore, the minimum one and maximum four diseases covered for classification and detection are one. One method is used, in which 14 diseases are investigated, but just the highlighting of diseases is done, rather than classification or detection.

TABLE I. ANALYSIS OF AN AVAILABLE DATASET AND METHODOLOGY USED

Papers	Variable detected	Total Images	Publicly available	Labeled Teeth	Labeled abnormalities	Methodology Used
[36]	Teeth	100	No	Yes	No	CNN
[37]	Teeth	100	No	Yes	No	CNN
[38]	Teeth	1500	Yes	Yes	No	Mask R-CNN
[39]	Teeth	1352	No	Yes	No	CNN
[40]	Endodontic	85	No	No	Yes	CNN
[41]	Endodontic	300	No	Yes	No	CNN
[42]	Teeth	1000[116]	Yes (Kaggle)	No	No	CNN
[32]	Teeth	1200	No	Yes	No	Yolov3
[43]	Teeth	1000	No	Yes	No	Deep Learning and Image Processing Techniques
Own Dataset	Teeth	500	No	No	Yes	Transfer Learning and XGBoost

III. DATASET DESCRIPTION

Researching any specific domain data availability is the primary step, but in the field of dental informatics dataset availability is a major concern. After analyzing all the related work in Dental Image Processing (DIP), we discovered that obtaining a suitable dataset for the early stages of research is challenging. As a result, we collaborated with many dental healthcare sections to acquire data. Rather than waiting for new patients for each ailment, we chose to screen panoramic dental X-rays from electronic medical records to collect data efficiently. There is no ethical issue because these images do not contain any personally identifiable information. The dataset consists of images and a questionnaire, which was supplementary at the initial level of research. Collecting symptoms and images for every patient is a time-consuming task so we focused on a collection of radiographs. Principally there are many types of radiographs available in dental radiography, such as Bitewing X-ray, Periapical X-ray, Occlusion X-ray, Panoramic X-ray, Cephalometric Projection, and Cone Beam X-ray. Among these categories, Panoramic OPG (Orthopantomogram) is considered as is two-dimensional radiograph that covers maximum all diseases with complete coverage. Dental radiographic images in OPG format were collected.

A. Panoramic (OPG) X-Ray and its Benefits

The maxillary and mandibular teeth, as well as the surrounding periodontium and certain anatomical landmarks like the maxillary sinus and the temporomandibular joints, are all visible on an OPG, a two-dimensional radiograph. Oral disorders such as periapical abscess, cysts, osteitis, and various tumors and cysts of mandibular and maxillary origin can be diagnosed using OPG X-ray. OPG is useful for determining how much bone has been lost due to periodontal disease and how severe it is. OPG X-ray is useful for diagnosing a wide variety of dental conditions, including cavities, root fractures, and periodontal inflammation. OPG X-rays can help identify and evaluate dental cusp fracture and enamel deterioration, but their two-dimensional nature has its limitations.

The data was collected from Government Dental College and several private dental clinics and laboratories in Nagpur and cities across India. Here image collection work is focused on the wear of the tooth to find erosion, attrition, and abrasion of the tooth and other oral diseases, including impacted tooth, periapical lesion, fractured tooth, missing teeth, etc.,

B. Process of Collecting, Preparing and Labeling the Data

This section describes the process of collecting, preparing, and labeling the dataset. The main contribution while collecting images and preparing the dataset are expressed with the help of the diagram given below:

a) From 2019 to 2021, we collected approximately 2,000 OPG radiographs from multiple sources.

b) All images were then checked for contrast and illumination quality.

c) Images of poor quality and nonrelevant, such as images with prostheses, multiple caps, braces, etc were removed from the dataset.

d) Diseases that can be seen through OPG are identified and discussed with clinicians.

e) OPG radiographs can be used to diagnose diseases such as caries, periodontitis, periapical infection, impacted teeth, tooth wear, bone fracture, missing teeth, and anatomy of the maxillary sinus and temporomandibular joint.

f) Following disease identification, images are labeled with the assistance of various dental experts (such as Oral and Maxillofacial Radiologists, Prosthodontists, Oral and Maxillofacial Surgeons, Periodontics, and endodontists,).

g) Image labeling is done in two ways: radiograph with multiple diseases and radiograph with a single disease.

h) After the image sorting process, it is discovered that the approximate probability of receiving a relevant radiograph is 50/500. Finally, we can filter 500 useful radiographs from available images.

i) Before developing any module, various pre-processing steps are performed to create the extension of the data set, In this automatic cropping of the image is done to extract the relevant portion by using an automatic cropping algorithm.

j) Our dataset is cross-validated and tested with the help of three different domain experts to make it more consistent, reliable, authentic, and useful.

k) The final label matrix is generated for different six diseases if the disease is present in that image so marked as 1 otherwise 0 which is shown in Table II: Labelled dataset. For example, in image id 92 wear, decay, missing, and impacted tooth diseases are present.

TABLE II. SAMPLE LABELLED DATASET

Image Id	Wear	Decay	Peri-Apical	Periodontitis	Missing	Impacted
1	0	1	0	0	0	0
92	1	1	0	0	1	1
104	1	0	0	0	0	1
238	1	1	0	0	1	0
347	1	1	1	0	0	0
457	1	0	1	1	0	0

To check the validity and authentication of dataset creation complete supplementary material is available. The panoramic dental images were obtained from a variety of sources, including dental laboratories. The image properties due to different acquisition systems vary in terms of contrast and brightness. The pre-processing challenge was to adjust the contrast and illumination levels of all the collected images so that subsequent stages produced the same set of features corresponding to each class or group. We gathered images that were prone to six major issues: dental decay, dental wear, periapical, periodontitis, missing teeth, impacts, and a combination of all six diseases. An experienced medical practitioner in the respective field. While creating the LABEL matrix great care was taken to avoid any false labeling in the LABEL matrix. Images with low contrast and illumination were discarded after being manually sorted one by one. Duplicated images were identified using image comparison and separate codes. Finally, 500 images were taken for this research project, while the remaining unclear images were discarded. Fig. 8 shows the analysis of the total number of images available in the dataset.

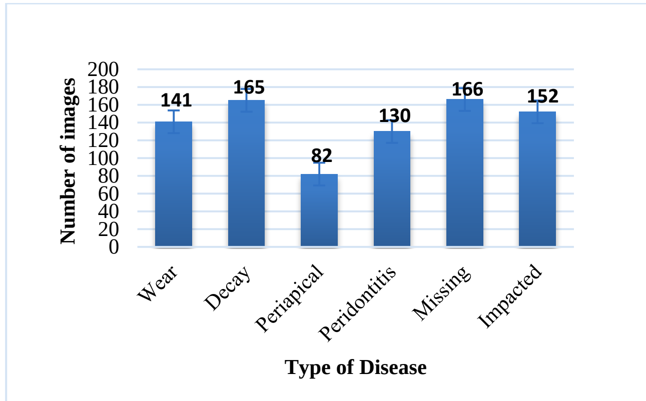

Fig. 8. Analysis of images available in the dataset

IV. TRANSFER LEARNING, XGBOOST, AND ENSEMBLE APPROACH

A. Transfer Learning

Deep feature extraction (DFE) and Transfer learning (TL) are two of the most effective alternatives to employing a small number of images as training samples. When a small sample size of images is present. TL refers to the process of applying previously taught models to unexpected challenges. TL is not a distinct classification of Machine learning algorithms; rather, it is a technique that can be used to create a new Machine learning model. The model will be able to use the knowledge and skills acquired from earlier training in new scenarios. In a manner like the preceding task, data will need to be organized based on the type of data. A further application of TL is the extraction of deep feature data.

Rather than manually modifying the activation layers of the CNN, it is possible to extract feature vectors by utilizing pre-trained CNN models. The deeper layers, which are activated by the activation of the lower-level layers, include the picture classification-critical higher-level features.

Transfer Learning is a technique in which we reuse a previously trained model as the foundation for a new model on a new challenge. In this case, a model trained on one task is repurposed for another. Transfer learning tries to improve target learners' performance on the target domain by transferring knowledge from distinct but related source domains. ResNet50V2, ResNet101V2', 'MobileNetV3Large', 'MobileNetV3Small', 'MobileNet', 'EfficientNetB0', 'EfficientNetB1', and 'EfficientNetB2' are examples of pre-trained networks. [33].

ResNet50V2 [34] is a changed version of ResNet50. It does better on the ImageNet dataset than ResNet50 and ResNet101. ResNet-101 is a 101-layer convolutional neural network.

MobileNetV3 [35] is a convolutional neural network that is tuned to mobile phone CPUs using a combination of hardware-aware network architecture search (NAS) and the NetAdapt algorithm. It was then improved by making new architectural advances.

EfficientNet-B0 is a convolutional neural network that has been trained on more than a million images from the ImageNet database [30]. The network can divide images into 1000 different kinds of objects, like a keyboard, mouse, pencil, and many animals.

B. *XGBoost*

Extreme Gradient Boosting, which is what XGBoost stands for, is a distributed gradient-boosted decision tree (GBDT) machine learning library that can be used on a large scale. It is the best machine learning library for regression, classification, and ranking problems, and it has parallel tree boosting. It's like Random Forest in that it builds a group of decision trees, but instead of training the models at the same time, XGBoost trains them one at a time. Each new decision tree learns from the mistakes made by the one before it. Boosting is the process of training models one after the other. A gradient in XGBoost stands for a type of boosting that uses weak learners. Weak learners are simple models that only do better than random chance. The algorithm starts with a weak learner at the beginning. Each new model tries to fix the mistakes that the previous decision tree made. This keeps happening until there are no more ways to make the model better. The result is a strong learner model. [31]

C. *Ensemble Modeling*

Ensemble modeling is a way to predict what will happen based on several different base models. The goal of using a group of models is to reduce the prediction's generalization error. When the ensemble approach is used, the prediction error goes down if the base models are different and can be used on their own. The method tries to figure out what will happen by asking a variety of individuals. Even though the ensemble model is made up of several base models, it works like a single model. [32]

D. *Weighted Average or Weighted Sum Ensemble*

Weighted average or weighted sum ensemble is a type of machine learning that uses a group of models to make a prediction. The contribution of each model is weighted according to how good it is. The voting ensemble has something to do with the weighted average ensemble. In this method, we didn't tune it; instead, we used the weights that were already there. In other words, a weight tells how much the input affects the output. Biases, which are always the same, are an extra input for the next layer that will always be 1[33][34].

V. Proposed Methodology

We have created a database of 500 images of panoramic radiographs which are explained in section III. Dataset is divided into a training set and a testing set. The training set goes to eight different pre-trained models.

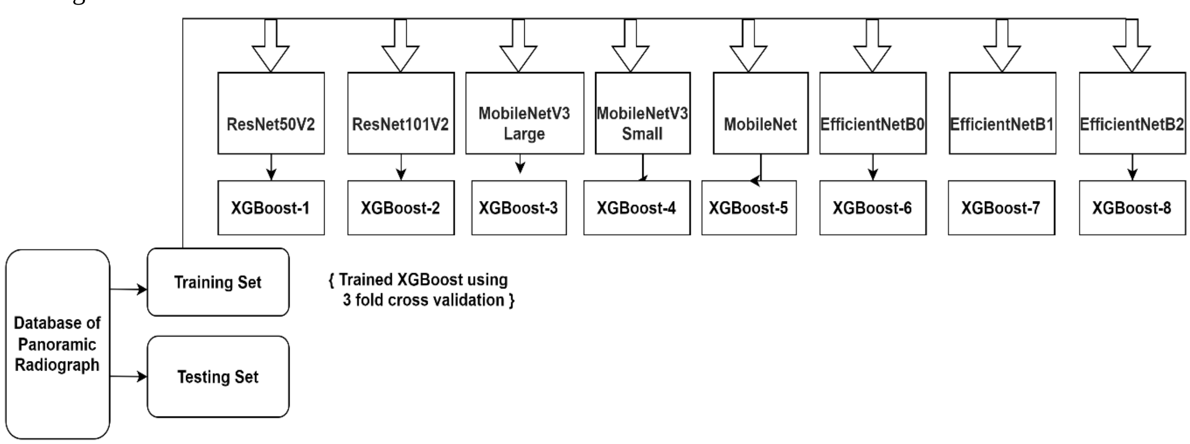

Fig. 9. Proposed methodology (training phase)

Different pre-trained models used for this phase are 'ResNet50V2', 'ResNet101V2', 'MobileNetV3Large', 'MobileNetV3Small', 'MobileNet', 'EfficientNetB0', 'EfficientNetB1', and 'EfficientNetB2'. Out of 500 samples 80% samples used for training. The detailed process of the training and evaluation phase is given in Fig. 9 and Fig. 10. While training we used 3-fold cross-validation. Cross-validation is a technique for testing models that involves training them on distinct subsets of the available input data and then testing them on the other. Dataset D is divided into three identically sized subsets. The process of fitting and evaluating the model is done three times, and each time a different subset is used as a training sample. At last, in the testing phase given in Fig. 10, we get eight tuned XGBoost. pre-trained networks are used to extract the features and then extracted features are given as input to the XGBoost model for training.

In comparison to other gradient-boosting approaches, XGBoost is nearly ten times faster and has a strong predictive power. Additionally, it contains a range of regularizations that minimize overfitting and enhance overall performance. These eight trained XGBoost will generate eight different classifier prediction outputs. To get the final prediction weighted ensemble module which is an extension of a model averaging ensemble where the contribution of each member to the final prediction is weighted by the performance of the model.

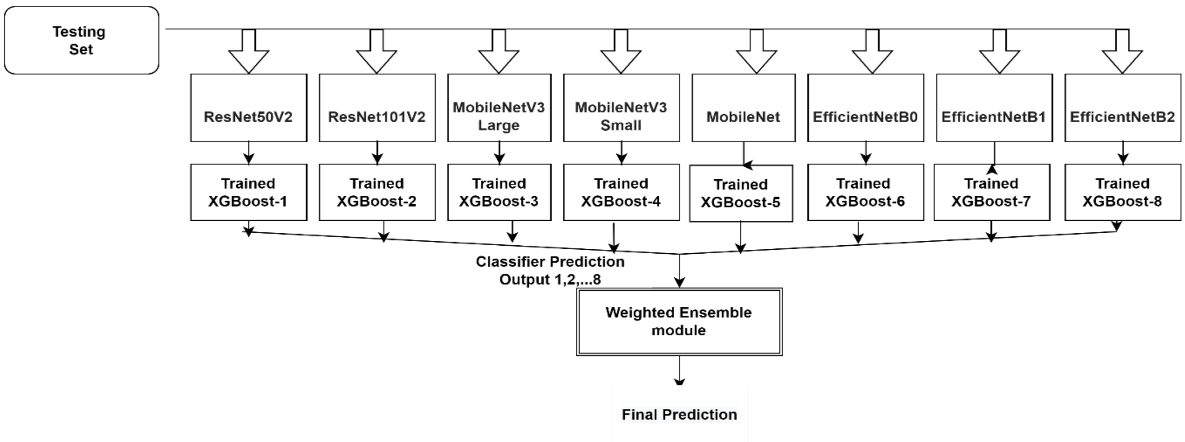

Fig. 10. Proposed methodology (evaluation phase)

VI. Results and Discussion

The developed methodology employing transfer learning is processed using Python; the predicted model's success rate is evaluated utilizing existing mechanisms in terms of accuracy, sensitivity, F-measure, precision, and recall. In this approach, 500 X-ray dental images are utilized by considering 80% for training and 20% for testing. The proposed system can classify multiple diseases which cover six different categories with 46 possible combinations. the dataset covers all types of images like one image with one disease and one image with multiple diseases. For example, this system can identify images where only wear or decay is present, apart from this it is also able to identify and classify images where different diseases are present such as wear, and decay as shown in image id.347. Thus, to prove the sustainability of the proposed work ablation study is performed which is expressed in Tables III, IV, V, and VI by considering different 8 pre-trained networks. Table III shows the performance of the model by considering seven different networks with 91% accuracy. Similarly, Table IV shows the performance of the classification of six different diseases by considering six different networks with 90% accuracy on average.

TABLE III. 7-Pre Trained Models [MobileNetSmall, MobileNet, MobileNetLarge, EfficientNetB0 EfficientNetB1 EfficientNetB2 ResNet50V2]

Performance Measure	Wear	Decay	Periapical	Periodontics	Missing	Impacted
Precision	0.776	0.843	0.717	0.799	0.826	0.773
Recall	0.957	0.912	0.855	0.925	0.91	0.944
F1-Score	0.857	0.876	0.78	0.857	0.866	0.85
Specificity	0.895	0.921	0.933	0.926	0.913	0.888
Accuracy	0.912	0.918	0.92	0.926	0.912	0.904

TABLE IV. 6-Pre Trained Models[MobileNetSmall, MobileNet, MobileNetLarge, EfficientNetB0 EfficientNetB1 EfficientNetB2]

Performance Measure	Wear	Decay	Periapical	Periodontics	Missing	Impacted
Precision	0.767	0.841	0.703	0.793	0.825	0.764
Recall	0.957	0.899	0.855	0.925	0.904	0.944
F1-Score	0.852	0.869	0.772	0.854	0.862	0.845
Specificity	0.89	0.921	0.928	0.924	0.913	0.882
Accuracy	0.908	0.914	0.916	0.924	0.91	0.9

TABLE V. 5-Pre Trained Models [MobileNetSmall, MobileNet, MobileNetLarge, EfficientNetB0 EfficientNetB1]

Performance Measure	Wear	Decay	Periapical	Periodontics	Missing	Impacted
Precision	0.765	0.831	0.689	0.786	0.815	0.756
Recall	0.942	0.899	0.855	0.917	0.904	0.944
F1-Score	0.844	0.864	0.763	0.846	0.857	0.84
Specificity	0.89	0.915	0.923	0.921	0.907	0.876
Accuracy	0.904	0.91	0.912	0.92	0.906	0.896

TABLE VI. 4-Pre Trained Models[MobileNetSmall, MobileNet, MobileNetLarge, EfficientNetB0]

Performance Measure	Wear	Decay	Periapical	Periodontics	Missing	Impacted
Precision	0.765	0.826	0.67	0.786	0.815	0.756
Recall	0.929	0.893	0.852	0.902	0.892	0.932
F1-Score	0.83	0.858	0.75	0.84	0.852	0.834

	9					
Specificity	0.88 9	0.912	0.919	0.921	0.906	0.876
Accuracy	0.9	0.906	0.908	0.916	0.902	0.892

A. Precision

Precision is the proportion of precise teeth affected region computed using Eqn. (1).

$$P = \frac{T_p}{T_p+F_p} \tag{1}$$

Fig. 11 shows the evaluation of precision for different diseases. it is observed that the precision of wear, decay, periapical, periodontitis, missing tooth, and impacted tooth is 79%, 85%, 73%, 82%, 84%, and 78% respectively.

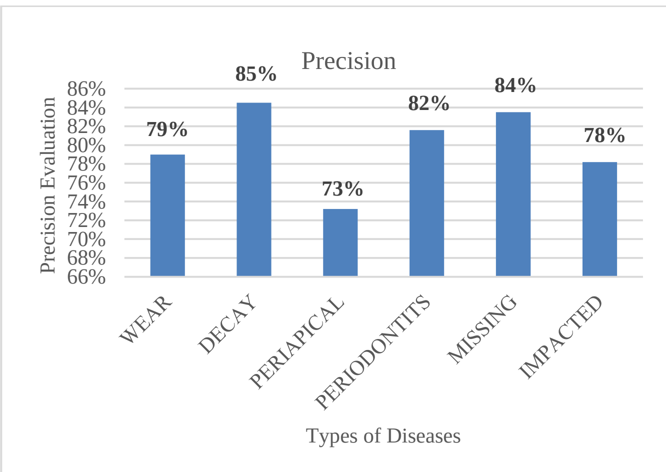

Fig. 11. Comparison of precision for different diseases

B. Recall

It expresses the proportion of predictions that have been correctly diagnosed as expected outcomes and is calculated using Eqn. (2).

$$R = \frac{T_p}{T_p+F_n} \tag{2}$$

It is observed from Fig. 12 that the Recall value for wear, decay, periapical, periodontitis, missing tooth, and impacted tooth is acceptable in the healthcare domain as it is above 95%,93%,86%,93%,91%, and 94% respectively.

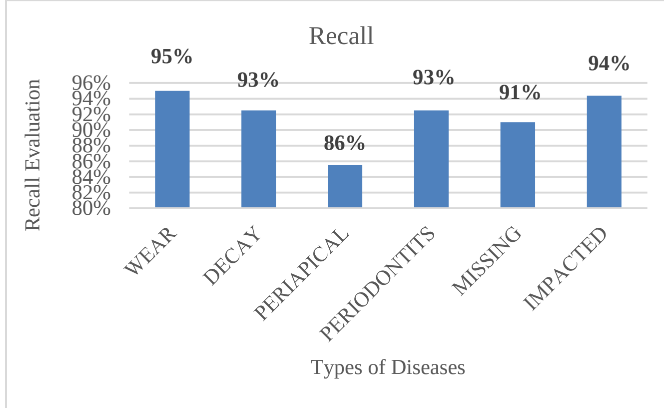

Fig. 12. Comparison of recall for different diseases

C. F-measure

F1-score is measured using Eqn. (3).

$$F1 - score = 2\left(\frac{P*R}{P+R}\right) \tag{3}$$

As per statistics shown in Fig. 13, it is observed that the F1-score for the detection of wear, decay, periapical, periodontitis, missing tooth, and impacted tooth is 86%, 88%, 79%, 87%, 87%, and 86% respectively.

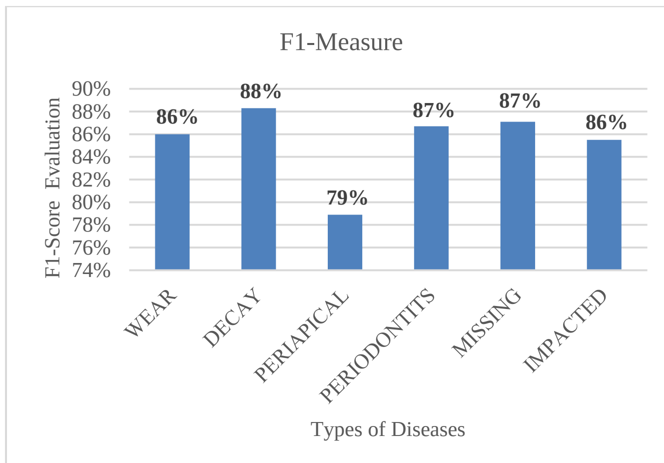

Fig. 13. Comparison of F1-measure for different diseases

D. Specificity

A recall is the probability of a negative diagnosis such as the patient being free from that disease. It is measured using eq. (4).

$$S = \frac{Tn}{(Tn+Fp)} \tag{4}$$

As per the statistical data presented in Fig. 14 specificity for all the diseases covering wear, decay, periapical, periodontitis, missing, and impacted teeth are found to be 90%,92%,94%,93%,92%, and 89% respectively.

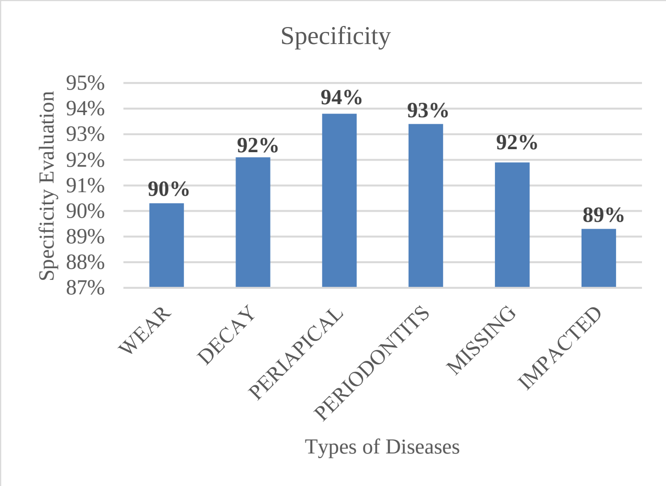

Fig. 14. Comparison of specificity for different diseases

E. Accuracy

Accuracy to the entire observations that are expressed in Eqn. (5),

$$A = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (5)$$

As per statistics shown in Fig. 15, it is examined that accuracy for classification of wear, decay, periapical, periodontitis, missing tooth, and impacted tooth is 92%, 92%, 92%, 93%, 92%, and 91% respectively.

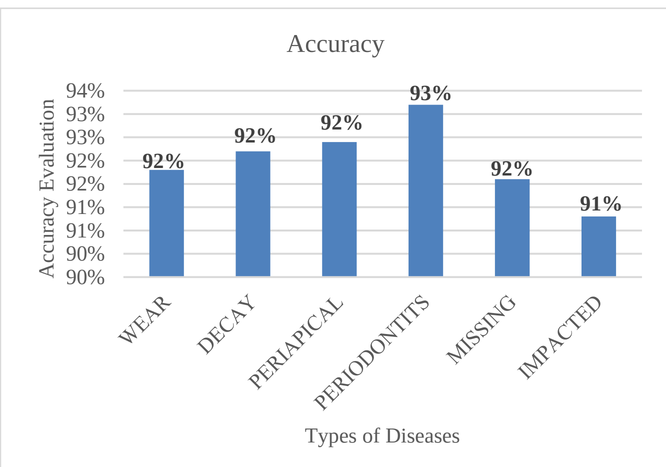

Fig. 15. Comparison of accuracy for different diseases

VII. Conclusion and Future Work

This system is implemented using the idea of transfer learning and different pre-trained networks, such as 'ResNet50V2', 'ResNet101V2', 'MobileNetV3Large', 'MobileNetV3Small', 'MobileNet', 'EfficientNetB0', 'EfficientNetB1', and 'EfficientNetB2', with XGBoost. To get the final prediction, a weighted ensemble module is used. Additionally, the developed technique has achieved better outcomes in terms of accuracy, precision, F-measure, recall, and sensitivity. Thus, it achieved 92-93% of accuracy in classifying multiple dental diseases including tooth wear, periapical, periodontitis, tooth decay, missing tooth, and impacted tooth. This approach is implemented on new diseases and compared the measuring parameters against different diseases. For some measuring parameters, values are less compared which can be improved in future work by using the extension of this dataset. In this present study, the given model is tested by classifying six diseases and their 46 possible combinations. this system can identify one disease present in an image as well as a combination of multiple diseases present in one image. The proposed method performs better in terms of accuracy and different measuring parameters than current state-of-the-art methods and has a variety of uses in computer-assisted multiple tooth disease identification and classification.

Acknowledgment

This work is supported by the following dental clinics for data collection, understanding, annotation, authentication, and validation of images.

a) Government dental college, Nagpur.

b) Dr. Jayaswal's clinic of Dentistry and Geriatric oral health care center, Nagpur, India.

c) Government Dental College, Aurangabad, India

d) Invasion Lab, Nagpur

Data Availability

On request, the corresponding author will provide the data used to generate the study's findings.

Funding Details

This work is being carried out with the assistance of an AICTE financing award (F No B-60/RFID/RPS/Policy-1/2017-18).

References

[1] Gu, Zaiwang, et al. "Ce-net: Context encoder network for 2d medical image segmentation." IEEE transactions on medical imaging 38.10 (2019): 2281-2292.

[2] Shiraisi, Ai, et al. "Improvement in Oral Health Enhances the Recovery of Activities of Daily Living and Dysphagia after Stroke." Journal of Stroke and Cerebrovascular Diseases 30.9 (2021): 105961.

[3] Lakshmi, M. Muthu, and P. Chitra. "Tooth Decay Prediction and Classification from X-Ray Images using Deep CNN." 2020 International Conference on Communication and Signal Processing (ICCSP). IEEE, 2020.

[4] Schwendicke, Falk, et al. "Convolutional neural networks for dental image diagnostics: A scoping review." Journal of dentistry 91 (2019): 103226.

[5] Hashem, Mohamed, and Ahmed E. Youssef. "Teeth infection and fatigue prediction using optimized neural networks and big data analytic tool." Cluster Computing 23.3 (2020): 1669-1682.

[6] Angst, Patrícia Daniela Melchiors, et al. "Oral health status of patients with leukemia: a systematic review with meta-analysis." Archives of Oral Biology (2020): 104948.

[7] Banu, Ancuta, et al. "Dental health between self-perception, clinical evaluation and body image dissatisfaction–a cross-sectional study in mixed dentition pre-pubertal children." BMC Oral Health 18.1 (2018): 1-9.

[8] Liu, Lizheng, et al. "A smart dental health-IoT platform based on intelligent hardware, deep learning, and mobile terminal." IEEE journal of biomedical and health informatics 24.3 (2019): 898-906.

[9] Glurich, Ingrid, et al. "Interdisciplinary care model: pneumonia and oral health." Integration of Medical and Dental Care and Patient Data. Springer, Cham, 2019. 123-139.

[10] Abdalla-Aslan, Ragda, et al. "An artificial intelligence system using machine-learning for automatic detection and classification of dental restorations in panoramic radiography." Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology 130.5 (2020): 593-602.

[11] Grischke, Jasmin, et al. "Dentronics: Towards robotics and artificial intelligence in dentistry." Dental Materials 36.6 (2020): 765-778.

[12] Mahmoud, Nourelhoda M., et al. "Detecting dental problem-related brain disease using intelligent bacterial optimized associative deep neural network." Cluster Computing 23.3 (2020): 1647-1657.

[13] Lee, Jeong-Hoon, et al. "Automated cephalometric landmark detection with confidence regions using Bayesian convolutional neural networks." BMC oral health 20.1 (2020): 1-10.

[14] Sonavane, Apurva, RohitYadav, and AdityaKhamparia. "Dental cavity Classification using Convolutional Neural Network." IOP Conference Series: Materials Science and Engineering. Vol. 1022. No. 1. IOP Publishing, 2021.

[15] Tandon, Divya, JyotikaRajawat, and Monisha Banerjee. "Present and future of artificial intelligence in dentistry." Journal of Oral Biology and Craniofacial Research 10.4 (2020): 391-396.

[16] P. Jaiswal and S. Bhirud, "Study and Analysis of an Approach Towards the Classification of Tooth Wear in Dentistry Using Machine Learning Technique," 2021 IEEE International Conference on Technology, Research, and Innovation for Betterment of Society (TRIBES), 2021, pp. 1-6, DOI: 10.1109/TRIBES52498.2021.9751650.

[17] Nripendra Kumar Singh, Khalid Raza, Progress in deep learning-based dental and maxillofacial image analysis: A systematic review, Expert Systems with Applications, Volume 199,2022,116968, ISSN 0957-4174,

[18] https://doi.org/10.1016/j.eswa.2022.116968. da Silva Rocha, É.; Endo, P.T. A Comparative Study of Deep Learning Models for Dental Segmentation in Panoramic Radiograph. Appl. Sci. 2022, 12, 3103. https://doi.org/10.3390/app12063103

[19] Ibrahim S. Bayrakdar, Kaan Orhan, Özer Çelik, Elif Bilgir, Hande Sağlam, Fatma Akkoca Kaplan, Sinem Atay Görür, Alper Odabaş, Ahmet Faruk Aslan, Ingrid Różyło-Kalinowska, "A U-Net Approach to Apical Lesion Segmentation on Panoramic Radiographs", BioMed Research International, vol. 2022, Article ID 7035367, 7 pages, 2022. https://doi.org/10.1155/2022/7035367

[20] Il-Seok Song, Imaging Science in Dentistry 2022, Deep learning-based apical lesion segmentation from panoramic radiographs.

[21] K. Panetta, R. Rajendran, A. Ramesh, S. P. Rao, and S. Agaian, "Tufts Dental Database: A Multimodal Panoramic X-Ray Dataset for Benchmarking Diagnostic Systems," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 4, pp. 1650-1659, April 2022, DOI: 10.1109/JBHI.2021.3117575.

[22] G. Silva, L. Oliveira, M. Pithon, "Automatic segmenting teeth in X-ray images: Trends, a novel data set, benchmarking and future perspectives", Expert Systems with Applications, 107, pp. 15-31, 2018.

[23] G. Jader, J. Fontineli, M. Ruiz, K. Abdalla, M. Pithon, L. Oliveira, "Deep instance segmentation of teeth in panoramic X-ray images" 31 st SIBGRAPI Conference on Graphics, Patterns and Images, October 2018.

[24] J. Kim, H. S. Lee, I. S. Song & K. H. Jung, "DeNTNet: Deep Neural Transfer Network for the detection of periodontal bone loss using panoramic dental radiographs", in Scientific reports, vol.9(1), pp. 1-9, Nov. 2019.

[25] K. Zhang, J. Wu, H. Chen, and P. Lyu, "An effective teeth recognition method using label tree with cascade network structure", Computerized Medical Imaging and Graphics, vol. 68, pp. 61-70, 2018.

[26] D. Abdelhafiz, C. Yang, R. Ammar, S. Nabavi, "Deep convolutional neural networks for mammography: advances, challenges, and applications", BMC Bioinformatics, 20(S11), 2019.

[27] [8] M.A. Kadampur, S. Al Riyaee, "Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images", Informatics in Medicine Unlocked, 18, 100282, 2020

[28] D.V. Tuzoff, L.N. Tuzova, M.M. Bornstein, A.S. Krasnov, M.A. Kharchenko, S.I. Nikolenko, M.M. Sveshnikov, G.B. Bednenko, "Tooth detection and numbering in panoramic radiographs using convolutional neural networks", Dentomaxillofacial Radiol, 48, 20180051, 2019.

[29] Saleh A. Albelwi, "Deep Architecture based on DenseNet-121 Model for Weather Image Recognition" International Journal of Advanced Computer Science and Applications(IJACSA), 13(10), 2022. http://dx.doi.org/10.14569/IJACSA.2022.0131065

[30] Krois, J., Cantu, A. G., Chaurasia, A., Patil, R., Chaudhari, P. K., Gaudin, R., Gehrung, S., & Schwendicke, F. (2021). Generalizability of deep learning models for dental image analysis. Scientific Reports, 11. https://doi.org/10.1038/s41598-021-85454-5

[31] Xin Yu Liew, Nazia Hameed, Jeremie Clos, An investigation of XGBoost-based algorithm for breast cancer classification, Machine Learning with Applications, Volume 6,2021,100154, ISSN 2666-8270,https://doi.org/10.1016/j.mlwa.2021.100154.

[32] Almalki, Y.E.; Din, A.I.; Ramzan, M.; Irfan, M.; Aamir, K.M.; Almalki, A.; Alotaibi, S.; Alaglan, G.; Alshamrani, H.A.; Rahman, S. Deep Learning Models for Classification of Dental Diseases Using Orthopantomography X-ray OPG Images. Sensors 2022, 22, 7370. https://doi.org/10.3390/s22197370

[33] Guo, Y., Ge, Y., Yang, C., Al-Garadi, M. A., & Sarker, A. (2022). Comparison of Pretraining Models and Strategies for Health-Related Social Media Text Classification. Healthcare, 10(8). https://doi.org/10.3390/healthcare10081478

[34] Raeho Lee and Haewon Byeon, "Application of Stacking Ensemble Machine in Big Data: Analyze the Determinants for Vitalization of the Multicultural Support Center" International Journal of Advanced Computer Science and Applications(IJACSA), 13(10), 2022. http://dx.doi.org/10.14569/IJACSA.2022.0131007

[35] Munindra Lunagaria, Vijay Katkar, and Krunal Vaghela, "Covid-19 and Pneumonia Infection Detection from Chest X-Ray Images using U-Net, EfficientNetB1, XGBoost and Recursive Feature Elimination" International Journal of Advanced Computer Science and Applications(IJACSA), 13(9), 2022. http://dx.doi.org/10.14569/IJACSA.2022.0130954

[36] A. B. Oktay, "Tooth detection with convolutional neural networks," in Proc. IEEE Med. Technol. Nat. Congr., 2017, pp. 1–4.

[37] C. Muramatsu et al., "Tooth detection and classification on panoramic radiographs for automatic dental chart filing: Improved classification by multi-sized input data," Oral Radiol., vol. 37, no. 1, pp. 13–19, 2021.

[38] G. Jader, J. Fontineli, M. Ruiz, K. Abdalla, M. Pithon, and L. Oliveira, "Deep instance segmentation of teeth in panoramic X-ray images," in Proc. IEEE 31st SIBGRAPI Conf. Graph., Patterns Images, 2018, pp. 400–407.

[39] D. V. Tuzoff et al., "Tooth detection and numbering in panoramic radiographs using convolutional neural networks," Dentomaxillofacial Radiol., vol. 48, no. 4, 2019, Art. no. 20180051

[40] T. Ekert et al., "Deep learning for the radiographic detection of apical lesions," J. Endodontics, vol. 45, no. 7, pp. 917–922.e5, 2019

[41] M. Fukuda et al., "Evaluation of an artificial intelligence system for detecting vertical root fracture on panoramic radiography," Oral Radiol., pp. 1–7, 2019

[42] K. Panetta, R. Rajendran, A. Ramesh, S. P. Rao, and S. Agaian, "Tufts Dental Database: A Multimodal Panoramic X-Ray Dataset for Benchmarking Diagnostic Systems," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 4, pp. 1650-1659, April 2022, DOI: 10.1109/JBHI.2021.3117575.

[43] M. P. Muresan, A. R. Barbara, and S. Nedevschi, "Teeth Detection and Dental Problem Classification in Panoramic X-Ray Images using Deep Learning and Image Processing Techniques," 2020 IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP), 2020, pp. 457-463, DOI: 10.1109/ICCP51029.2020.9266244.