Syrian Arab Republic

Ministry of Higher Education

Syrian Virtual University

Master in Bioinformatics (BIS)

Deep Learning-based System for Automated Classification and Detection of Lesions in Dental Panoramic Radiography

A thesis submitted in partial fulfillment of the requirements for the degree of Master in Bioinformatics

Kenana Hejazi

Kenana_190324

F23

Supervisor:

A.Prof. Yasser Khadra

2024

Dedication

This work is dedicated to the memory of my father for his principles, thoughts, and constant love of science, which made me proud to be his daughter.

I also dedicate this work to my soulmate Tarek and my lovely kids Maysoun and Ayham, I hope they will be proud of me when they grow up.

To my family - My mother, Yasser, Thuraya, Abd Alhamid, Hazar and Arwa - for their always goodness, support, and love.

To my best friend Obada, for his cooperation and support in the most difficult times.

I also dedicate this to the great friends that I met along the way – Shahd and Shaza. Thank you for always being here for me. May we always be together.

Acknowledgements

Deepest regard and thanks to Professor Yasser Khadra for his supervision.

Abstract

The field of dentistry is highly dependent on the analysis of dental radiographs to guide diagnosis of disease. The integration of AI applications based on deep learning in analyzing panoramic radiographs aims to offer highly accurate, efficient, and automated detection of dental lesions, aiding dentists in diagnosing these lesions while reducing their time and risk of misdiagnosis. These applications enhance the accuracy and allow for early intervention, ultimately improving patient outcomes in dental care. The aim of this project is to develop a deep learning-based system for the automated classification and detection of lesions in dental panoramic radiography. Utilizing the Tufts Dental Database (TDD), which comprises 1000 dental panoramic images, a classification step using traditional CNN model was implemented to classify X-Ray images into normal and abnormal classes, then an object detection system was developed using Faster R-CNN model, to detect abnormal dental lesions. The dataset was augmented to enhance the model's robustness and performance, then split into training, validation, and testing sets in a 70-15-15 ratio, respectively. After classifying the images as abnormal by CNN, the Faster RCNN model provides bounding boxes around detected lesions in abnormal regions. The system showed promising results for the classification step with accuracy of 84.56 and for the detection stage with a precision of 82.35%, recall of 75%, F1 score of 78.5% and IoU of 55%.

Table of Contents

Contents

List of Figures

List of Tables

Introduction

X-rays, or radiography, are a form of electromagnetic radiation that uses a very small dose of ionizing radiation to produce images of the body's internal structures. It was discovered in 1895 by Wilhelm Conrad Roentgen (1845-1923), X-rays are the oldest and most frequently used form of medical imaging. They are commonly used to diagnose fractured bones, detect injuries or infections, and locate foreign objects in soft tissues.

Dental radiography (dental X-ray) is a crucial tool in modern dentistry, playing an essential role in diagnosing dental issues and detecting lesions that might not be visible during a routine dental examination. The primary purpose of dental radiography is to identify various dental conditions such as cavities, gum disease, infections, cysts, tumors, decay, bone loss, developmental lesions and abnormalities in tooth structure or positioning. By revealing what lies beneath the surface, dental radiographs help dentists formulate accurate diagnoses and develop effective treatment plans tailored to each patient's needs.

There are several types of dental radiographs, each has its advantages and serving specific purposes in diagnosing dental conditions. Some of the most common types include Bitewing X-rays, Periapical X-rays, Panoramic X-rays, Occlusal X-rays and Cone Beam Computed Tomography (CBCT).

Among the various types of intra-oral and extra-oral radiographs, panoramic X-ray images is the best for this kind of studies due to their advantage over other X-ray imaging methodologies. They show all the structures in mouth region including the upper and lower teeth, the positions of emerging teeth, jaws joints, nerves, sinuses and supporting bone. That is, A panoramic X-ray allows the dentist to get an overview of any existing oral health issues. Furthermore, faster operation time, more diagnostic capabilities, less radiation exposure, better patient acceptance, and fewer infection control procedures make it ideal for dentistry.

Dentists face many challenges when manually reading panoramic images. These images capture a wide view of the entire mouth, which can be complex to focus on specific areas of interest. Distortions due to the curved nature of the dental arch can affect the accuracy of measurements and interpretation. Additionally, the superimposition of various anatomical structures can hide important details, making the distinguish between different lesions and tissues is difficult and demand a lot of time. Subtle lesions, such as early-stage caries, small cysts, or fine fractures, can be hard to detect. The quality of images varies depending on the equipment used, the technique of the operator, and the positioning of the patient, with poor quality images hindering accurate diagnosis. Interpreting these images also requires significant expertise and experience, as dentists need to identify normal anatomical variations and pathological conditions accurately. These challenges highlight the importance of training, careful image acquisition, and the use of advanced technologies such as artificial intelligence to assist in the accurate interpretation of panoramic radiographs.

Introducing artificial intelligence (AI) into the field of dentistry holds significant importance due to its potential to transform various aspects of dental care. AI algorithms enhance diagnostic accuracy by automate the analyzing of dental images, early detection of abnormalities such as cavities, periodontal disease, bone loss, and oral cancers, facilitating accurate diagnosis and treatment planning.

Advancements in deep learning algorisms such as convolutional neural networks (CNNs), and computer vision techniques like semantic segmentation, object detection, and classification of lesions offer promising ways for dentists to enhance patient care efficiency. These technologies also enable dentists to see more patients while increases diagnostic accuracy.

Overall, the integration of AI into dentistry promotes higher standards of care, efficiency, and patient satisfaction, making it a vital advancement in the field.

Aim of the study:

The aim of this project is developed and optimize a deep learning based automatic system have the ability to accurately detecting and classifying lesions in dental panoramic radiography.

This project highlights the need to use deep learning techniques to enhance the automation, accuracy, and efficiency of lesion detection and classification in dental panoramic radiography, finally lead to improved patient care and outcomes in dentistry.

Methods:

In this project, a deep learning-based system was developed to automate the detection and classification of lesions in dental panoramic radiography. The Tufts Dental Database (TDD), containing 1000 dental panoramic images, was utilized for this purpose. An object detection framework based on Faster R-CNN and simple CNN was implemented to classify and identify dental lesions. To improve the model's performance and robustness, the dataset underwent augmentation. The model classifies images as normal or abnormal and provides bounding boxes around the detected lesions in abnormal images. The system produced promising results and can be developed for practical use in dental diagnostics.

Chapter 1: Theoretical Review of Medical Aspect

1- Anatomy of the Oral Cavity:

The oral cavity is the first part of the digestive system. It is responsible for speech, digestion, secondary respiration, along with its being a major chemosensory organ, and the use of teeth for digestion. Therefore, any problems that might affect the oral cavity can have serious repercussions on the quality of life [1].

In general, the oral cavity shown in figure (1) is divided into the lips, tongue, jaws (maxilla and mandible) in which the teeth are positioned, and the gingiva (gums) which surround the teeth [2].

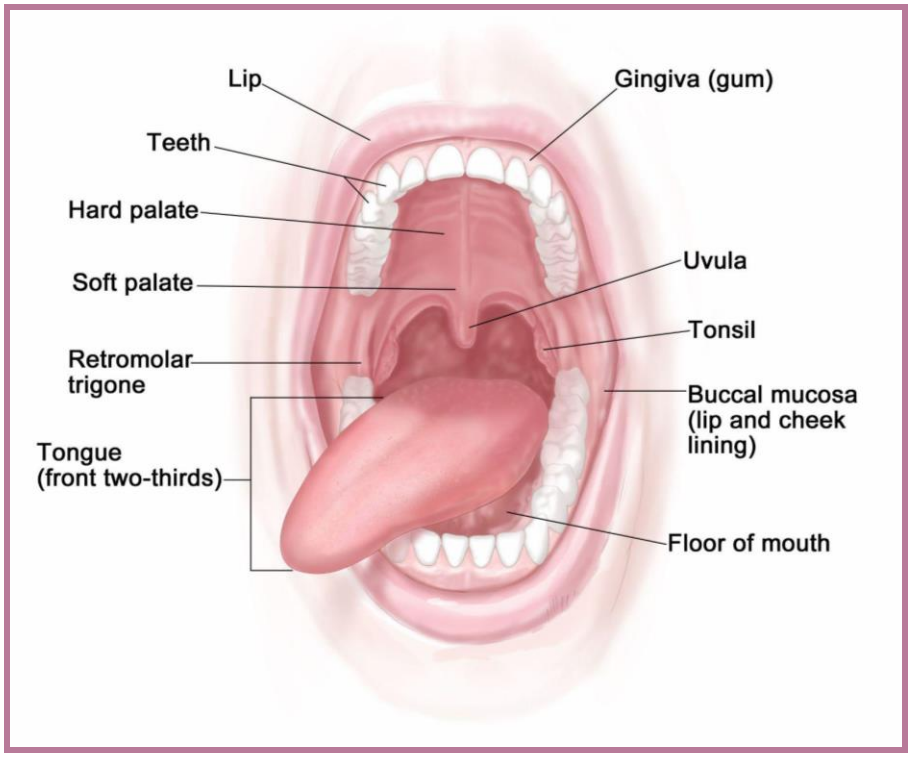


Figure (1): Anatomy of the Oral Cavity [2].

Diving deeper into the dental component of the oral cavity, the average adult has 32 permanent teeth. 16 teeth in the maxilla (upper jaw), and 16 teeth in the mandible (lower jaw). The main purpose of teeth is for digestion through mastication (chewing), but they also aid in speech. Each jaw has four incisors (two central incisors and two lateral incisors), two canines (or cuspids), four bicuspids (premolars), and six molars (including third molars or wisdom teeth).

Histologically, all teeth are composed of four tissues. The tooth's crown is covered by an outer layer that is called the enamel, which is the tooth's first line of defense against caries or erosion and is the hardest tissue of the body. Beneath the enamel is the dentin, which is softer than enamel but still a hard substance. It covers and protects the third tissue which is the center of the tooth - the pulp. The pulp is a fibrovascular structure and is responsible for vascularization, innervation, and repair. Most harm that comes to this tissue is irreversible, deeming the tooth in need of immediate dental treatment. The fourth tissue is found along the root of the tooth and is the cementum and is bound by connective tissue fibers to the alveolar bone (the jaw) to form the periodontal ligament [3].

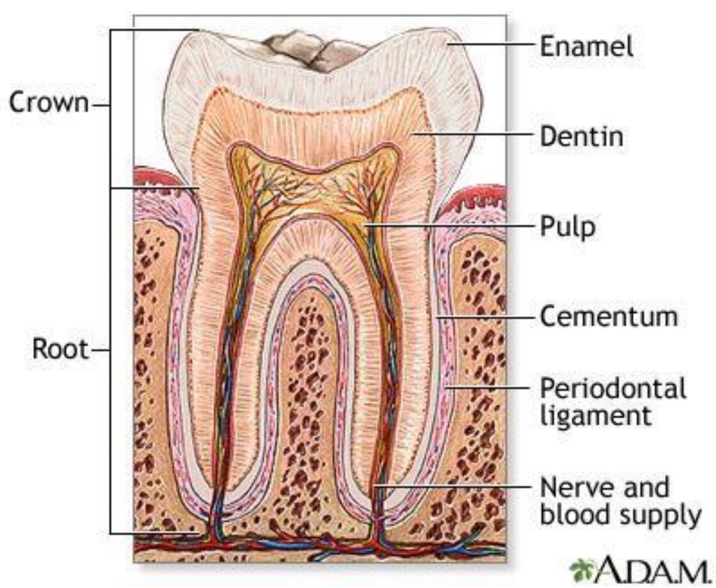
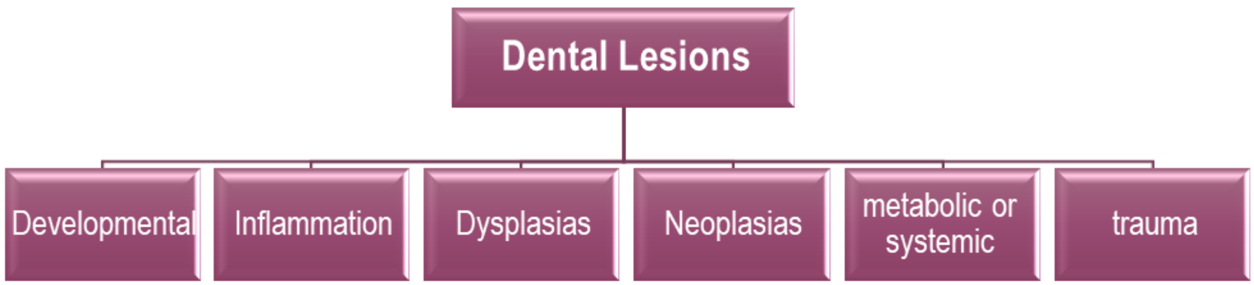


Figure (2): Dental Tissues [3].

2- Pathophysiology of Common Dental Lesions:

As dental health is an essential part of all-around health, it is necessary to be aware of the most common lesions that can affect this area. These lesions can be developmental, of inflammatory origin, traumatic, dysplasias, neoplasias (benign or malign) and can be one side to metabolic or systemic disorders.



2-1 Developmental Lesions:

They are normal lesions that are found from birth or later in life [4].

2-1-1 Dense bone island (Idiopathic Osteosclerosis): is basically asymptomatic radiopaque foci that do not occur because of local infections or systemic disease. These mostly appear in the alveolar processes of the jaws, especially the mandible [5]. They are classified as developmental intraosseous anatomic variations and differ from those of inflammatory origin or systemic disease [6].

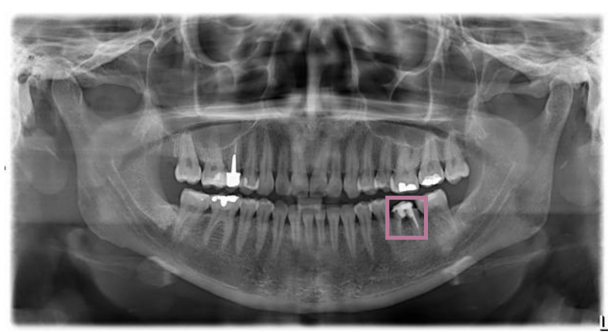
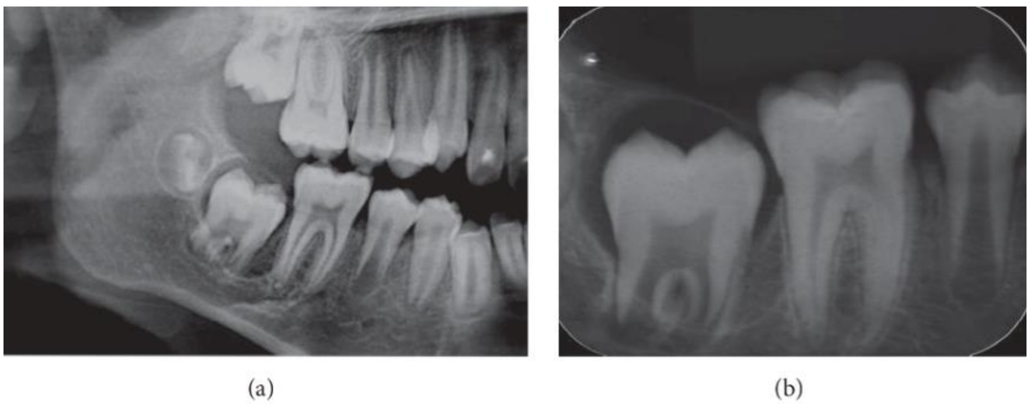


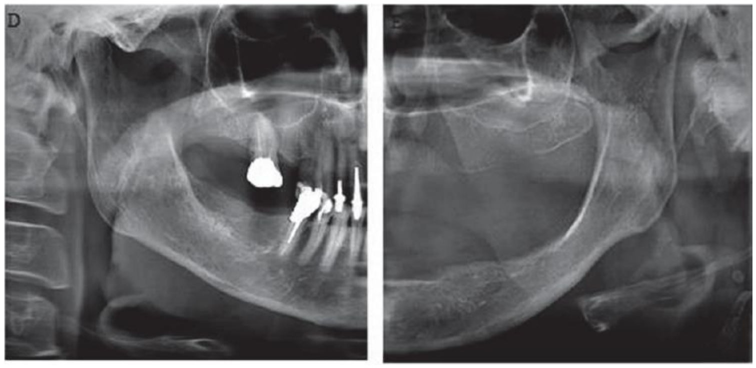
Figure (3): Idiopathic Osteosclerosis [7].

2-1-2 Follicular hyperplasia (hyperplastic dental follicles): are described as the impaction of certain teeth, surrounded with well-demarcated radiolucency [8]. These follicles are pericoronal dental lesions that are either single or multiple and appear in pericoronal tissue of unerupted teeth and they mimic the appearance of dentigerous cysts [9].



Figure(4): (a) Sectioned panoramic radiography: radiolucent well-defined area surrounding the crown of tooth 47, extending to apical region in the anterior area. (b) Periapical radiography: delicate sclerotic border, normal enamel and radicular formation, and absence of visible calcifications in pericoronal space[10].

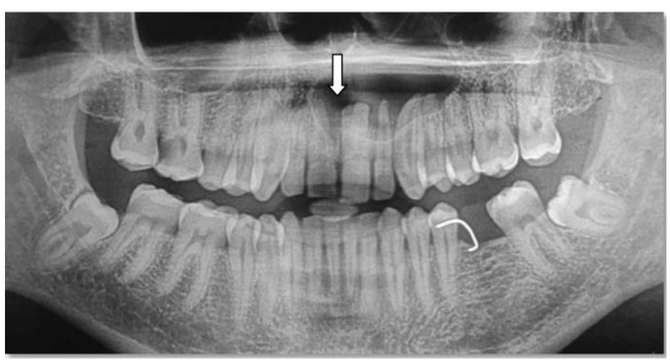
2-1-3 Bone marrow defect: usually an asymptomatic radiolucency (poorly defined margin) that is only discovered incidentally during radiographic examination. It causes complications for dental implants or osseointegration processes. It consists of “hematopoietic red marrow with varying amounts of fatty yellow marrow” [11]. It is basically an unusual hematopoietic tissue in the maxillary bone (maxillary tuberosity), posterior region of the mandible, and the condylar process. In radiographs, it appears as a focal radiolucency with poorly-defined margin and thin central trabeculae [12].



Figure(5): Osteoporosis – Panoramic radiographs indicating generalized radiolucency and lower mandibular cortical presenting erosion [13].

2-1-4 Inverted Mesiodens: is a supernumerary tooth that falls between the two maxillary central incisors and is caused by excess dental lamina during dental development. If the tooth is normal in appearance, then it is a supplementary tooth, but if abnormal then it is referred to as:

- Mesiodens (if between the central incisors).
- Peridens (if near the premolar region).
- distodens (near the molar region) [14].



Figure(6): Panoramic radiograph showing large well-circumscribed, unilocular radiolucent lesion with corticated margins, with radio-opaque irregular structure in relation to maxillary left central incisor[14].

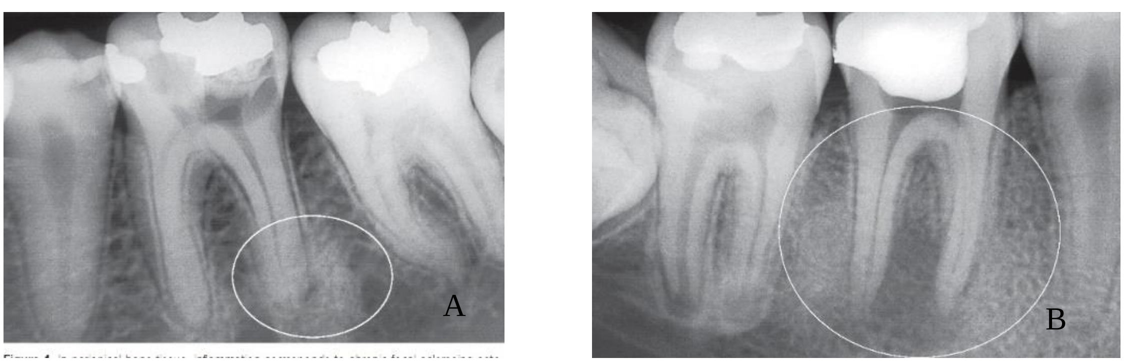
2-2 Inflammation:

Lesions of inflammatory origin can be caused by trauma, substance abuse, poor dental treatment, neglect of dental hygiene, and can also be sequelae to pulpitis or gingivitis. The easiest way to spot inflammation in radiographic imaging is to notice any widening in the periodontal ligament area (PDL).

Some aspects of inflammation are also apical rarefying osteitis and sclerotic osteitis, and more developed stages are **osteomyelitis**. Other specific cases include **peri-implantitis**.

The most common form of inflammation is **periapical granuloma** which is made of granulation tissue and bacteria which is formed after the necrosis of pulp tissue in the tooth that results from caries or trauma [15].

Sclerotic osteitis results from chronic pulpal infections when deep caries or large restorations are present [16], whereas **rarefying osteitis** is given as a name to radiolucent lesions that are periapical abscesses or cysts [17].



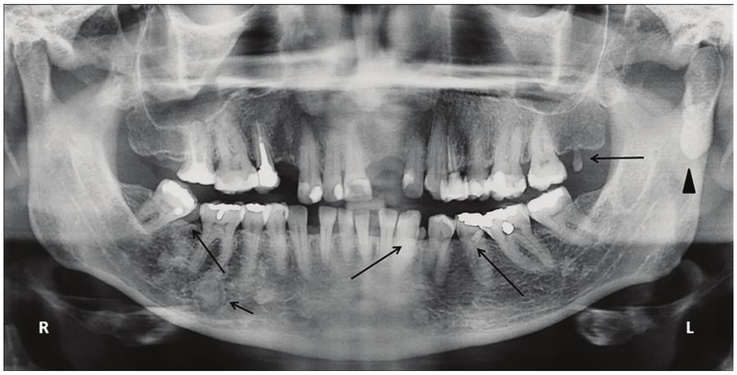
Figure(7): A-In periapical bone tissue, inflammation corresponds to chronic focal sclerosing osteitis. B- In bone environment, the lesion diagnosed as chronic focal rarefying osteitis [18].

2-3 Neoplasia (benign/malignant):

Neoplasias are much more serious, both benign and malignant. Many neoplasias are random and erratic and can be triggered by poor dental hygiene and bad oral habits such as smoking and drug abuse.

2-3-1 Benign Neoplasias and Cysts: include osteoma, odontogenic keratocyst, dentigerous cyst, odontoma, mucous retention pseudocyst, ameloblastoma, and incisive canal cyst.

- ***Osteomas:*** are benign neoplasms and evolve from the proliferation of mature compact bone, and they are most common in the mandible. They are slow growing with unknown etiology [19].



Figure(8): The presence of osteoma in the left condylar region (arrow head)[19].

- ***Odontogenic keratocysts (OKC):*** are also benign but they are aggressive intraosseous tumors, and they have the ability to become cancerous. It often has no symptoms at first which proves the need for routine radiographic imaging [20].

Figure(9): Odontogenic keratocyst in a 36-year-old woman. Panoramic radiograph shows an ellipsoid, expansile, well-corticated , lucent lesion in left mandibular body (arrow) [21].

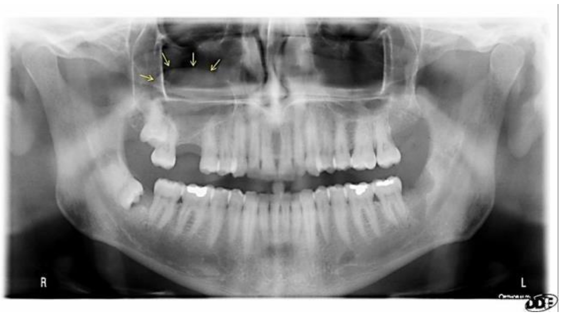
- ***Dentigerous cysts:*** accompany teeth that are unable to complete eruption and are found around the crown and have fluid accumulation. They have well-defined sclerotic borders and well-demarcated unilocular radiolucencies surrounding the crown [22].

Figure(10): Panoramic X-ray of Dentigerous Cysts [22].

- ***Odontomas:*** are “hamartomas” or aborted tooth formations and are common benign odontogenic tumors and are growths with both the epithelial and mesenchymal components completely differentiated [23].

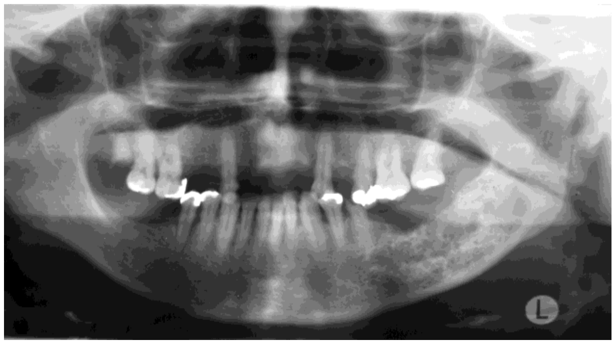
Figure(11): Giant compound odontoma [23].

- ***Mucous retention pseudocysts***: are benign and self-limiting. They result from outflow of mucus due to duct obstruction and they are of non-odontogenic origin. Radiographically, they are well-defined, homogenous, circular radiopacities located mostly on the floor of the maxillary sinus [24].



Figure(12): Mucous retention pseudocyst of the right maxillary sinus on panoramic view [24].

- ***Ameloblastomas***: are benign and odontogenic and originate from the epithelium of tooth germ. Large sizes in the mandible can cause facial asymmetry, dental displacement, and malocclusion [25].



Figure(13): Panoramic radiograph depicting a large ameloblastoma of the left mandible [25].

- ***Incisive canal cysts***: (nasopalatine duct cyst) are of embryologic origin and develop in the midline of the anterior maxilla. They are usually asymptomatic and develop in a single location [26].

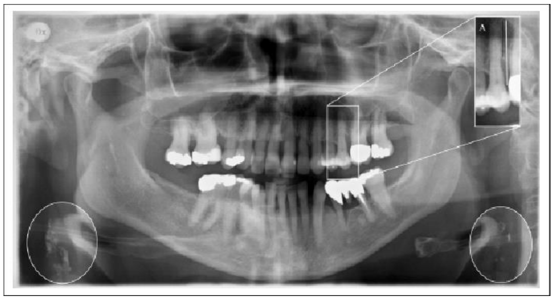
2-3-2 Malignant Neoplasia:

Malignant neoplasia are cancerous and could be mistaken for benign which enforces the need for proper detection and diagnosis. Common findings may be rapid growth, ill-defined margins, bone absorption, bleeding, heterogeneous tissue upon biopsy [27].

Figure(14): Cropped panoramic radiographic image showing the tumor extending from the right first premolar up to the second molar [27].

2-4 Metabolic/Systemic:

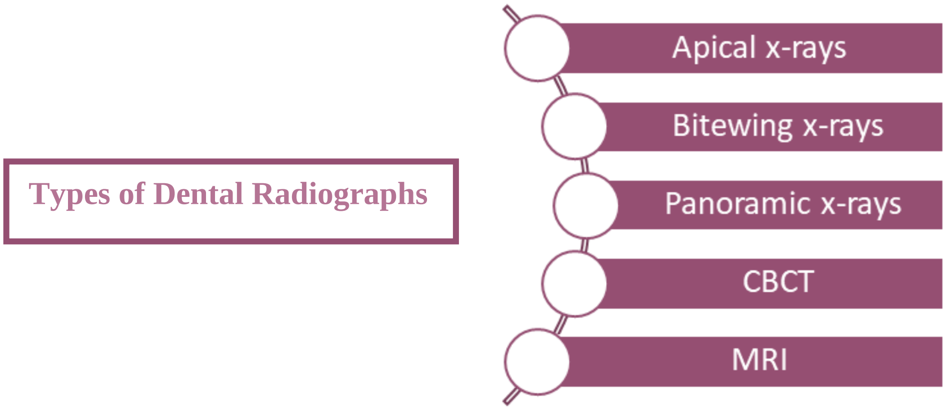
Many metabolic or systemic diseases can have oral manifestations that can be detected radiographically. These include osteopenia/osteoporosis and calcified carotid atheromatous plaque [28].



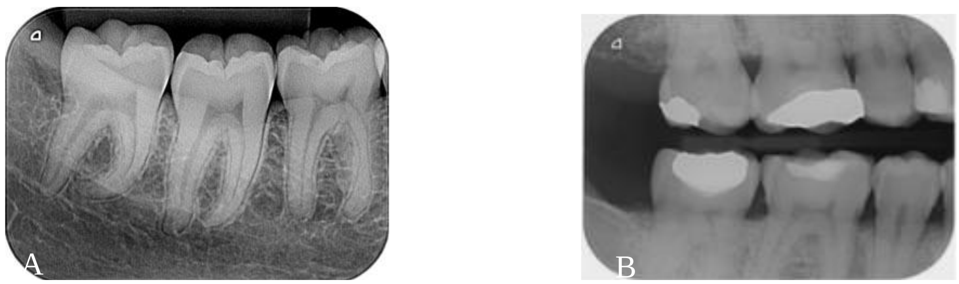
Figure(15): Panoramic radiograph shows bilateral calcified carotid artery atheromas (white rings) [28].

3- Types of Dental Radiographs:

For diagnosis of most lesions in the oral cavity, a clinical examination is necessary but not always sufficient, since many lesions are simply hidden beneath the mucosal/bone structure. Therefore, several imaging techniques are used for diagnosis in dentistry. The most common are apical x-rays, bitewing x-rays, and panoramic x-rays. Also for benefit are cone-beam CT (CBCT) and magnetic resonance imaging (MRI) [29].

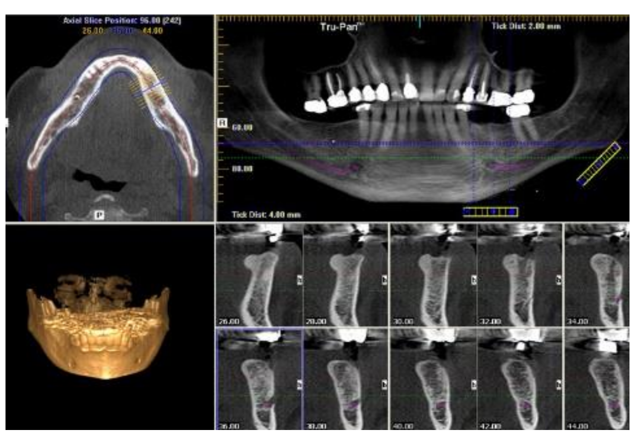


3-1 Periapical and bitewing x-rays: both use x radiation and only cover a limited area of the oral cavity, but they ensure good quality and resolution for endodontics use, which doesn't make them ideal for the detection of large structures in the maxillofacial area. The main difference between bitewing and periapical is the angle and positioning of the film, which offers a different focus on certain aspects between the two Image [29].



Figure(16): The different kinds of dental radiographs. (A) periapical x-ray, (B) bitewing x-ray [29].

3-2 Cone-Beam CT (CBCT): is a more advanced technique that is three-dimensional in range and is very useful in implant-planning and complex cases and diseases [30].

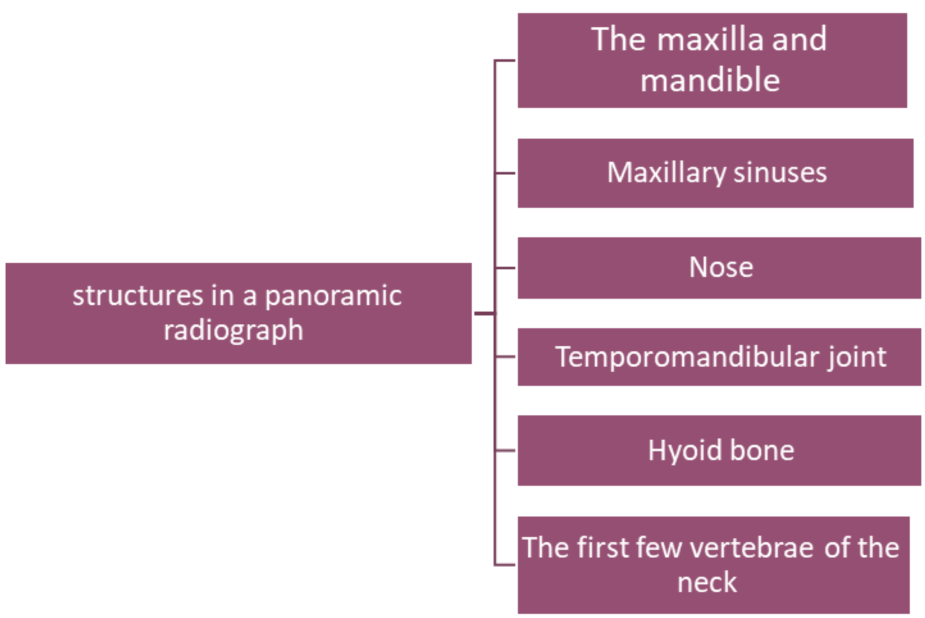


Figure(17): Cone Beam CT Scanning [30].

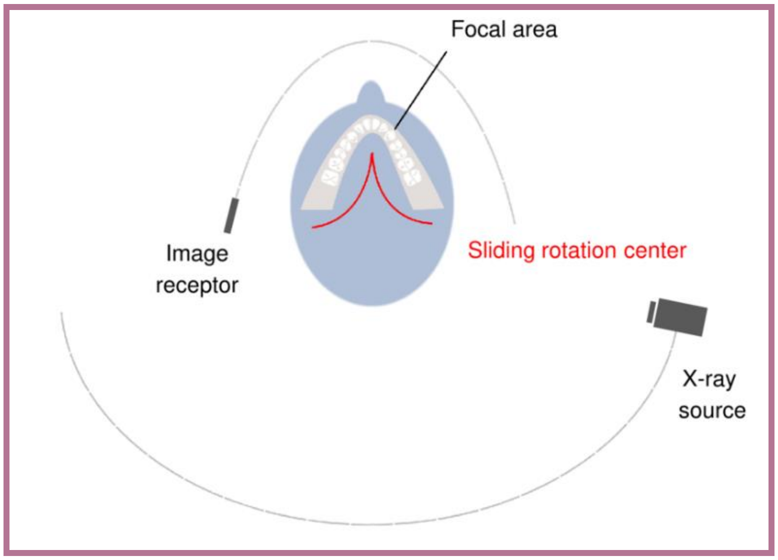
3-3 Magnetic Resonance Imaging (MRI): though not commonly used, is helpful in detection of abnormalities of the soft tissue and temporomandibular joint (TMJ) [30].

3-4 Panoramic X-Rays: which is the main focus of this study, they offer a much wider view of the oromaxillofacial area than that covered in bitewing or periapical x-rays. The technology used includes a device that revolves around the head (hence the name panorama from Greek origin pan παν meaning 'all' + horama όραμα meaning 'view'). It includes an x-ray tube which emits ionizing radiation, and an x-ray film or detector on the opposite side of the tube. The imaging results after the radiation passes through the patient while the tube revolves in a semicircle around the patient from one side of the jaw and ending with the other [31].

The structures detected in a panoramic radiograph include:



Indications of panoramic radiographs include the diagnosis of periodontal disease, orthodontic evaluation, diagnosis of lesions like cysts and tumors, and implant planning [32].



Figure(18): The working principle of panoramic imaging [32].

3-4-1 The advantage of Panoramic Radiography:

The main advantage of using panoramic radiography, as described above, is the capability of capturing a wider area consisting of multiple tissues that might be previously unnoticed. Unlike periapical x-rays for example, in which only a small area is covered, it is rarer for coincidences to capture previously unnoticed lesions.

Besides panoramic radiography's main advantage, it is also beneficial for its somewhat low radiation dosage. It is easier than intra-oral radiography for patients with gag reflexes, and the device can be available in many dental practices. It also facilitates use because there is no need for a technician to be close by the patient for proper positioning of the film.

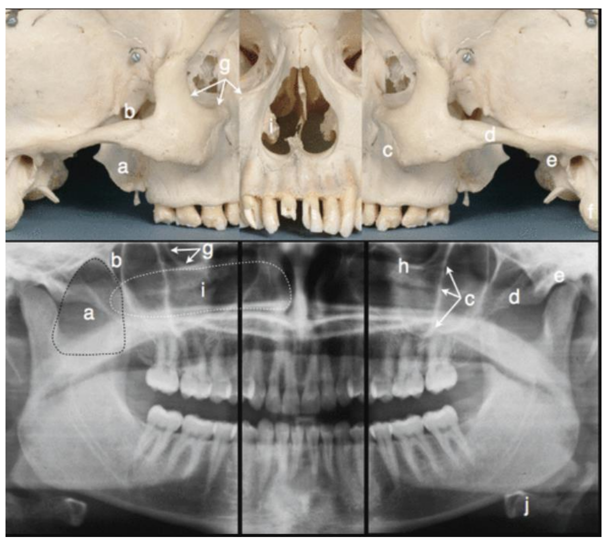
All cases, however, cannot rely solely on a panoramic radiograph for correct diagnosis. It can only direct attention to previously unnoticed structures. Most lesions require biopsies, clinical history, and CBCT for correct diagnosis to ensure that the nature of the lesion is of benign or malignant nature [33].

3-4-2 The Interpretation of Dental Radiographs:

Proper interpretation of dental radiographs, no matter the type or device used, is always achieved by a dental professional. Since panoramic radiographs contain a wider coverage and more varied structures, it requires special attention to detail. Each professional reaches a different methodology regarding their interpretation sequence in reading the image. The example in figure (19) may include:

1. begin with the corners of the image and peripheral structures (because the main focus of a dentist is the teeth, one can start with everything else to avoid neglect) such structures include the temporomandibular joint TMJ, the cervical spine, the hyoid bone, and even the orbital bone.
2. The outer cortex of the mandible to assess the continuity of the bone and any abnormalities in a circular fashion from one end to the other and back.
3. Assessment of the maxilla cortex which also includes the maxillary sinus which is a structure of major importance.
4. Zygomatic bones and arches.

5. Density of the sinuses (could indicate inflammation or more serious conditions).
6. Nasal cavity and palate.
7. Bone pattern and density of maxilla and mandible (metabolic disease could be indicated) along with examining the mental foramen, inferior alveolar canal, and mandibular foramen.
8. Finally, one can begin assessing the teeth and alveolar processes along with follicles of any unerupted teeth [33].



Figure(19): Angles taken in the panoramic radiograph and anatomical structures covered [33].

3-4-3 Current Approach in the use of Panoramic Radiography and the new shift toward AI:

Panoramic radiography is not the most high-definition radiography method available now, but its benefit in offering a wide area of coverage makes it ideal in taking regular radiographs for reference and routine checkups at the dental office, whereas intra-oral radiographs are more targeted and are focused on a certain condition, which often leads to neglect of other structures in the oral cavity. Routine panoramic radiographs are more common nowadays, which makes tracking of certain lesions over time more accessible. The purpose of such radiographs is often superficial but can (with the expert's eye) lead to uncovering lesions or conditions of the cavity that often lie unnoticed till after it is too late for intervention. For example, many malignant lesions begin small in size and remain silent in regard to symptoms, but when they reach a bigger size that makes them noticeable in a clinical manner, it is often too late for safe intervention and good prognosis. Therefore, many practitioners have started to recommend routing panoramic radiographs to keep record on any changes and to quickly step in when deemed necessary.

Here we suggest an automated method for detection of such lesions, since panoramic radiographs have many variables that can sometimes intervene with the expert's eye for detail, such variables which may be due to image settings (gamma correction, contrast, brightness), and artefacts due to poor exposure, movement during imaging, issues such as ghost images from jewelry or piercings, and air bubble from the airways or even from the temporomandibular joint.

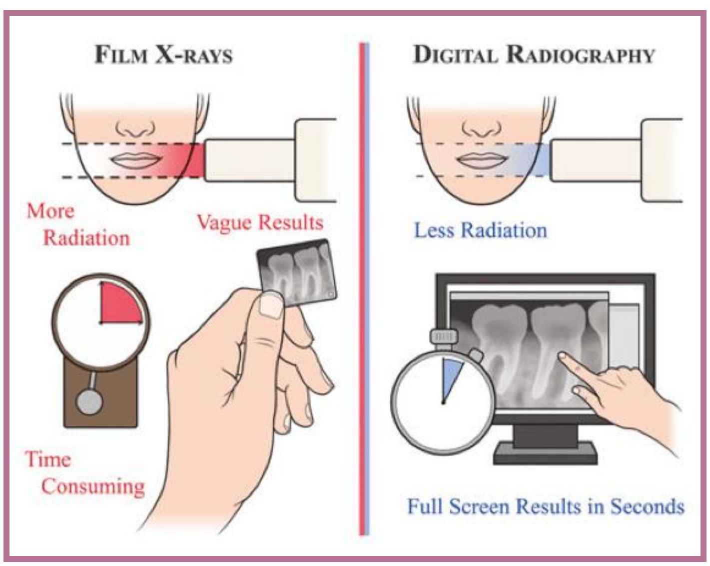
Since the expert's eye may not always be able to adjust to all these variables with human capacity, it was suggested by that models can be trained to learn the many variables present in a panoramic radiograph, and then be able to adjust and adapt with the differences between images to detect anomalies in the oral cavity [34].

4- Traditional Approaches for Lesion Detection and Classification:

The traditional methods for detecting and classifying lesions in dental radiography depended on dental professionals' expertise and the use of various imaging techniques and interpretative methods. These are some of the main traditional ways:

4-1 Visual Inspection of Radiographs:

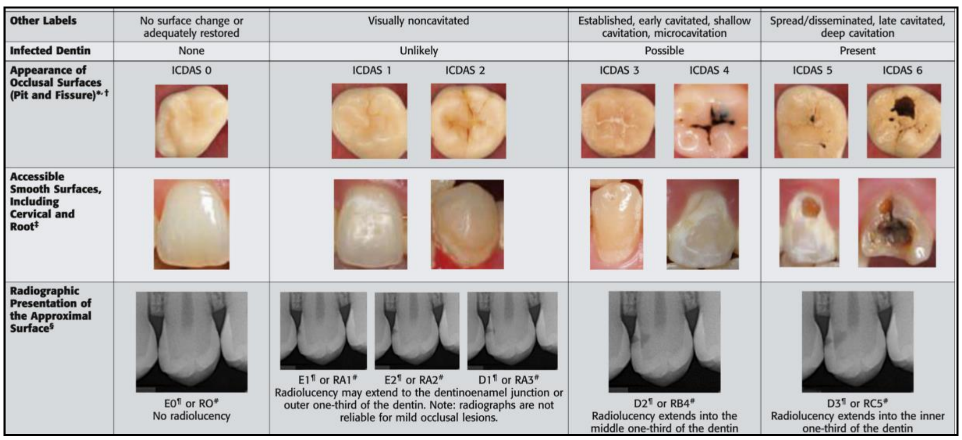
- **Manual Review:** Dentists and radiologists manually review dental X-rays, such as bitewing, periapical, and panoramic radiographs, to identify caries, periodontal disease, bone loss, cysts, tumors, and other abnormalities.
- **Film Radiography:** requires careful visual inspection on a lightbox. The dentist looks for variations in the density and contrast of the film to detect lesions [35].



Figure(20): The differences between Film X-Ray and Didital Radiography [36].

4-2 Standardized Criteria:

- **Caries Classification Systems:** Systems like the International Caries Detection and Assessment System (ICDAS) and the Radiographic Caries Classification (RCC) provide standardized criteria for classifying caries based on their appearance in radiographs.
- **Periodontal Classification**: Periodontal bone loss is classified based on the extent and pattern of bone destruction visible on radiographs [35].



Figure(21): American Dental Association Caries Classification [37].

4-3 Magnification Tools:

magnifying glasses or loupes, these tools help in examining radiographs more closely to identify subtle changes in tooth and bone structures.

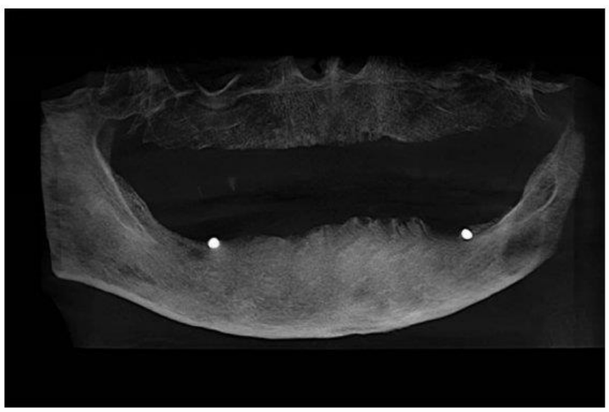
For high-detail analysis microscopes are used [35].

Figure(22): Magnifying Glasses [35].

4-4 Comparison with Previous Radiographs: to detect changes over time, such as the progression of caries or bone loss.

4-5 Diagnostic Aids:

- **Radiopaque Markers:** Sometimes used to help delineate areas of interest on radiographs.
- **Contrast Media:** In some cases, contrast agents are used to enhance the visibility of certain structures or lesions [35].



Figure(23): Denture with radiopaque markers used as template [38].

4-6 Training and Experience:

Dentists and radiologists undergo extensive training in radiographic interpretation to develop the skills needed to detect and classify lesions accurately. In addition, ongoing education to stay updated with the latest advancements and techniques in radiographic interpretation [35].

4-7 Symptom Correlation:

Correlating clinical findings and patient-reported symptoms taking into consideration the patient's overall health and dental history, with radiographic findings to make a more accurate diagnosis.

4-8 Consultation and Second Opinions:

- **Peer Review:** Consulting with colleagues or specialists to get second opinions on difficult or ambiguous cases.
- **Interdisciplinary Teams:** Collaboration between general dentists, radiologists, periodontists, endodontists, and other specialists to ensure comprehensive evaluation and diagnosis.

These traditional methods require a high level of expertise and are often time-consuming. While they have been effective, the subjectivity and variability in interpretation can lead to inconsistencies. The integration of AI in dental radiography aims to reduce these inconsistencies by providing objective, consistent, and automated analysis, improving diagnostic accuracy and efficiency [35].

Summary:

This chapter includes an explanation of the anatomical structures in the oral cavity, mentioning the most important types of dental lesions, their characteristics and features, in addition explaining the types of medical images used in this field, with emphasis on the details of panoramic radiography in terms of the characteristics and advantages, it also includes the traditional approaches for lesions detection and classification.

Chapter 2: Artificial Intelligence in Dentistry

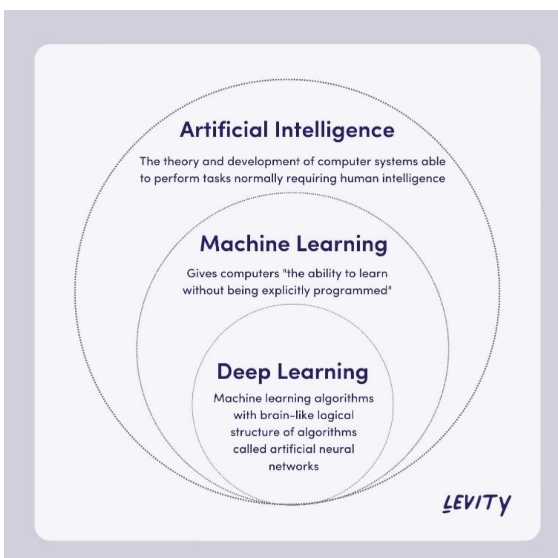
5-The definition of Artificial Intelligence:

Artificial intelligence (AI) means making machines that can think and learn like humans. Expert systems, natural language processing (NLP), speech recognition, and computer vision are examples about AI applications. Typically, AI systems operate by processing large amount of labeled training data, analyzing that data for correlations and patterns, and utilizing these patterns to predict future outcomes. In many fields, AI can perform tasks more efficiently and accurately than humans [39].

Machine Learning is computer learning from data by algorithms to perform tasks without being explicitly programmed. It describes the relation between computer science and statistics.

Deep Learning includes algorithms that analyze data like the human brain. This enables algorithms to process unstructured data such as images, documents, and text.

deep learning is a specialized subset of machine learning which, in turn, is a subset of artificial intelligence [40].



Figure(24): Machine Learning vs Deep Learning vs Artificial Intelligence [40].

5-1 Comparison of Machine learning and Deep learning:

Machine learning (ML) and deep learning (DL) are both subsets of artificial intelligence (AI) with different structures and applications. ML involves simpler models such as decision trees and support vector machines, which work better with smaller datasets and require manual feature extraction with less computational power making them faster for training. These models are often used for predictive analytics, decision support systems, and administrative tasks like appointment scheduling and patient record management. In contrast, DL uses neural networks with multiple layers, which automatically extract features from raw data and achieve superior performance in complex tasks involving large datasets that need huge computational power. DL models, particularly convolutional neural networks (CNNs), are better in analyzing dental radiographs and 3D images, identifying abnormalities, segmenting different structures, and providing highly accurate diagnostics [41-42].

Table (1): Comparison of Machine learning and Deep learning [40].

Difference	Machine learning	Deep learning
Application scenarios	Fewer	Wider
Data volume requirements	Smaller	Larger
Data dependency	Less	Higher
Hardware dependence	Equipment requirements	High equipment

		requirements
Feature processing	Manual extraction	Automatic execution
Execution time	Short	Long
Problem-solving method	Split and solved collection	End to end

5-2 The Role of Artificial intelligence in Dentistry:

Artificial intelligence (AI) is expanding and making significant advancements in all fields including dentistry, offering numerous benefits and applications. because it can learn from large data and human expertise to make works which needs human intelligence.

AI is considered a valuable tool to help dentists and clinicians in many aspects. where applications of AI can be classified into diagnosis, decision-making, treatment planning, and prediction of treatment outcomes. diagnosis is the most popular and AI can make accurate and efficient diagnoses, that help dentists and reduce their efforts, for that AI applications have spread over all fields in dentistry, like operative dentistry, periodontics, orthodontics, oral and maxillofacial pathology, and prosthodontics. AI has demonstrated accuracy and precision. However, before incorporating AI models into routine clinical operations, it is still important to further certify the cost-effectiveness, dependability, and applicability of these models.

Artificial intelligence is one of the most promising technologies that give impressive results if algorithms are properly trained on unbiased training data [43-44].

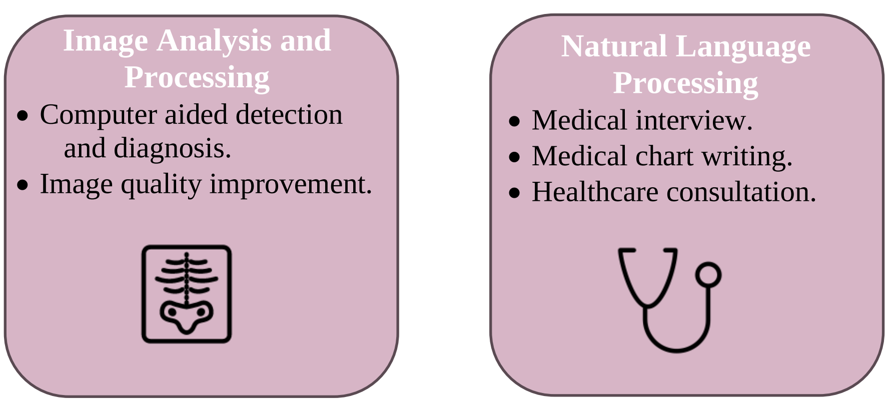
Figure(25): The Application prospects of Artificial Intelligence Technology in Dentistry [43].

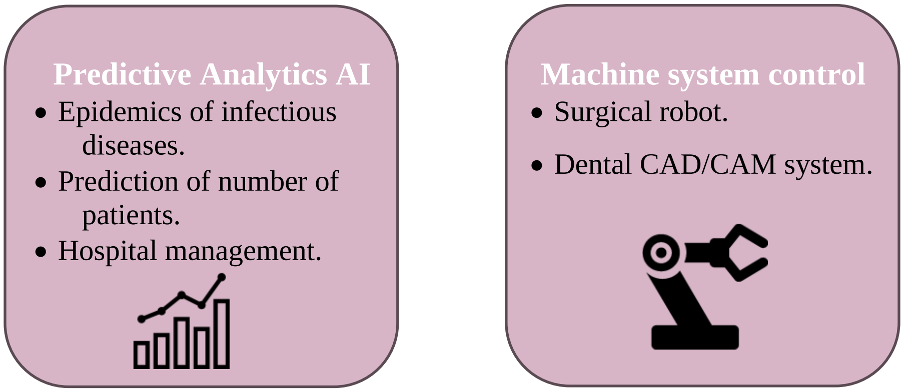
At the beginning, classical Machine Learning (ML) techniques were used in many applications at the field of dentistry such as mathematical morphology and active contour have been used for teeth segmentation. Bayesian techniques, linear models, and support vector machines with hand-crafted features have been used for classification tasks. However, these methods need to extract features from raw data and transform them to a suitable representation for the algorithms to detect or classify the input images. The efficiency of these methods greatly depends on the image preprocessing techniques used and the quality of the extracted features. The methods give favorable results when the test data is small while they perform poorly with large dataset. The limitations of the traditional approaches were overcome with the introduction of Artificial Intelligence (AI) based on deep learning that has the ability to consider raw image data as input, process this data in different layers to automatically identify the suitable representations, and produce a meaningful output in the form of detection or classification. The main advantage of the deep learning system over the classical ML systems is its ability to learn features directly from the raw data instead of the hand-crafted features [58].

6- Deep Learning and its Applications in Dentistry:

Deep learning (DL) is a computer technology that has triggered the development of AI applications. DL is a subset of machine learning method that uses Neural Networks. A Neural Network is a set of algorithms having mathematical models that mimic the networks of neurons in the brain and is widely studied since the 20th century [45].

Deep learning, has shown remarkable success in various domains, including dentistry. It performs well in difficult tasks lead to accurate identification of oral conditions and abnormalities [43].





Figure(26): Possible clinical application of deep learning in dentistry [43].

Some notable applications and benefits of deep learning in dentistry:

1. Radiographic Image Analysis and Diagnostics:

Automated Detection with deep learning algorithms which can analyze dental radiographs (X-rays) to detect cavities, periodontal disease, periapical lesions, and other dental lesions with high accuracy. These systems can help dentists to identify problems that might be missed during a manual examination. Another application is segmentation: it can segment different parts of dental X-rays, such as teeth, bones, and lesions, to provide more accurate diagnostics.

2. 3D Imaging and Orthodontics:

Deep density and enhance the interpretation of Cone Beam Computed Tomography CBCT scans, aiding in the detection of abnormalities, assessing bone density, and planning surgical procedures. The algorithms can also assist in the analysis of 3D models from intraoral scanners, improving the precision of orthodontic treatment planning.

3. **Predictive Analytics:**

   Using patient data, deep learning models can predict the risk of dental caries (cavities) development. It can predict the success of different dental treatments based on historical data, helping dentists to choose the most effective treatment plans for their patients.

4. **Custom Prosthetics:**

   Deep learning can help in designing personalized dental prosthetics, such as crowns, bridges, and dentures, ensuring a better fit and improved patient comfort.

5. **Workflow Automation:**

   this includes organizing administrative tasks in dental practices and driven virtual assistants to improving overall patient experience.

6. **Educational Tools:**

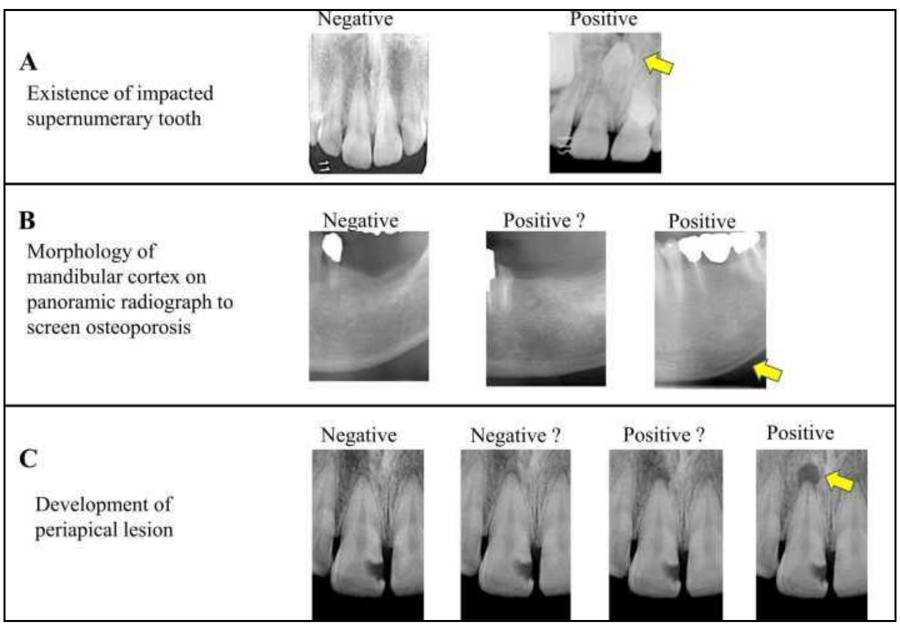
   It can be used to develop advanced training simulations for dental education, allowing students to practice procedures in a virtual environment then evaluates the performance of them and providing detailed feedback to help them improve their skills.

Deep learning is transforming dentistry by its ability to handle complex data and integrate with modern dental tools makes it an invaluable asset in improving the overall quality of dental care. However, there are still many challenges that need to be addressed [43].

6-1 Deep Learning and its Applications in Dental Diagnostic Imaging:

In dentistry, imaging plays a crucial role in both diagnosis and treatment therefore dental diagnostic imaging is the most popular application of artificial intelligence AI based on deep learning. Many tasks have been applied in this fields:

6-1-1 Classification Task: it categorizes dental images into predefined classes or categories. These tasks typically aim to identify and differentiate between various types of dental conditions, structures, or abnormalities or to evaluate the progress of a lesion based on imaging findings [41].



Figure(27): Examples objectives for classification task [41].

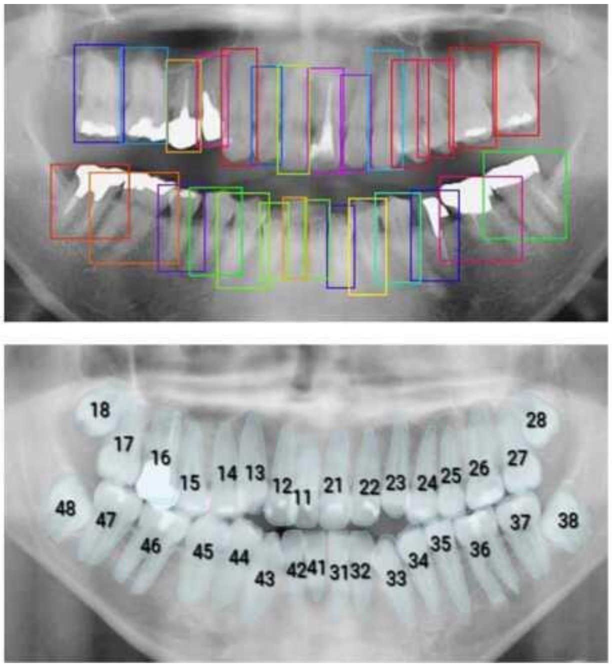
6-1-2 Region (Object) Detection Task:

Region detection in dental radiography involves identifying and locating specific objects or regions within the image. With the aim of detecting the presence and position of various dental structures or abnormalities. for examples:

- **Tooth Detection:** Identifying the location and boundaries of individual teeth.
- **Caries Detection:** Locating areas affected by dental caries [41].

Region detection of teeth

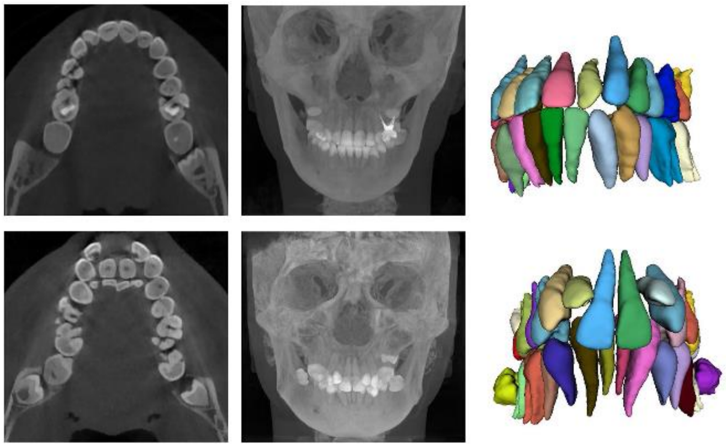
Segmentation of teeth



Figure(28): Automatically Region Detection and Segmentation of Teeth in Panoramic Radiography [41].

6-1-3 Segmentation Tasks:

It involves dividing the dental radiographic image into meaningful segments, usually to outline specific structures or regions in detail. Segmentation provides more precise information about the shape, size, and boundaries of the detected regions. like tooth segmentation and dental lesions segmentation [41].



Figure(29): Automatic Segmentation of Individual Tooth in Dental CBCT Images From Tooth Surface Map [41].

Deep learning methods can also be used for detecting and evaluating anatomical structures of interest from images. Furthermore, generative AI based on natural language processing can automatically create written reports from the findings of diagnostic imaging [41].

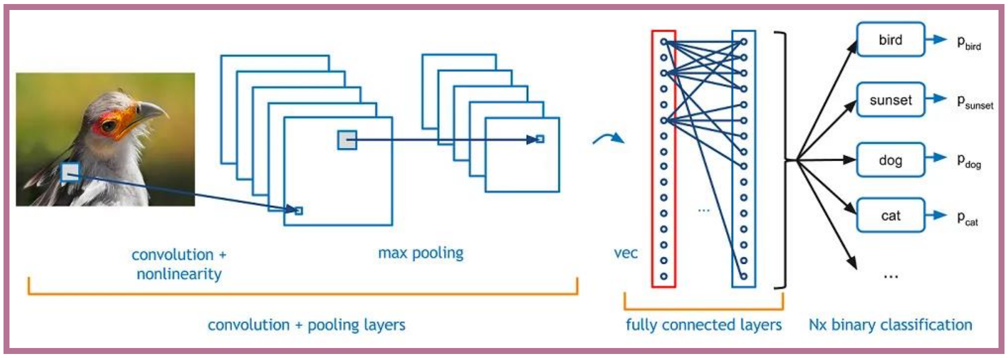
6-2 Deep Learning Architectures for Image Analysis in Dentistry:

Deep learning has revolutionized various fields, including dentistry, by providing powerful tools for image analysis. the key deep learning architectures used for image analysis in dentistry:

6-2-1 Convolutional Neural Networks (CNNs):

CNNs are increasingly utilized for dental image diagnostics, assisting dentists in a more comprehensive, systematic, and faster evaluation and documentation of dental images. CNNs can perform tasks such as:

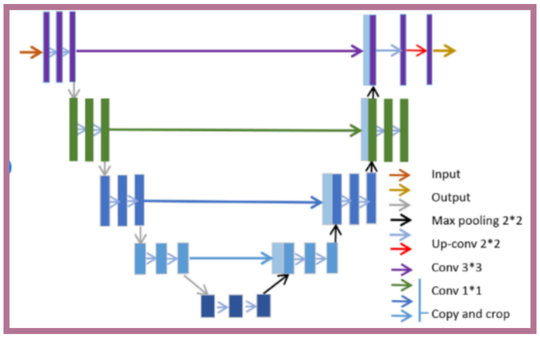
1) detecting structures (identifying the presence of a tooth in an image) or pathologies (detecting caries lesions on teeth).
2) segmenting these structures (determining the exact shape of the tooth).
3) classifying them (labeling each tooth in a dentition and distinguishing between acquired and developmental enamel white spot lesions) [46].



Figure(30): CNN Architectures [46].

6-2-2 U-Net:

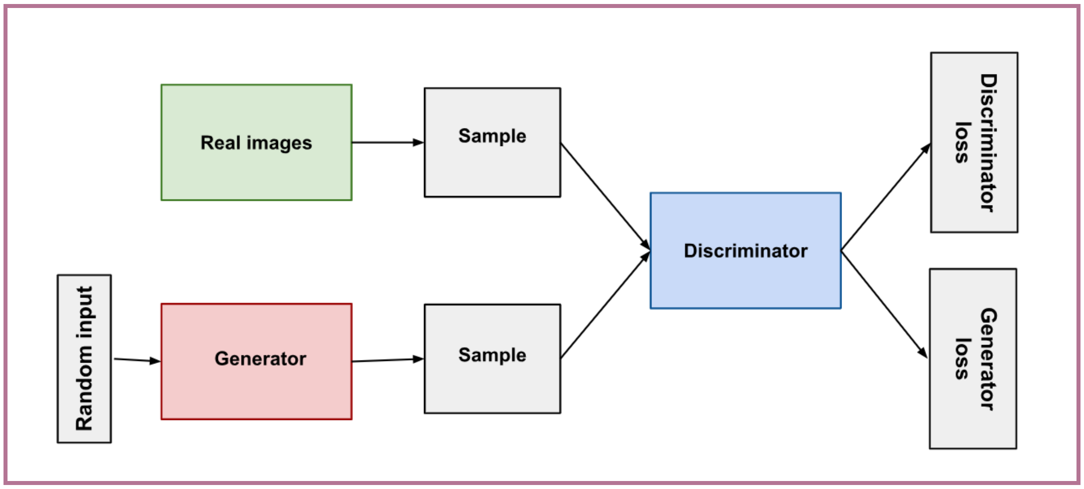
U-Net architecture originally designed for biomedical image segmentation, it has seen many variations and advancements aimed at improving its performance. in dentistry it uses in segmentation of dental structures in 3D scans, precise extraction of tooth and bone regions from radiographs [47].



Figure(31): The architecture of Unet [47].

6-2-3 Generative Adversarial Networks (GANs):

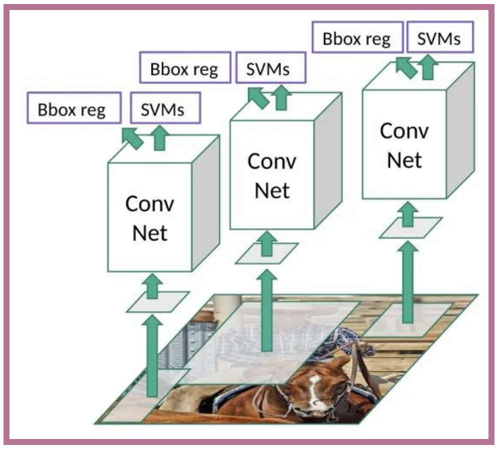
GANs were used for different imaging forms, including two-dimensional and three-dimensional images. In dental imaging, GANs were utilized for tasks such as reduction, denoising, and super-resolution, image enhancement, image generation for augmentation, outcome prediction, and identification. The generated images were incorporated into tasks such as object detection and classification [48].



Figure(32): GAN Architecture [48].

6-2-4 Region-based CNNs (RCNNs):

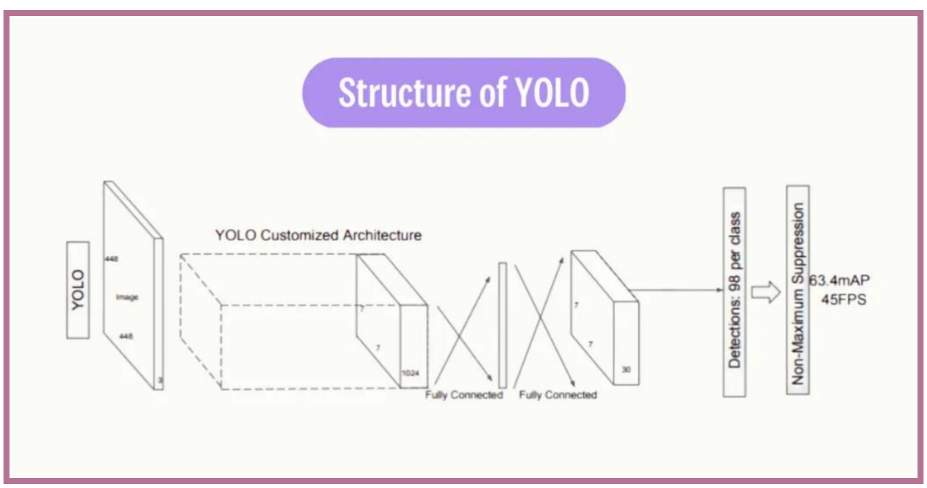
Region-based Convolutional Neural Network was one of the pioneering models that helped advance the object detection field. In dentistry it plays a significant role by enhancing the precision and efficiency of dental image analysis, where it works on detecting and localizing specific dental structures and pathologies within images. It also has an important role in segmentation tasks of dental images. There are many types of R-CNNs like, Fast R-CNN, Faster R-CNN, Mask R-CNN, R-FCN (Region-based Fully Convolutional Network), and Cascade R-CNN [49].



Figure(33): R-CNN Architecture [49].

6-2-5 YOLO (You Only Look Once):

It is a popular Real-time object detection model known for its speed and accuracy. It has been used for detection of dental anomalies during clinical procedures [50].



Figure(34): YOLO Structure [50].

There are many deep learning tools that hold great promise for enhancing the accuracy, efficiency, and accessibility of dental diagnostics and treatments, ultimately leading to improved patient care.

Summary:

This chapter includes a definition of artificial intelligence and its subsets, with a comparison between machine learning and deep learning, and the role of each in the field of dentistry, in addition to explaining the most important applications in the field of dental diagnostic imaging and the most important algorithms used for this purpose.

Chapter 3: Literature Studies

7- Reference studies:

Hyunwoo Yang and Eun Jo [51] introduced A deep learning tool for real-time object detection by using YOLO (You Only Look Once) algorithm, they worked on 1602 lesions on panoramic radiographs, and it is the first study that comprising both maxilla and mandible datasets. The images classified in for categories: dentigerous cysts, odontogenic keratocyst, ameloblastoma, and no lesion. They compared the performance between YOLO, oral and maxillofacial surgeons, and general practitioners where the differences were statistically insignificant. The results for YOLO were (precision = 0.707, recall = 0.680). Detecting and classifying performance was measured in many ways to evaluate the suitability of YOLO as a CAD system.

Ba-Hattab R, Barhom N et al. [52] proposed an AI model made of two convolutional neural networks (CNNs) based on “Faster-RCNN” and Inception-v3, one for detection and the other for classification. The dataset contains 713 panoramic radiographs with 18618 periapical root areas (PRA). The detector found the locations of the periapical region of the teeth using a bounding-box, while the classifier classified the extracted PRAs into healthy and periapical lesions (as PL or not-PL). Results after integrating both detection and classification models, the proposed method accuracy, sensitivity, and specificity were 84.6%, 72.2%, and 85.6%, respectively.

Langston Nashold et al [53] proposed two hybrid approaches for the analysis of dental panoramic radiographs: the first approach is multi-head CNN model with one head for teeth segmentation and another for abnormality detection, While the fusion approach, worked on abnormality detection by using teeth segmentation masks or gaze plots in addition to the radiographs. The output was predicted binary teeth mask with binary classification for whether the given radiograph contains an abnormality or not. Tufts Dental Database was used for this purpose. ResNet-50 for teeth segmentation and training a standard ResNet-18 for abnormality detection gave better results.

Çelik, B et al [54] used 454 objects in 357 panoramic radiographs. Also 10 different deep learning-based detection frameworks were applied to periapical lesion detection problem. Detection performance, mean average precision, varied between 0.832 and 0.953 while accuracy was between 0.673 and 0.812 for all models. F1 score was between 0.8 and 0.895. RetinaNet performed the best detection performance, After the training was completed, they were tested with an external public data set Tufts Dental Database with 41panoramic radiographs to compare their performances.

Sangyeon Lee et al [55] built a model to detect 17 fine-grained dental abnormalities. they used about 23,000 anonymized panoramic dental images taken from local dental clinics. this model used faster R-CNN to detect these abnormal signs and filter out normal images with high sensitivity of about 0.99.

Ali Mahran et al [56] presented the importance of U-Net architecture with teeth segmentation. The model demonstrated atypical performance through training and testing on Tufts Dental X-Ray Dataset with 10-fold cross-validation, it achieved an average dice coefficient of 95.01%, intersection over union of 90.6%, and pixel accuracy of 98.82%. It outperformed other models that were applied to the same dataset.

Giulia Rubiu et al [57] proposed a model based on Mask Region-based Convolutional Neural Network (Mask-RCNN) for instance, segmentation of the different teeth in panoramic dental X-rays. the Tuft dental database was used for this goal with predicted label (20 deciduous and 32 permanent). the percentage of correctly classified teeth was 98.4%, while the Dice score was 0.87.

Conclusion:

Panoramic radiographs are routinely used in dental practice due to their ability to capture a vast area of the oral cavity, which allows easy examination of the complete dentition. However, the ability of dental professionals' abilities to read panoramic radiographs usually varies, which is affected by their individual skills and expertise.

The remarkable developments in AI, especially in deep learning models, has yielded important results that mostly exceed human experts, this is because deep learning has the ability to learn from data with high accuracy in detection and classification tasks to provide worthy support to dental professionals in the aim of improving patient care.

The studies referenced provide a comprehensive overview of the advancements in the use of deep learning for the detection and classification of lesions in dental panoramic radiography. Each study gives unique insight into methodologies that help to achieve this purpose. They have shown diversity of deep learning methodologies, from YOLO to Faster R-CNN and U-Net architectures, and their efficacy in various aspects of dental radiography. The ability of these models to achieve high accuracy, sensitivity, and specificity in detecting and classifying dental lesions, is reached by taking advantage of the proposed ideas. This project aims to create an automatic System for Detection and Classification of Lesions in Dental Panoramic Radiography using Deep Learning.

Chapter 4: Practical Aspect

8- Tufts Dental Database:

After extensive search through common database the Tufts Dental Database was chosen for this project.

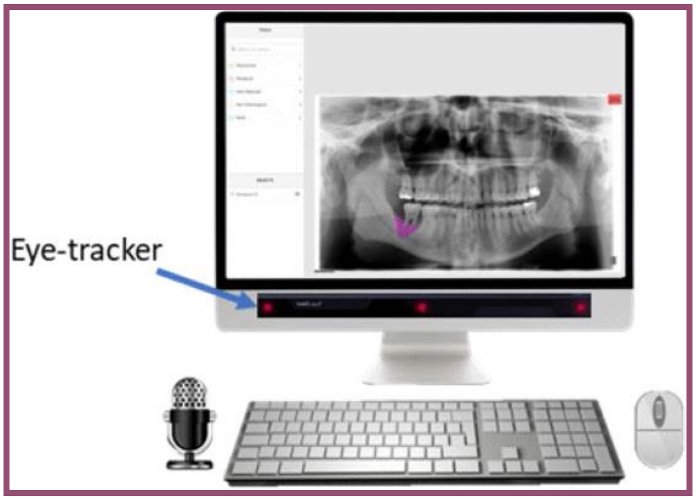
8-1 Introduction to Tufts Dental Database:

TDD is a new X-ray panoramic radiography image dataset. It consists of 1000 panoramic radiographs with the abnormalities, teeth, and the maxillomandibular region of interest outlined.

These panoramic radiography images were for Tufts University School of Dental Medicine patients, that were randomly selected from the electronic patient database (axiUm) in the period from January 1, 2014, to December 31, 2016.

Images were selected according to special inclusion criteria: optimum diagnostic quality of the image with minimal or no technical errors in the image. The selected radiographs were de-identified and saved in a generic image format (TIFF/JPEG) with a unique identifier. Each of these radiographs were annotated by an expert and a student from the Tufts University School of Dental Medicine.

This database is considered as multimodal dataset because it also includes the expert and student expertise captured in the form of eye-tracking and think-aloud protocol. gaze maps are presented the eye-tracking while oral dictation is provided as text within .json file. The multimodal nature of the dataset allows for a much more comprehensive range of applications. Fig. 2 displays the acquisition system setup used for this project [58].



Figure(35): The acquisition system setup. The system consists of an eye-tracker and an audio recording device [58].

8-2 Tools for Collecting Database:

Labelbox was used to annotate and label each radiograph by expert and student. Tobii Eye Tracker 4C captured the eye movements as x and y coordinates while the audio was recorded using microphone and synchronized with the eye tracker. These audio files acquired from the expert and the student describing each radiograph were converted to text using Google and Dragon automatic speech recognition platform. The text generated for each radiograph is added to the .json file.

All panoramic images were obtained using OP100 Orthopantomograph (Instrumentarium Imaging/Kavo Kerr) and Plammeca Promax 2D radiographic units. Automatic exposure control was utilized to obtain clinically diagnostic images with optimal image density and contrast [58].

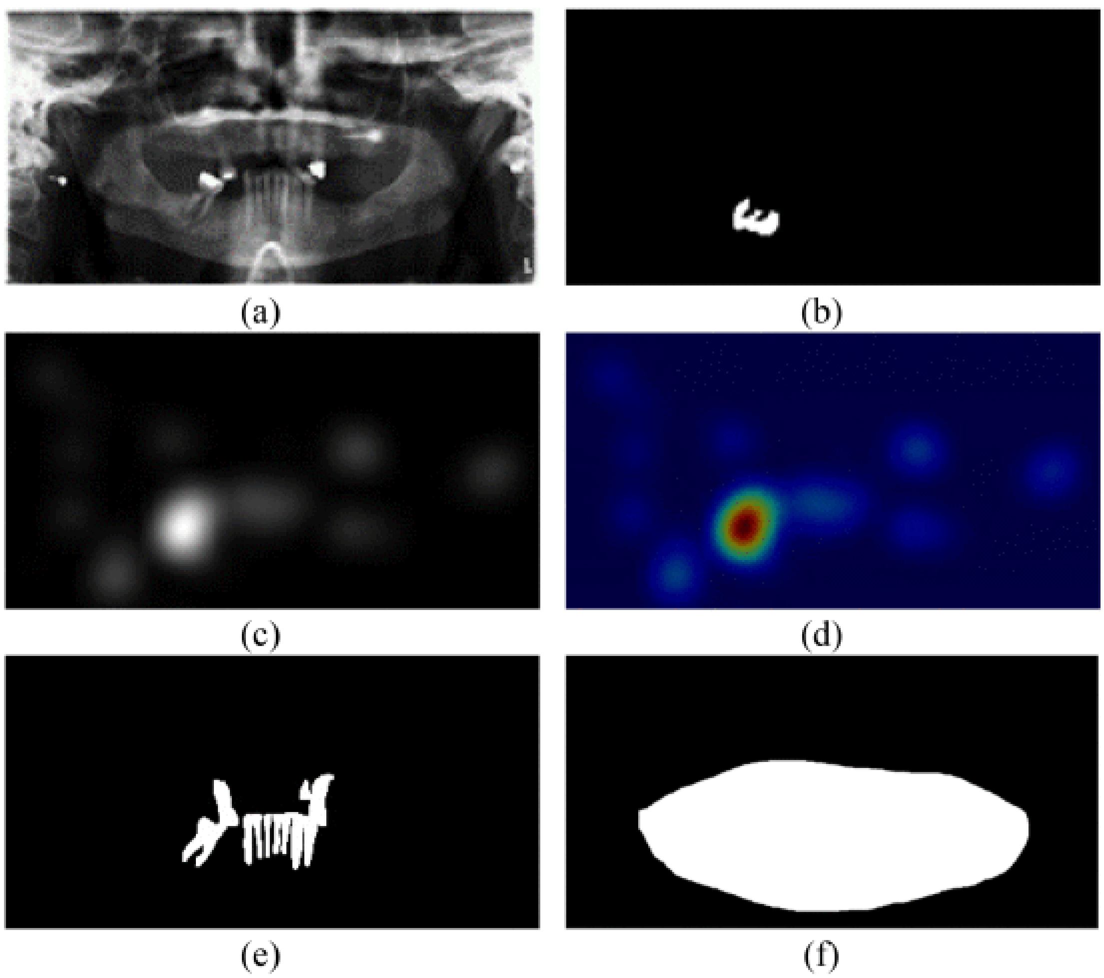
8-3 The Components of Tufts Dental Database:

The dataset consists of 6 major components:

1) radiographs.

2) labeled masks (segmented mask outlining the abnormality).
3) eye tracker generated maps (gray and quantized).
4) text information describing each radiograph.
5) teeth mask for each radiograph with labels.
6) maxillomandibular region-of-interest mask.

Figure (36) shows an example of each imaging component [58].



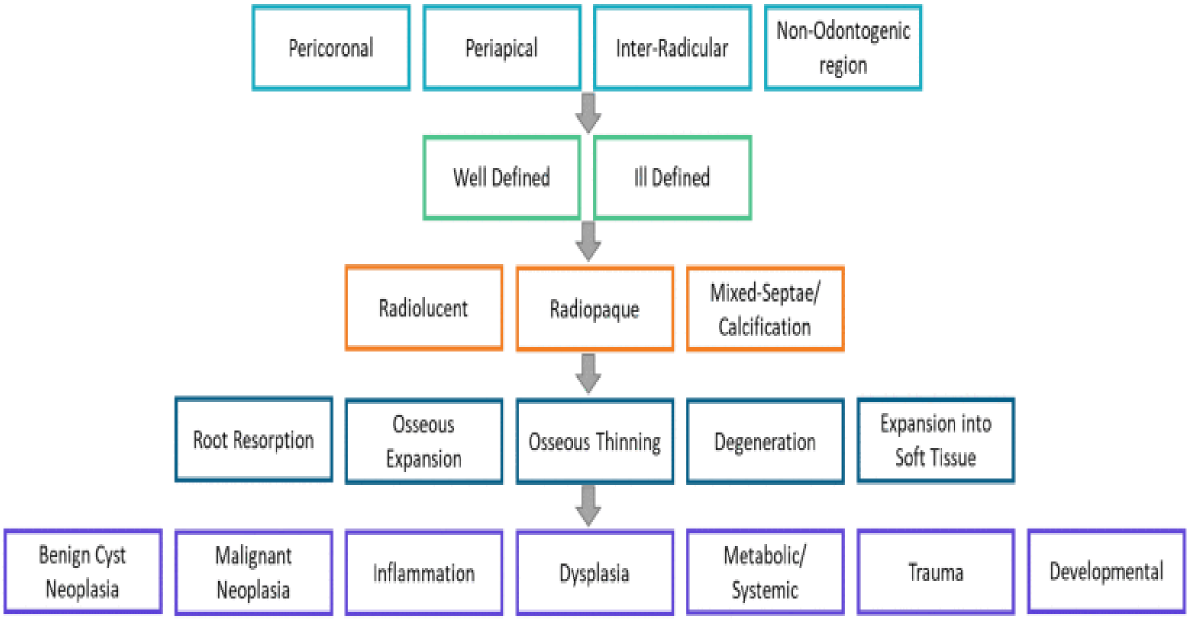
Figure(36): Illustration of the different imaging components. (a) panoramic radiograph, (b) segmented mask outlining the abnormality, (c) grayscale eye-tracking gaze plot, (d) color quantized eye-tracking gaze plot (e) teeth mask, (f) maxillomandibular region of interst [58].

The expert and student utilized an approach to analyze each radiograph, first identify the abnormality, then describe the radiographic characteristics of the abnormality, which allowed for categorization and differential diagnosis of the entity. Then, the classification of radiography images was performed based on five different levels:

1) The first level described the anatomic location relative to the jaw and teeth.
2) The second level described the periphery or marginal characteristics of the abnormality (well-defined vs ill-defined).
3) The internal architecture was the third level and focused on the radiodensity- Radiolucent, Radiopaque, and Mixed with Septae or Calcifications.
4) The fourth level focused on the effects of the abnormality on adjacent structures and was recorded in terms of tooth displacement, root resorption, osseous thinning and expansion, and extension into the adjacent soft tissue or degenerative changes.
5) Based on the overall descriptions, a category (fifth level) for the abnormality was selected from Trauma, Inflammation, Dysplasia, Developmental, Benign tumor or Cyst, Malignant Neoplasia, Systemic or Metabolic conditions.

The five-level classification of each radiograph is recorded in the .json files for the expert and the student. In conclusion the structure of the .json file containing the label and text information for each radiograph.

Figure (37) shows the five-level classification of each radiograph and figure (38) shows the structure of the .json file [58].



Figure(37): Logical sequence followed while evaluating each radiograph. It depicts the five different characteristics recorded for a radiograph with abnormalities [58].


```
"Label": {
    "objects": [
        {
            "title": "None",
            "value": "none",
            "classifications": "none" ,
            "polygons": "none"
        }
        ],
        "classifications": []
},
"Description": [Within normal limits]
"External ID": "570.JPG"
},
```


Figure(38): Structure of the .json file. The file contains information on abnormalities pertaining to each radiograph [58].

8-4 Performance Review:

The dataset was evaluated with modern image enhancement and segmentation techniques.

8-4-1 Image Enhancement:

In diagnostic radiography, the primary aim of image enhancement is to create visually appealing images. This is achieved by increasing contrast, optimizing brightness, improving sharpness, and reducing noise. To evaluate the effectiveness of different methods typically used for medical image enhancement, images from a dataset were tested. Each image underwent enhancement, and their quality scores were evaluated. Among the various algorithms, the Contrast Limited Adaptive Histogram Equalization CLAHE algorithm demonstrated the best performance in enhancing image contrast. Additionally, a visual comparison of the results revealed that the Guided Filtering GF algorithm excelled in enhancing edges.

8-4-2 Segmentation of teeth from radiographs:

They evaluated the performance of different deep learning-based algorithms for segmenting teeth from panoramic radiographs. The data consisted of the 1000 radiographs and ground truth masks. The results showed that the performance of UNet and UNet++ was the best [58].

8-5 Conclusion:

Tufts Dental Database, a new dataset with the potential to revolutionize the dental AI practice. TDD consists of a tooth segmentation mask, abnormality mask, maxillomandibular region of interest mask, eye-tracking gaze map, and text description of the abnormality. it provided a detailed evaluation of dental radiographic image enhancements and tooth segmentation. Furthermore, TDD opens the opportunity to integrate the radiologist's expertise into the AI system, making it more efficient and reliable. This dataset also paves the way for a new generation of AI systems that help dentists saving their effort and time. After understanding the database structure and its contents, radiograph folder, expert's mask folder and expert's json file were chosen for this work [58].

9- Importance of Python:

Choosing the right programming language very important and can make all the difference in the application of Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL).

Python stands out as the best language for AI applications that make it popular and widely used. According to a Stack Overflow Developer Survey in 2023, Python is among the most widely used software programming languages worldwide [59].

There are many reasons make Python the preferred choice for AI applications:

1) **Simplicity and readability** are the key reasons coders use Python for AI, ML, and DL. Python is designed to be easy to understand and write. It helps developers concentrate on the problem-solving aspects in these field.
2) Python allows you to run the script on GPU that can be comparatively faster than CPU. GPUs are preferred because most AI applications require parallel processing of multiple calculations, for example, deep learning operations with massive parallel inputs of data.
3) **Libraries:** Another advantage is the diversity of Python standard libraries that cover a lot of aspects, reducing the need to code everything

from scratch. Libraries such as NumPy, Pandas, Tensorflow, Pytorch and Matplotlib take care of the numerical aspects and data visualization, while SciPy brings in additional scientific computing capabilities.

4) **Interoperability**: is another advantage of Python. It leads to easy communication with other languages like C and C++, enabling it to leverage optimized code pieces for computationally intensive tasks. This results in a better performance. Moreover, Python's extensive community support ensures that developers can quickly find solutions and get help when needed. [60]

All these characteristics allows it to adapt with various AI, ML, and DL applications. Whether it's data preprocessing, developing complicated algorithms, or creating sophisticated neural networks, Python effortlessly caters to the diverse needs of these domains [59]. Because of all the above, I used Python as programming language for this project.

10- Implementation and Workflow:

10-1 Labels Extraction from JSON file:

10-1-1 JSON File Definition:

JSON stands for JavaScript Object Notation; it is a text format to store and transport simple data structures and objects. JSON is used to send data between computers, primarily between a web application and a server. JSON files are characterized by lightweight, text-based, human-readable, and can be edited using a text editor [61].

10-1-2 The Structure of JSON File:

It stores data in key-value pairs and arrays, in away make it readable where the keys serving as names and the values containing related data [62].

JSON syntax is derived from JavaScript object notation syntax:

1) Data is in key/value pairs.
2) Data is separated by commas.
3) Curly braces hold objects.
4) Square brackets hold arrays.

The JSON syntax is restricted, Keys must be strings written with double quotes, and values must be a valid JSON data type string, number, object, array, Boolean, or null [62].

I worked on the expert's json file, where I studied its nested structure to know how to deal with this type of files, with the aim of extracting the labels (lesions associated with panoramic radiography image), knowing their number and types, and collecting this data in a CSV file to be use in the next stages. Within this step, I found that there was some panoramic radiography image duplicated in the CSV file. Some of these duplicates were related with different labels, and others with the same label. So it was necessary to find the reason of these duplicates and find out whether they were caused by the presence of several areas in one image that carry different labels, or the labels are belong to one area and compared this with the areas covered by the mask, associated that with the oral dictation recorded by the expert as a text in the same json file in order to understand the condition of each image and the method based on which the annotation was made and thus determine how to deal with these panoramic radiography image in the best way. For collecting labels json and csv libraries was used, and the result was labels_extraction.csv that contains images ID and labels.

10-2 Importing Libraries:

Essential libraries for handling images (PIL, OpenCV), data manipulation (pandas, numpy), deep learning (PyTorch, torchvision), evaluation metrics (sklearn), and visualization (matplotlib) are imported to provide the necessary functionalities for the project.

10-2-1 Python Imaging Library (PIL):

It is a free and open-source Python Imaging Library that is used to opening, manipulating, and saving many different images file formats. This image library is designed for fast access to data stored in a few basic pixel formats. It is considered as a general image processing tool that provides solid foundation [63-64].

10-2-2 OpenCV:

(Open-Source Computer Vision Library) is a huge open-source library for computer vision, machine learning, and image processing. It is used for process images and videos to build applications in fields like robotics, facial recognition, object detection, and much more [65].

10-2-3 Pandas:

Pandas is a powerful and widely used open-source data analysis and manipulation library in Python, designed to make working with structured data both easy and efficiently. It provides two primary data structures: Series and DataFrame. These data structures allow for a variety of operations such as data cleaning, transformation, and aggregation [66].

10-2-4 NumPy (Numerical Python):

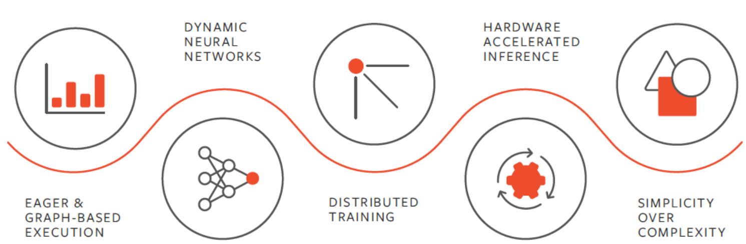
It is an open-source Python library that provides support for large, multidimensional arrays and matrices with fast operations including mathematical, logical, shape manipulation, sorting, selecting and much more [67].

10-2-5 PyTorch:

PyTorch is an open-source deep learning framework Written in Python. It is widely used for building and training neural networks and has gained popularity due to its ease of use, flexibility, and strong support for dynamic computation graphs, which allow for more intuitive model building and debugging. it is commonly used in applications like image recognition and language processing.

The two main features of PyTorch are:

- Tensor Computation with strong GPU (Graphical Processing Unit) acceleration support.
- Automatic Differentiation for creating and training deep neural networks.



Figure(39): Features of PyTorch [68].

PyTorch is currently the preferred library for AI (Artificial Intelligence) researchers and practitioners around the globe, both in industry and academia. It is favoured over other Deep Learning frameworks like TensorFlow and Keras since it uses dynamic computation graphs and is completely Pythonic. It allows scientists, developers, and neural network debuggers to run and test parts of the code in real-time. For that, users don't have to wait for the entire code to be implemented to check if a part of the code works or not [68-69].

Table (2): Comparison of PyTorch, TensorFlow, and Keras [70].

Criteria	PyTorch	TensorFlow	Keras
Ease of Use	Intuitive, flexible, Pythonic	Complex API, improved with eager execution	Simple, user-friendly, high-level API
Flexibility	Dynamic computational graph	Originally static, now supports dynamic	High-level, integrated with TensorFlow
Performance	Strong for dynamic neural networks, GPU support	Highly optimized for large-scale production	Benefits from TensorFlow optimizations
Community	Rapidly growing, strong in research	Large, comprehensive ecosystem	Large, benefits from TensorFlow's community
Deployment	Tools like TorchServe, ONNX support	Robust tools (TFX, TF Lite, TF Serving)	Simplified deployment via TensorFlow
Best For	Research, rapid prototyping	Production environments, large-scale deployment	Beginners, quick prototyping, ease of use

10-2-6 Torchvision:

It is a library that is part of the PyTorch ecosystem. It provides tools specifically tailored for computer vision tasks. Torchvision includes popular datasets, model architectures, and image transformations, which streamline the process of developing and testing computer vision models [71].

10-2-7 scikit-learn (also known as sklearn):

It is a free and open-source machine learning library written with Python. It contains various classification, regression and clustering algorithms [72].

10-2-8 Matplotlib:

It is a powerful plotting library for the Python programming language and its numerical mathematics extension NumPy. it used for creating static, animated, and interactive visualizations. it also can add plots into applications using general-purpose GUI to represent data graphically which make it easier to analyze and understand [73].

10-3 Bounding Box Extraction:

A function 'extract_boxes' processes the masks of images to extract bounding boxes using morphological operations and contour detection. This function is very important for identifying regions of interest in the mask images, which are later used as targets for the object detection model.

10-4 Dataset loading and preprocessing:

10-4-1 Custom Dataset Class:

Loading datasets efficiently is a critical step in any deep learning application. PyTorch provides powerful tools to facilitate loading custom datasets through its torch.utils.data.Dataset and torch.utils.data.DataLoader. PyTorch handled data loading through two classes:

1) **Dataset:** A base class for representing all datasets. The custom dataset should inherit Dataset and implement three functions:

__init__	function is run once when instantiating the Dataset object.
__len__	Returns the size of the dataset.
__getitem__	Retrieves a sample and its corresponding label at a given index.

2) **DataLoader:** Combines a dataset and a sampler, providing an iterable over the dataset to enable easy access to the samples, supporting batching, shuffling, and parallel data loading. [74-75].

A custom dataset class ‘DentalXrayDataset’ is defined, inheriting from torch.utils.data.Dataset. This class handles loading images and masks, applying transformations, and preparing target data for the model. It implements methods for getting the length of the dataset and fetching individual data items, including image-mask pairs and their corresponding labels. This process depending on csv file which contain images ID and their corresponding labels.

Table (3): The Input and Output of DentalXrayDataset.	
Inputs	Images, masks and CSV file contains labels.
Outputs	Images with Targets contains bounding box and labels.

10-4-2 Data Transformations:

Data Transformations: is an essential step in any deep learning pipeline. it involves preprocessing steps applied to each sample in the dataset and change it to a format suitable for training a model. This includes many operations provided by PyTorch which has a powerful library called torchvision.transforms that makes these operations easy and efficient.

In this work Several transformation classes are defined to handle image and mask preprocessing. These transformations are composed using the "ComposeWithMask" class to make sure that the same transformations are applied consistently to both images and masks [76].

Most transformations accept both PIL images and tensor images, although some transformations are PIL-only and some are tensor-only. The Conversion may be used to convert to and from PIL images, or for converting dtypes and ranges.

The transformations that accept tensor images also accept batches of tensor images. A Tensor Image is a tensor with (C, H, W) shape, where C is a number of channels, H and W are image height and width. A batch of Tensor Images is a tensor of (B, C, H, W) shape, where B is a number of images in the batch [77].

The following is the list of transformations applied:

1. **ToTensorWithMask:**

- The Role: Converts images and masks from PIL format or numpy arrays to PyTorch tensors.
- Preprocessing Function: This is a main preprocessing step that prepares the data in a tensor format that is compatible with PyTorch where tensors are the primary data structure used in PyTorch for modeling.

2. **NormalizeWithMask:**
   - The Role: Standardizes the pixel values of images and masks to have a specific mean and standard deviation.
   - Preprocessing Function: Normalization is an important preprocessing step that helps in stabilizing and accelerating the training process. By making sure that the input data has a uniform distribution it allows the model to learn more effectively.
3. **RandomHorizontalFlipWithMask:**
   - The Role: Applies horizontal flipping to images and masks with a certain probability for randomness.
   - Preprocessing function: This step improves the training data set by adding variations, making the model stable in horizontal directions. It is a form of data augmentation and is an integral part of the preprocessing pipeline to improve generalization.
4. **RandomRotationWithMask:**
   - The Role: Rotates images and masks by a random angle within a specified range.
   - Preprocessing Function: Like the flipping process, rotation introduces variations in the dataset, which helps the model learn rotational

invariance. It’s another augmentation technique that simulates different possible orientations of dental X-rays.

5. **ColorJitterWithMask:**
   - The Role: Randomly adjusts the brightness, contrast, saturation, and hue of images while leaving masks unchanged since they are in binary.
   - Preprocessing Function: This transformation simulates different lighting conditions and enhances the robustness of the model to variations in image acquisition settings. It ensures the model can handle variations in real-world scenarios.
6. **RandomNoiseWithMask**:
   - The Role: Adds random noise to images while leaving masks unchanged (binary masks)
   - Preprocessing Function: Adding noise is a preprocessing step that makes the model robust to artifacts in the images. It helps the model to generalize better by training on slightly noisy data.
7. **Composite Transformation (ComposeWithMask)**
   - The Role: Chains multiple transformations into a single composite transformation.
   - Preprocessing Function: This composite transformation ensures that all individual preprocessing steps are applied in sequence to each image-mask pair. It streamlines the preprocessing pipeline, making it more efficient and consistent.

10-4-3 Data Augmentation:

Data augmentation is very important in healthcare applications especially in medical imaging because it helps improve diagnostic models that detect, recognize, and diagnose diseases based on images. the main reason for augmented images is making more training data for models especially for application that lack source data variations.

Data augmentation is the artificial process for generating new data from existing data. it is primarily used to train machine learning (ML) and deep learning models because these models require large and varied datasets, but sourcing sufficiently diverse real-world datasets can be challenging because of data silos, regulations, and other limitations. Augmented images are made by making small changes to the original data [78].

The benefits of data augmentation:

1- Enhanced model performance:

Data augmentation techniques enrich datasets by generating numerous variations of the existing data. This results in a larger dataset for training, allowing the model to encounter a diverse range of features. Consequently, the augmented data enables the model to generalize better to new, unseen data, improving its overall performance in real-world environments.

2- Reduced data dependency:

The collection and preparation of large data is costly and time-consuming. Data augmentation techniques make small data sets useful and effective by supplementing the main set with artificial data point.

3- Mitigate overfitting in training data:

Data augmentation helps prevent overfitting, that is unfavorable when models accurately provide predictions for training data, but it struggles with new data. This happened with small and narrow dataset where the model overfit and give predictions related to specific data type, for that using data augmentation is important larger and more comprehensive dataset for model training. it makes training sets unique to deep neural networks, preventing them from overfitting to only specific characteristics.

4- Improved data privacy:

Augmentation techniques help in training deep learning models on sensitive data by using existing data to create synthetic data. This augmented data keeps the input data's statistical properties and weights while protecting and limiting access to the original [78].

For this work, TDD contains 1000 images which considered relatively small. it is also having normal cases more than abnormal ones for this reason, data augmentation techniques used to expand the dataset and combining it with additional data sources to improve model performance and reduce overfitting.

The function "custom_augmentation" is defined to apply random data transformations to images and masks. This adds variability to the training data and helps the model generalize better by simulating different viewing angles and positions of the dental structures.

- The Role: Applies random data transformations.

- Preprocessing Function: This step is integrated into the preprocessing pipeline to simulate diverse real-world conditions, aiding in the creation of a robust model.

10-5 Dataset Initialization and Splitting:

The dataset is instantiated and split into training, validation, and test sets using an 70-15-15 ratio. Data loaders are created for batching and shuffling the datasets, facilitating efficient and randomized access to the data during training and evaluation.

10-6 Model Definition:

10-6-1 Overview of the Object Detection Pipeline:

Traditional object detection techniques consist of 3 major steps given in the figure (40):

1. The first step in the process involves generating numerous region proposals, which are potential candidates for containing objects. Typically, the number of these regions is in the thousands, often around 2,000 or more. Algorithms like Selective Search and EdgeBoxes are commonly used for generating these region proposals.
2. For each region proposal, a fixed-length feature vector is extracted using various image descriptors, such as the histogram of oriented gradients (HOG). This feature vector is crucial for the success of object detectors, as it needs to effectively describe an object even when it undergoes transformations like scaling or translation.
3. The extracted feature vector is then used to classify each region proposal as either a background class or one of the object classes. As the number

of object classes increases, the complexity of building a model that can accurately distinguish between all these objects also increases [79].

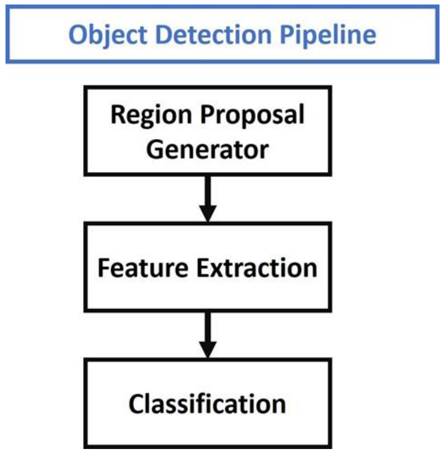


Figure (40): Object Detection Pipeline [79].

10-6-2 Faster R-CNN:

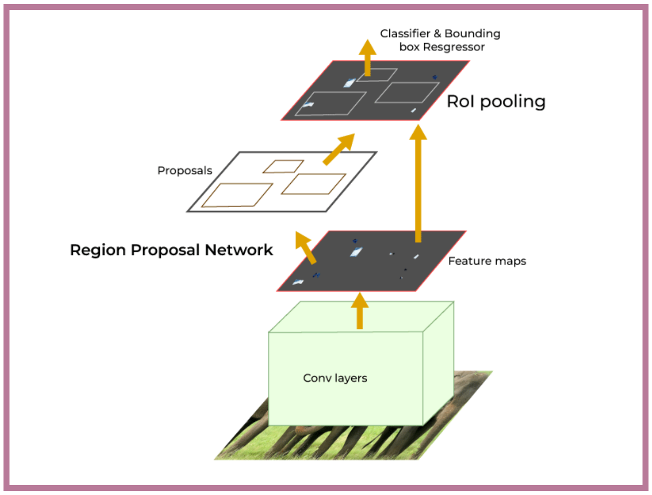
Identify the objects and their location inside images or video streams is one of the key tasks in computer vision, with the development of deep learning significant growth was made in this field. The R-CNN family is a popular example about algorithms have used for object detection.

Faster Region-Convolutional Neural Network or Faster RCNN is a modern object detection architecture of the R-CNN family, introduced by Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun in 2015. The primary goal of this algorithm is to detect objects within an image and find the location of them accurately. It combines the advantages of deep learning, convolutional neural networks (CNNs), and region proposal networks (RPNs) into one network, which greatly improves the speed and accuracy of the model.

Faster R-CNN architecture consists of 3 components:

1. Convolutional Neural Network (CNN) Backbone.
2. Region Proposal Network (RPN).

3. Fast R-CNN detector [80].



Figure(41): Faster R-CNN architecture [80].

10-6-2-1 Convolutional Neural Network (CNN) Backbone:

the starting layers of Faster R-CNN architecture is Convolutional Neural Network (CNN) Backbone that work on extract the relevant features from the input image. It consists of multiple convolutions layers that apply different convolutions kernel to extract the features. the kernels capture the hierarchical representations of the input image means he first layers of CNN captures the low-level features likes edges and tectures, and while deeper layers capture the high level semantic features like objects parts and shapes. the extracted features are used by both RPN and Fast R-CNN detector which redact computing time and memory use [80].

10-6-2-2 Region Proposal Network (RPN):

It is the main part of Faster R-CNN that generates possible regions of interest (region proposals) in images that may contain objects. it uses the concept of attention mechanism in neural networks that instruct the subsequent Fast R-CNN detector where to look for objects in the image. Previously, generating region proposals with traditional Selective Search algorithm took more time in computations. Faster R-CNN fixes these issues by introducing a convolutional-based network RPN, which reduces proposal time for each image from 2 seconds to 10 ms and improves feature representation by sharing layers with detection stages. The key components of the Region Proposal Network are as follows:

1. **Anchors boxes:**

   Faster R-CNN model uses a set of predefined anchor boxes to generate region proposals. These anchor boxes are placed at different positions on the feature maps. each anchor box has 2 key parameters scale and aspect ratio.

2. **Sliding Window approach:**

   RPN slides over the feature map obtained from the CNN backbone (sliding window mechanism). It uses a small convolutional network (typically a 3×3 convolutional layer) to process the features within the receptive field of the sliding window. This produces scores that indicate the probability of an object's presence and regression values for adjusting the anchor boxes.

3. **Objectness Score:**

   It represents that a given anchor box contains an object of interest (positive) rather than being just background (negative). this value is given for each anchor. The objectness score reflects the confidence that the anchor corresponds to a meaningful object region.

4. **IoU (Intersection over Union):**

It is a metric used to measure the degree of overlap between two bounding boxes. It calculates the ratio of the area of overlap between the two boxes to the area of their union. Mathematically, it is represented as:

$$IoU = \frac{Area\ of\ Intersection}{Area\ of\ Union}$$

5. **Non-Maximum Suppression (NMS):**

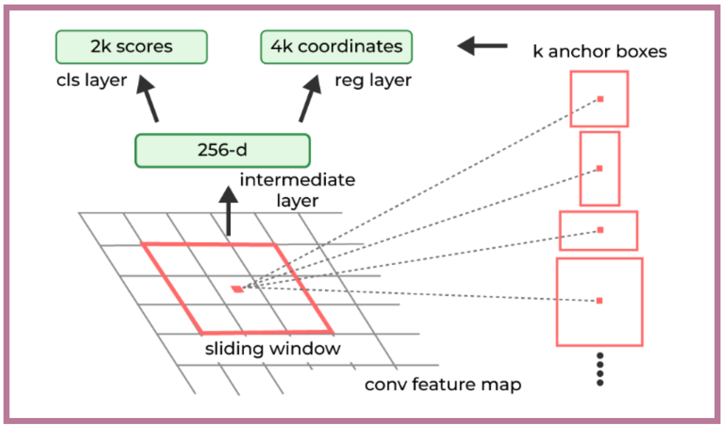
It is used to remove redundancy and select the most accurate proposals, based on the objectness scores of overlapping proposals and keeps only the proposal with the highest score while suppressing the others.

On feature maps obtained from the CNN backbone a sliding window approach with anchor boxes of different scales and shapes uses to detect potential object positions. The network refines these anchor boxes throughout training to better match actual object positions and sizes. the RPN predicts two parameters for each anchor:

- The probability of the anchor containing an object (objectness Score)
- Adjustments to the coordinates of anchor that match the actual object's shape

This operation generates a large number of region proposals, many of them may overlap and related to the same object. Here the Non-Maximum

Suppression (NMS) work on ranks the anchor boxes according to their objectness probabilities and selects the top-N anchor boxes with the highest scores where NMS make sure that the final selected proposals are both accurate and non-overlapping. As a result of all these steps the selected anchor boxes are considered as possible region proposals [80].

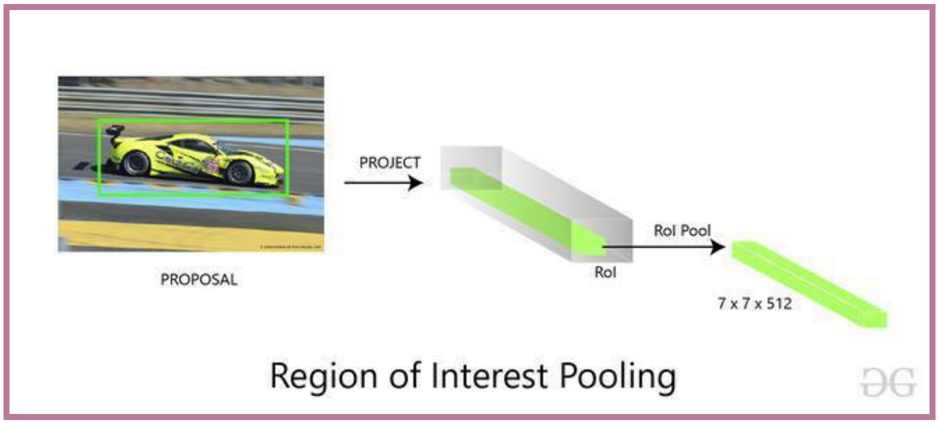


Figure(42): *Region Proposal Network* [80].

10-6-2-3 Fast R-CNN detector:

The Fast R-CNN detector is essential component of the Faster R-CNN architecture, responsible for object detection within the region proposals suggested by the Region Proposal Network. The workflow of Fast R-CNN detector:

1- **Region of Interest (RoI) Pooling:** The first step is applying RoI pooling on region proposals suggested by the RPN. This transforms the region proposals with variable size into fixed-size feature maps that fed into the network's subsequent layers. each region proposal is divided into a grid of equal-sized cells by RoI pooling then max pooling apply within each cell. This process gives fixed-size feature map for each region proposal, which can be further processed by the network.



Figure(43): Region of interest pooling [80].

2- **Feature Extraction:** the fixed-size feature maps are fed into the CNN backbone to extract meaningful hierarchical features from region proposals that capture object-specific information. These features retain spatial information while abstracting away low-level details, this allows the network understanding the content of proposed regions.

3- **Fully Connected Layers:** The RoI-pooled and feature-extracted regions then pass through a series of fully connected layers. These layers are responsible for object classification and bounding box regression tasks.

- **Object Classification:**

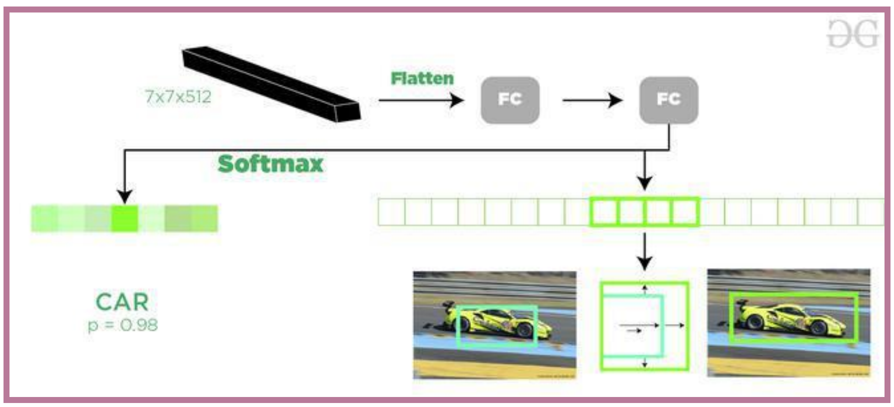
  For each region proposal, the network predicts class probabilities that point out the proposal contains an object of a specific class.

combining the features retrieved from the region proposal with the shared weights of the CNN backbone, the classification is made.

- **Bounding Box Regression:**

 The network also predicts bounding box adjustments for each region proposal that refine the position and size of the bounding box to make it more accurate and aligned with the actual object boundaries.

1. the first layer is a softmax layer with N+1 output parameters where N represent the number of class labels and background. this layer predicts objects in region proposal.
2. the second layer is a bounding box regression layer that has 4* N output parameters. it regresses the location of bounding box location that contain the object in the image.



Figure(44): fully connected layer [80].

4- Multi-task Loss Function: Fast R-CNN detector combines classification and regression losses. The classification loss computes the difference between expected and true class probabilities. The regression loss computes the difference between expected and actual bounding box adjustments.

$$L(p_i, t_i, v_i) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, v_i)$$

Where:

N_{CLS}: is the number of RoIs used for classification.

N_{reg}: is the number of RoIs used for bounding box regression.

p_i: is the predicted probability of classifying the (i) − th RoI.

p_i^*: is the ground truth indicator (1 or 0) for the (i) − th RoI being a foreground or background object.

t_i: represents the ground-truth bounding box parameters for the (i) −th RoI

v_i: represents the ground-truth bounding box parameters for the (i) − th RoI

L_{cls}: is the classification loss function, often computed using cross-entropy loss.

L_{reg}: is the regression loss function, often computed using smooth L1 loss.

λ : is a balancing parameter that controls the trade-off between the two components of the loss.

5- **Post-Processing:** this procedure refines the final detection results (class probabilities and bounding box changes) by non-maximum suppression (NMS) which used to reduce redundant detections while retaining the most confident and non-overlapping detections [80].

For this project's model:

- A provided CNN architecture is a simple Convolutional Neural Network (CNN) designed for binary classification of dental X-ray images into normal and abnormal cases. The architecture consists of two convolutional layers followed by max pooling, flattening, and two fully connected layers. Each convolutional layer applies a set of filters to extract features from the images, and the pooling layers downsample the feature maps to reduce the spatial dimensions. The fully connected layers then map the extracted features to the final class scores.

Table (4): CNN Architecture.

Layer	Parameters Description
conv1	3 input channels, 32 output channels, 3x3 kernel size, padding of 1.
conv2	32 input channels, 64 output channels, 3x3 kernel size, padding of 1.
Pooling Layer	2x2 kernel size, stride of 2.
Fully Connected Layers.	
fc1	Input Features: 64 * 64 * 64 (after two rounds of convolution and pooling). Output Features: 512.
fc2	Input Features: 512 (output from fc1). Output Features: 2 (binary classification: normal vs abnormal).

- Then the function 'get_model_instance_segmentation' loads a pre-trained Faster R-CNN model and changes it to detect the regions of abnormalities in the abnormal images which are the output of CNN architecture. This involves replacing the last layer with a custom layer designed for this specific case.

10-7 Training and Optimization:

10-7-1 Faster RCNN Training:

Faster R-CNN is trained by optimizing the entire network, including the Region Proposal Network (RPN) and shared convolutional layers, using backpropagation and stochastic gradient descent (SGD) to minimize the loss function which combines classification loss, which measures how well the network distinguishes objects from the background, and regression loss, which measures the accuracy of predicted bounding box coordinates. Each mini-batch is created from a single image, containing 256 randomly selected anchors (potential bounding boxes) balanced with up to 128 positive (object-containing) and 128 negative (background) examples. If there are fewer than 128 positive samples, negative samples are added to maintain balance. New RPN layers are initialized with weights drawn from a Gaussian distribution, while shared convolutional layers use weights pretrained on ImageNet. The learning rate starts at 0.001 for the first 60,000 mini-batches and then decreases to 0.0001 for the next 20,000, with a momentum of 0.9 and a weight decay of 0.0005 to prevent overfitting. This approach ensures the network is trained efficiently and accurately [80].

10-7-2 Optimizer:

The optimizer is a main element that fine-tunes a neural network's parameters during training to minimize the model's error or loss function, enhancing performance. These algorithms facilitate the learning process of neural networks by iteratively refining the weights and biases based on the feedback received from the data [81].

10-7-2-1 Stochastic Gradient Descent Optimizer:

SGD is a useful method for optimizing deep learning models, especially with large datasets. Unlike traditional gradient descent, which processes the entire dataset at once, SGD uses randomness by selecting random batches of data for each iteration. It chooses initial parameters (w) and a learning rate (n). Then, Shuffle the data randomly at each iteration.

$$w \coloneqq w - \eta \nabla Q_i(w)$$

Because SGD uses only small batches of data, the optimization path is noisier compared to gradient descent, leading to more iterations to find the minimum. However, even with more iterations, SGD is still faster and less computationally expensive than processing the entire dataset at once. In summary, for large datasets and when computation time matters, SGD is a better [81].

10-7-2-2 Adam Optimizer:

The Adam optimizer (short for Adaptive Moment Estimation optimizer) is a popular algorithm in deep learning, used to update the weights of a neural network during training. It is an advanced version of the stochastic gradient descent (SGD) algorithm. It adjusts the learning rate for each weight individually, unlike SGD which uses the same learning rate for all weights. Adam calculates learning rates based on past gradients and their squared values, helping the neural network train faster and perform better. This algorithm is straightforward to implement, has a faster running time, low memory requirements, and requires less tuning than any other optimization algorithm.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \left[\frac{\delta L}{\delta w_t}\right] v_t = \beta_2 v_{t-1} + (1 - \beta_2) \left[\frac{\delta L}{\delta w_t}\right]^2$$

The above formula represents the working of adam optimizer. Here B1 and B2 represent the decay rate of the average of the gradients [81].

For this work:

- An Adam optimizer is used to adjust the model's weights during training. The 'ReduceLROnPlateau' scheduler lowers the learning rate when validation loss stops improving, making training more efficient.
- The EarlyStopping class monitors the validation loss and stops training if there is no improvement over a set number of epochs. This prevents overfitting by stopping the training once the model stops getting better on the validation set.
- Label Distribution Check is functions are created to check and print the number of normal and abnormal images in the training and testing sub-dataset. This ensures that the dataset is balanced and contains both classes for training and testing.

- Model Training and Validation

  The model is trained for number of epochs. Each epoch involves training on batches of images, calculating loss, and updating the model's parameters. Validation is performed at the end of each epoch to monitor the model's performance. Early stopping is used here to stop training if the validation loss does not improve which prevents overfitting.

10-8 Testing and Statistical Analysis:

The model is tested on a separate dataset. Predictions are made, and several metrics are calculated. A confusion matrix is used to summarize the model's performance, showing how well it distinguishes between normal and abnormal cases.

10-8-1 Visualizing Predictions:

The "display_prediction" function is used to visualize the actual and predicted bounding boxes for abnormal cases. This helps in assessing the model's performance by comparing predictions with actual annotations visually.

10-8-2 Metrics calculations:

1) Accuracy:

It is an essential metric for evaluating the performance of a classification model, providing a quick information of how well the model is performing in terms of correct predictions. This is calculated as the ratio of correct predictions (both true positives and true negatives) to the total number of input instances.

$$Accuracy = \frac{No.\,of\,correct\ predictions}{Total\ number\ of\ input\ samples}$$

- **The Importance:** It gives an overall measure of how often the model is correct. However, in imbalanced datasets, accuracy alone can be misleading as it may hide the poor performance of the minority class [81].

2) Precision:

The precision is the ratio between the number of Positive samples correctly classified to the total number of samples classified as Positive (either correctly or incorrectly). Precision measures the accuracy of model in classifying a sample as positive.

$$\text{Precision} = \frac{True_{positive}}{True_{\ positive} + False_{positive}}$$

- **The Importance:** Precision shows the ability of the model to correctly identify positive cases. High precision means that the model has a low false positive rate, which is important in medical images to avoid false alarms [81].

3) Recall (Sensitivity):

The recall is the ratio between the number of Positive samples correctly classified to the total number of Positive samples. The recall measures the model's ability to detect Positive samples. The higher the recall, the more positive samples detected.

$$\text{Recall} = \frac{True_{positive}}{True_{positive} + False_{negative}}$$

- **The Importance:** Recall measures the ability of the model to identify all relevant instances of the positive class. High recall is critical in medical images to ensure that no positive cases (e.g., abnormal conditions) are missed [82].

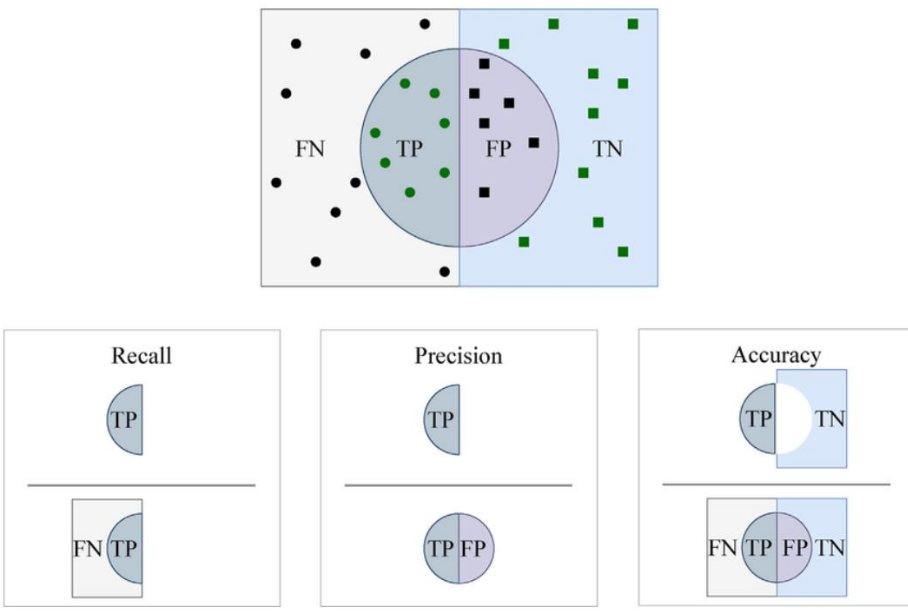


Figure (45): Visualizing accuracy, recall, and precision, which are the common performance measures for classification tasks. Given samples from two categories [83].

4) F1 Score:

It is calculated as a harmonic mean of recall and precision. Its range is [0,1]. This metric usually tells us how precise (It correctly classifies how many instances) and robust (does not miss any significant number of instances), meaning it provides a single metric to evaluate performance.

$$F_1 = 2 \times \frac{Precision \, . \, Recall}{Precision + Recall}$$

$$F_1 = \frac{True_{positive}}{True_{positive} + \frac{1}{2}\left(False_{positive} + False_{negative}\right)}$$

- **The Importance:** The F1 score provides a balance between precision and recall, especially useful in situations where the dataset is imbalanced. It gives a single metric that combines both precision and recall, making it easier to evaluate the model’s performance [82].

5) IoU (Intersection over Union):

Intersection over Union is a popular metric in object detection models to calculate localization accuracy and compute localization errors. IoU is the ratio of the intersection of the two boxes' areas (predicted bounding box and a ground truth bounding box) to their combined areas. Mathematically, it is written as:

$$Intersection\ over\ Union\ (IoU) = \frac{A \cap B}{A \cup B}$$

$$Intersection\ over\ Union\ (IoU) = \frac{True_{positive}}{True_{positive} + False_{\ negative} + False_{positive}}$$

- **The Importance:** IoU measures the accuracy of the predicted bounding boxes. It is a key metric in object detection tasks, indicating how well the predicted regions match the actual regions [84].

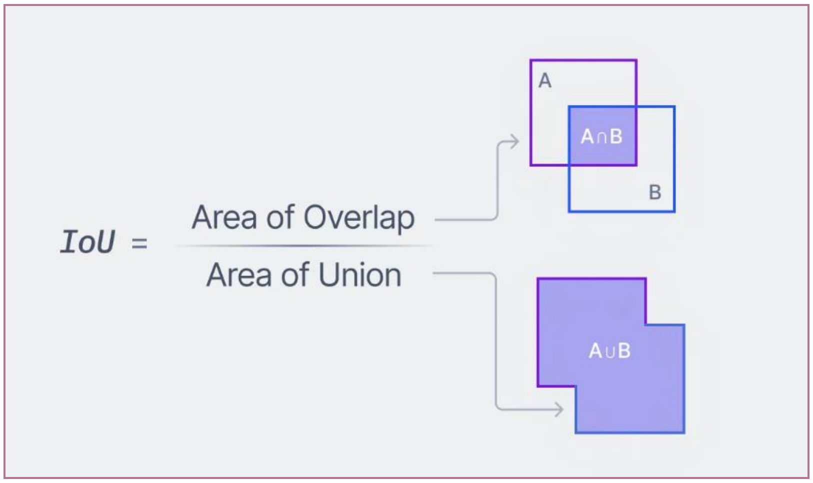


Figure (46): Visual Representation of IoU [84].

6) Confusion Matrix:

The confusion matrix (error matrix) is a special kind of table that allows visualization of the performance of an algorithm.

Each column of the matrix represents the instances in an actual class while row represents the instances in a predicted class, or vice versa. The diagonal of the matrix therefore represents all instances that are correctly predicted.

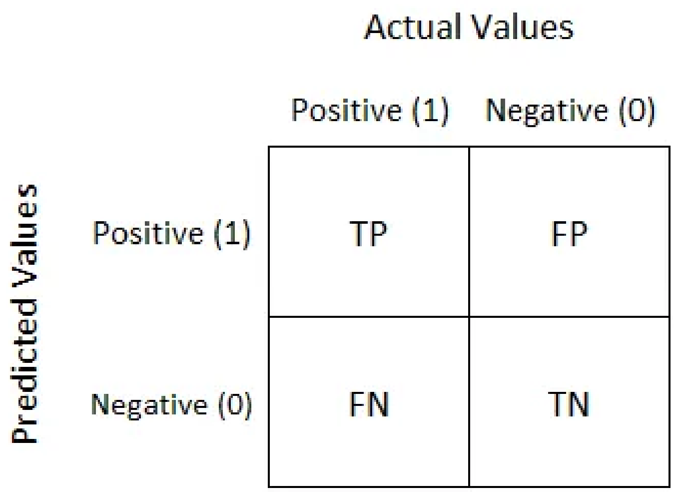


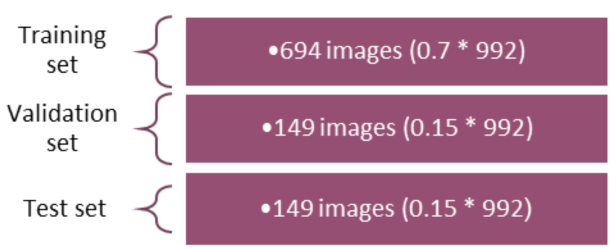
Figure (48): Confusion Matrix [86].

- **The Importance:** The confusion matrix provides details of the model's performance. It helps identify where the model is making errors (false positives and false negatives) and gives an idea into specific areas that need improvement [86-87].

10-9 Results and Explanation:

10-9-1 CNN Architecture Results:

The input images to CNN architecture for binary classification (as normal and abnormal) were 992 images (661 normal and 331 abnormal). The data split as 70-15-15 (training-validation-test):



Classification Confusion Matrix shown in figure (49):

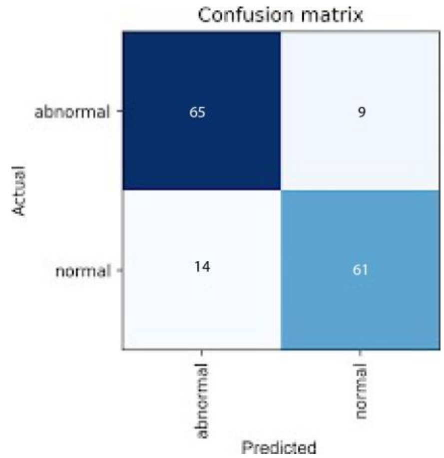


Figure (49): classification confusion matrix.

Classification Accuracy = (65 + 61) / 149 ≈ 0.8456 or 84.56%

Where:

- True Positives (TP): 65
- True Negatives (TN): 61
- False Positives (FP): 14
- False Negatives (FN): 9

10-9-2 Faster RCNN Results:

Images classified as true positives by CNN were 65 abnormal images. They fed to the Faster R-CNN for abnormality region detection, the results shown in figure (50) and metrics calculations shown in table (5) represent the detection results:

Figure (50): Detection Results.

Confidence Score: is a value given by the model to determine its ability identify the bounding box correctly and it is between [0-1].

True Positive TP: if the confidence score value is 0.5 and more the image considered as TP which mean the bonding box detect correctly and the overlapping with ground truth is big.

False Positive FP: if the confidence score value is less than 0.5, the image considered as FP which mean the bonding box detect but the overlapping with ground truth is small.

False Negative FN: when confidence score value is equal to 0, meaning there is no overlapping at all and the detection is incorrect.

Table (5): Detection Metrics Calculations.

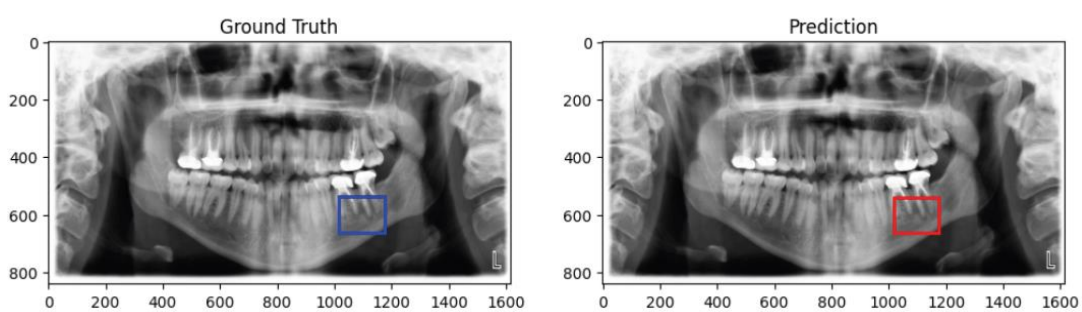
Metrics Calculations	Result
Precision	0.8235 or 82.35%
Recall	0.75 or 75%
F1 Score	0.785 or 78.5%
IoU	0.55 or 55%

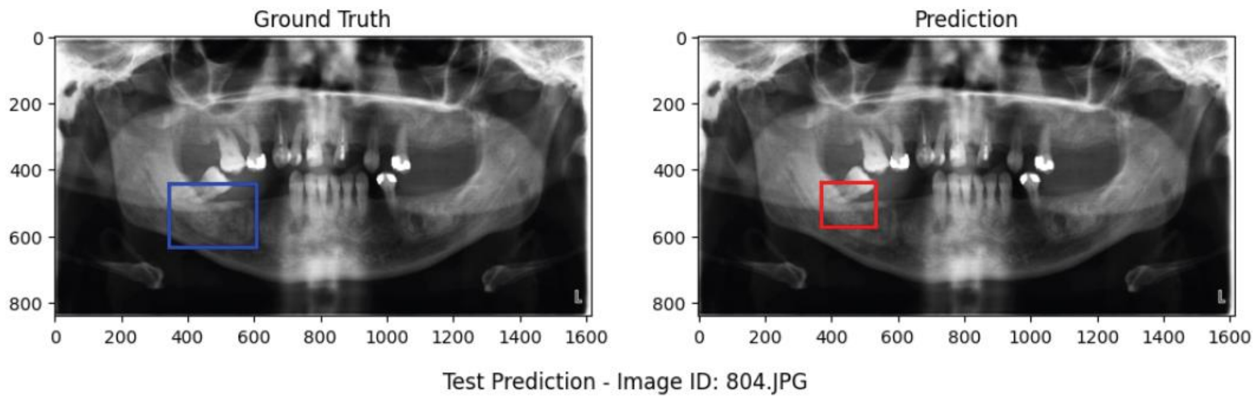
The Explanation:

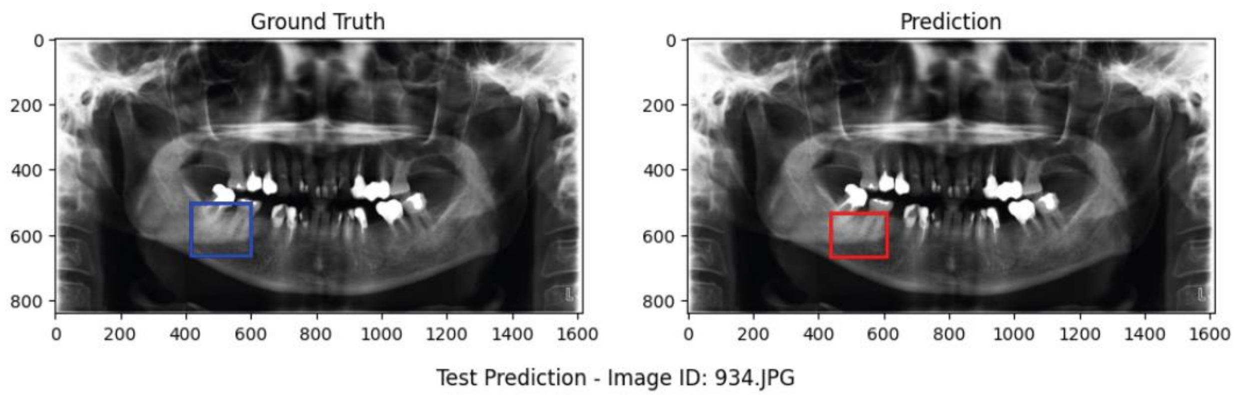
1. Precision of 82.35% indicates that when the model predicts an abnormality region, 82.35% it is correct. This suggests that the model is reasonably good at identifying true abnormality regions while minimizing false positives.
2. Recall of 75% means that the model successfully detects 75% of actual abnormality regions. This indicates that the model has a good detection rate for abnormalities and misses a few true positives.
3. F1 score of 78.5% indicates a good balance between precision and recall, meaning the model maintains a high detection rate (recall) while also being accurate in its positive predictions (precision).
4. IoU means the model predicted bounding boxes are 55% correctly overlapping with the ground truth data, the model can detect the abnormality regions but sometimes the bonding box was smaller or bigger than the actual regions.

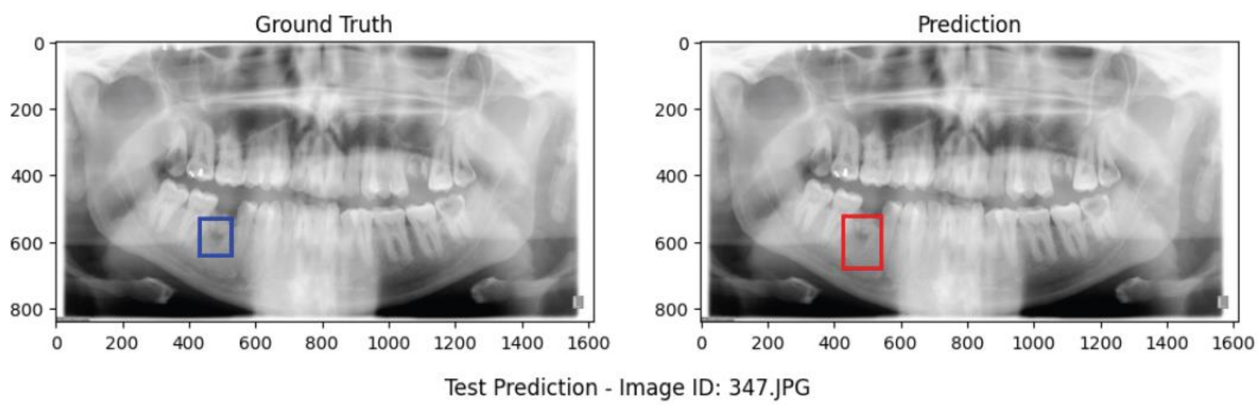
The model demonstrates good performance in detecting abnormality regions, with high precision and good recall. This reflects a good detection system, though the presence of some false positives and false negatives indicates need for improvement.

10-9-3 Detection Results on images:




Ground Truth
Prediction
Test Prediction - Image ID: 804.JPG


Ground Truth
Prediction
Test Prediction - Image ID: 934.JPG


Ground Truth
Prediction
Test Prediction - Image ID: 347.JPG

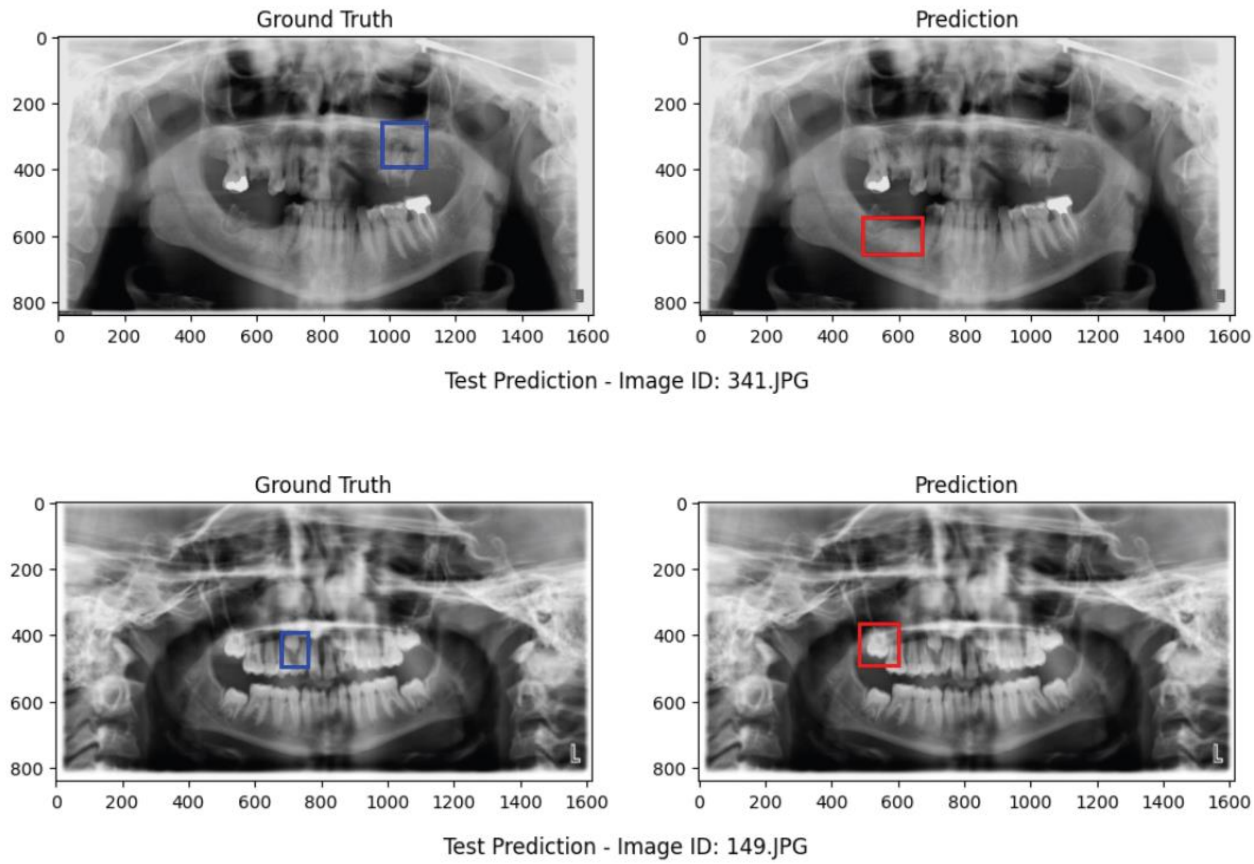


Figure (51): Detection Results om images.

11- Future Prospects:

1. Improving the accuracy of the model and its results in both classification and detection stages.
2. Generalizing the model to test external data.
3. Developing the model to accurately classify the detected object into different dental lesions (more classes).
4. Presenting the results to dentists and experts in the field for comparison.

Summary:

This chapter includes an explanation of the database with all tools that were used practically to achieve the goal in the required manner. It also includes the practical steps applied with their specific details then represents the results that were reached for the detection and classification stages.

References

[1] Medscape. (June 26). Mouth Anatomy. https://emedicine.medscape.com/article/1899122-overview

[2] National cancer institute. (2019). NCI Dictionary of Cancer Terms. National Cancer Institute; Cancer.gov. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/oral-cavity

[3] MedlinePlus [Internet]. Bethesda (MD): National Library of Medicine (US); [updated 2020 Jun 24]. Heart attack; [updated 2020 Jun 10; reviewed 2016 Aug 25; cited 2020 Jul 1]; [about 5 p.]. Available from: https://medlineplus.gov/heartattack.html

[4] Zwiri, A. M.., Patil, S., AL- Omair, F., Mousa, M. A., & Ali Ahmad, I. (2016). Prevalence Of Developmental Oral Mucosal Lesions Among A Sample Of Denture Wearing Patients Attending College Of Dentistry Clinics In Aljouf University. European Scientific Journal, ESJ, 12(24), 352. https://doi.org/10.19044/esj.2016.v12n24p352

[5] Syed, A. Z., Yannam, S. D., & Pavani, G. (2017). Research: Prevalence of Dense Bone Island. Compendium of continuing education in dentistry (Jamesburg, N.J.: 1995), 38(9), e13–e16.

[6] Sisman, Y., Ertas, E. T., Ertas, H., & Sekerci, A. E. (2011). The frequency and distribution of idiopathic osteosclerosis of the jaw. European journal of dentistry, 5(4), 409–414.

[7] Uribe, S. (2024). *Mandibular idiopathic osteosclerosis | Radiology Case | Radiopaedia.org*. Radiopaedia. https://radiopaedia.org/cases/mandibular-idiopathic-osteosclerosis

[8] Sun, C. X., Ririe, C., & Henkin, J. M. (2010). Hyperplastic dental follicle - review of literature and report of two cases in one family. The Chinese journal of dental research, 13(1), 71–75.

[9] Schmitd, L. B., Bravo-Calderón, D. M., Soares, C. T., & Oliveira, D. T. (2014). Hyperplastic dental follicle: a case report and literature review. Case reports in dentistry, 2014, 251892. https://doi.org/10.1155/2014/251892

[10] Schmitd, L.B., Bravo-Calderón, D.M., Soares, C.T., & Oliveira, D.T. (2014). Hyperplastic Dental Follicle: A Case Report and Literature Review. *Case Reports in Dentistry, 2014*.

[11] Juodzbalys G. (2022). Dental Implant Placement in Focal Osteoporotic Bone Marrow Defect: a Case Report and Treatment Recommendations. Journal of oral & maxillofacial research, 13(3), e5. https://doi.org/10.5037/jomr.2022.13305

[12] Santos, A., Santos, M., Barros, F. B. A., Gimenez, T., & Cavalcanti, M. G. P. (2023). Are imaging exams relevant in aiding the diagnosis of focal osteoporotic bone marrow defect: A systematic review. the Saudi Dental Journal. https://doi.org/10.1016/j.sdentj.2023.10.010

[13] Patel, V., Harwood, A., & McGurk, M. (2010). Osteomyelitis presenting in two patients: a challenging disease to manage. *British Dental Journal*, *209*(8), 393–396. https://doi.org/10.1038/sj.bdj.2010.927

[14] Rajaram Mohan, K., Pethagounder Thangavelu, R., & Fenn, S. M. (2022). Bilateral Inverted Mesiodens: A Rare Case Evaluated by Cone-Beam Computed Tomography. Cureus, 14(7), e26629. https://doi.org/10.7759/cureus.26629.

[15] Central Park Dentistry. (2024 June 26). Periapical Granuloma. https://centralparkdentistry.com/site/dental-health-rx-library/oral-pathology/periapical-granuloma/

[16] Al-Habib M. A. (2022). Prevalence and Pattern of Idiopathic Osteosclerosis and Condensing Osteitis in a Saudi Subpopulation. Cureus, 14(2), e22234. https://doi.org/10.7759/cureus.22234

[17] Santos, A., Santos, M., Barros, F. B. A., Gimenez, T., & Cavalcanti, M. G. P. (2023). Are imaging exams relevant in aiding the diagnosis of focal osteoporotic bone marrow defect: A systematic review. the Saudi Dental Journal. https://doi.org/10.1016/j.sdentj.2023.10.010

[18] Alberto, & Bianco, R. (2013). Bone reaction capability and the names of inflammatory bone disorders in endodontic clinical practice.

[19] Bhatt, G., Gupta, S., Ghosh, S., Mohanty, S., & Kumar, P. (2019). Central Osteoma of Maxilla Associated with an Impacted Tooth: Report of a Rare Case with Literature Review. Head and neck pathology, 13(4), 554–561. https://doi.org/10.1007/s12105-018-0994-3

[20] Chen, P., Liu, B., Wei, B., & Yu, S. (2022). The clinicopathological features and treatments of odontogenic keratocysts. American journal of cancer research, 12(7), 3479–3485.

[21] Radiographical approach to jaw lesions - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Odontogenic-keratocyst-in-a-36-year-old-woman-Panoramic-radiograph-shows-an-ellipsoid_fig4_5550311 [accessed 25 Jul 2024]

[22] Demiriz, L., Misir, A. F., & Gorur, D. I. (2015). Dentigerous cyst in a young child. European journal of dentistry, 9(4), 599–602. https://doi.org/10.4103/1305-7456.172619

[23] Satish, V., Prabhadevi, M. C., & Sharma, R. (2011). Odontome: A Brief Overview. International journal of clinical pediatric dentistry, 4(3), 177–185. https://doi.org/10.5005/jp-journals-10005-1106

[24] Niknami, M., Mirmohammadi, M., & Pezeshki, A. (2018). Evaluation of the Prevalence of Mucous Retention Pseudocyst and its Correlation with the Associated Risk Factors Using Panoramic Radiography and Cone-Beam Computed Tomography. Journal of dentistry (Tehran, Iran), 15(2), 123–129.

[25] Masthan, K. M., Anitha, N., Krupaa, J., & Manikkam, S. (2015). Ameloblastoma. Journal of pharmacy & bioallied sciences, 7(Suppl 1), S167–S170. https://doi.org/10.4103/0975-7406.155891

[26] Shylaja, S., Balaji, K., & Krishna, A. (2013). Nasopalatine duct cyst: report of a case with review of literature. Indian journal of otolaryngology and head and neck surgery: official publication of the Association of Otolaryngologists of India, 65(4), 385–388. https://doi.org/10.1007/s12070-011-0242-6

[27] Aksakal, G. (2024, April 10). *Neoplasm and Malignant Neoplasm*. Massive Bio.

https://massivebio.com/neoplasm-malignant-neoplasm/

[28] Adrain C. (2021). Systemic and cellular metabolism: the cause of and remedy for disease?. *The FEBS journal*, *288*(12), 3624–3627. https://doi.org/10.1111/febs.16033

[29] Shah, N. (2014). Recent advances in imaging technologies in dentistry. World Journal of Radiology, 6(10), 794. https://doi.org/10.4329/wjr.v6.i10.794

[30] Csurga. (2023, September 11). Imaging Techniques in Dentistry: A Continuously Evolving Path to Precise Diagnoses. Diamond DentArt. https://diamonddentart.hu/en/blog-en/kepalkoto-eljarasok-a-fogaszatban-egy-folyamatosan-fejlodo-ut-a-pontos-diagnozis-felallitasahoz/

[31] Radiology (ACR), R. S. of N. A. (RSNA) and A. C. of. (2022, June 1). Panoramic Dental X-ray. Radiologyinfo.org. https://www.radiologyinfo.org/en/info/panoramic-xray

[32] Różyło-Kalinowska, I. (2021). Panoramic radiography in dentistry. Clinical Dentistry Reviewed, 5(1). https://doi.org/10.1007/s41894-021-00111-4

[33] Perschbacher, S. (2012). Interpretation of panoramic radiographs. Australian Dental Journal, 57, 40–45. https://doi.org/10.1111/j.1834-7819.2011.01655.x

[34] Surlari, Z., Budală, D. G., Lupu, C. I., Stelea, C. G., Butnaru, O. M., & Luchian, I. (2023). Current Progress and Challenges of Using Artificial Intelligence in Clinical Dentistry-A Narrative Review. *Journal of clinical medicine*, *12*(23), 7378. https://doi.org/10.3390/jcm12237378

[35] American Psychological Association. (2020). Traditional approaches for lesion detection and classification. In *Diagnostic methods in dentistry: A comprehensive guide* (pp. 123-125). Retrieved from https://www.apa.org/pubs/books/diagnostic-methods-dentistry

[36] DDS, J. C. (2020, July 30). *Digital Dental Radiographs*. Jennifer Chiang, DDS. https://jenchiangdds.com/sunnyvale-dentist-patient-safety/dental-radiographs/

[37] Young, D. A., Nový, B. B., Zeller, G. G., Hale, R., Hart, T. C., Truelove, E. L., Ekstrand, K. R., Featherstone, J. D. B., Fontana, M., Ismail, A., Kuehne, J., Longbottom, C., Pitts, N., Sarrett, D. C., Wright, T., Mark, A. M., & Beltran-Aguilar, E. (2015). The American Dental Association Caries Classification System for Clinical Practice. The Journal of the American Dental

Association, 146(2), 79–86. https://doi.org/10.1016/j.adaj.2014.11.018

[38] Success and High Predictability of Intraorally Welded Titanium Bar in the Immediate Loading Implants - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Denture-with-radiopaque-markers-used-as-template_fig2_263394656 [accessed 25 Jul 2024]

[39] Laskowski, N., & Tucci, L. (2022, July 1). *What Is Artificial Intelligence (AI)?* TechTarget; TechTarget. https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence

[40] Wolfewicz, A. (2022, April 21). *Deep learning vs. machine learning – What's the difference?* Levity.ai. https://levity.ai/blog/difference-machine-learning-deep-learning

[41] Huang, C., Wang, J., Wang, S., & Zhang, Y. (2023). A review of deep learning in dentistry. *Neurocomputing*, *554*, 126629–126629. https://doi.org/10.1016/j.neucom.2023.126629

[42] Arsiwala-Scheppach, L. T., Chaurasia, A., Müller, A., Krois, J., & Schwendicke, F. (2023). Machine Learning in Dentistry: A Scoping Review. *Journal of Clinical Medicine*, *12*(3), 937. https://doi.org/10.3390/jcm12030937

[43] Ding, H., Wu, J., Zhao, W., Matinlinna, J. P., Burrow, M. F., & Tsoi, J. K. H. (2023). Artificial intelligence in dentistry—A review. *Frontiers in Dental Medicine*, *4*. https://doi.org/10.3389/fdmed.2023.1085251

[44] Agrawal, P., & Nikhade, P. (2022). Artificial Intelligence in Dentistry: Past, Present, and Future. *Cureus*, *14*(7), e27405. https://doi.org/10.7759/cureus.27405

[45] Katsumata A. (2023). Deep learning and artificial intelligence in dental diagnostic imaging. *The Japanese dental science review*, *59*, 329–333. https://doi.org/10.1016/j.jdsr.2023.09.004

[46] Schwendicke, F., Golla, T., Dreher, M., & Krois, J. (2019). Convolutional neural networks for dental image diagnostics: A scoping review. *Journal of Dentistry*, *91*, 103226. https://doi.org/10.1016/j.jdent.2019.103226

[47] Rafiatul Zannah, Bashar, M., Rahil Bin Mushfiq, Chakrabarty, A., Hossain, S., & Yong Ju Jung. (2024). Semantic Segmentation on Panoramic X-ray Images Using U-Net Architectures. *IEEE Access*, *12*, 44598–44612. https://doi.org/10.1109/access.2024.3380027

[48] Yang, S., Kim, K. D., Ariji, E., & Kise, Y. (2024). Generative adversarial networks in dental imaging: a systematic review. *Oral radiology*, *40*(2), 93–108. https://doi.org/10.1007/s11282-023-00719-1

[49] *What is R-CNN?* (2023, September 25). Roboflow Blog. https://blog.roboflow.com/what-is-r-cnn

[50] Kundu, R. (2023, January 17). *YOLO: Real-Time Object Detection Explained*. Www.v7labs.com. https://www.v7labs.com/blog/yolo-object-detection

[51] Yang, H., Jo, E., Kim, H. J., Cha, I. H., Jung, Y. S., Nam, W., Kim, J. Y., Kim, J. K., Kim, Y. H., Oh, T. G., Han, S. S., Kim, H., & Kim, D. (2020). Deep Learning for Automated Detection of Cyst and Tumors of the Jaw in Panoramic Radiographs. Journal of clinical medicine, 9(6), 1839. https://doi.org/10.3390/jcm9061839

[52] Ba-Hattab R, Barhom N, Osman SAA, Naceur I, Odeh A, Asad A, Al-Najdi SARN, Ameri E, Daer A, Silva RLBD, et al. Detection of Periapical Lesions on Panoramic Radiographs Using Deep Learning. Applied Sciences. 2023; 13(3):1516. https://doi.org/10.3390/app13031516

[53] Nashold, L., Pandya, P., & Lin, T. (2022). Multi-Objective Processing of Dental Panoramic Radiographs. https://cs231n.stanford.edu/reports/2022/pdfs/118.pdf

[54] Çelik, B., Savaştaer, E. F., Kaya, H. I., & Çelik, M. E. (2023). The role of deep learning for periapical lesion detection on panoramic radiographs. Dento maxillo facial radiology, 52(8), 20230118. https://doi.org/10.1259/dmfr.20230118

[55] Lee, S., Kim, D. & Jeong, HG. Detecting 17 fine-grained dental anomalies from panoramic dental radiography using artificial intelligence. Sci Rep 12, 5172 (2022). https://doi.org/10.1038/s41598-022-09083-2

[56] Helmi Mahran, A. M.; Hussein, W.; Saber, S. E. D. M. Automatic Teeth Segmentation Using Attention U-Net. Preprints 2023, 2023061468. https://doi.org/10.20944/preprints202306.1468.v2

[57] Rubiu G, Bologna M, Cellina M, Cè M, Sala D, Pagani R, Mattavelli E, Fazzini D, Ibba S, Papa S, et al. Teeth Segmentation in Panoramic Dental X-ray Using Mask Regional Convolutional Neural Network. Applied Sciences. 2023; 13(13):7947. https://doi.org/10.3390/app13137947

[58] 1. 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. https://ieeexplore.ieee.org/document/9557804/references#references

[59] Sheremetov, D. (2023, November 9). 7 Reasons Why Python is Best for AI, ML, and Deep Learning. Onix-Systems.com. https://onix-systems.com/blog/python-is-best-for-ai-ml-and-deep-learning

[60] Yadoshchuk, V. (2024, March 1). Why Use Python for AI and Machine Learning. Waverley. https://waverleysoftware.com/blog/python-for-ai-and-ml/

[61] Crockford, D. (2018). JSON File Extension - What is a .json file and how do I open it? Fileinfo.com. https://fileinfo.com/extension/json

[62] Ozanich, A. (2022, September 2). What Are JSON Files & How Do You Use Them? Blog.hubspot.com. https://blog.hubspot.com/website/json-files

[63] Pillow — Pillow (PIL Fork) 6.2.1 documentation. (2011). Readthedocs.io. https://pillow.readthedocs.io/en/stable/

[64] Python Imaging Library. (2021, May 31). Wikipedia. https://en.wikipedia.org/wiki/Python_Imaging_Library

[65] Team, G. L. (2021, August 5). OpenCV Tutorial in Python. GreatLearning Blog: Free Resources What Matters to Shape Your Career! https://www.mygreatlearning.com/blog/opencv-tutorial-in-python/

[66] ActiveState. (2022, August 9). What Is Pandas in Python? Everything You Need to Know. ActiveState. https://www.activestate.com/resources/quick-reads/what-is-pandas-in-python-everything-you-need-to-know/

[67] NumPy. (2022). What is NumPy? — NumPy v1.19 Manual. Numpy.org. https://numpy.org/doc/stable/user/whatisnumpy.html

[68] What is PyTorch? (2024). NVIDIA Data Science Glossary. https://www.nvidia.com/en-us/glossary/pytorch/

[69] Simplilearn. (2021, December 2). What is PyTorch, and How Does It Work? | Simplilearn. Simplilearn.com. https://www.simplilearn.com/what-is-pytorch-article

[70] Terra, J. (2020, July 27). Keras vs Tensorflow vs Pytorch: Popular Deep Learning Frameworks. Simplilearn.com. https://www.simplilearn.com/keras-vs-tensorflow-vs-pytorch-article

[71] torchvision — Torchvision master documentation. (2024). Pytorch.org. https://pytorch.org/vision/stable/index.html

[72] Wikipedia Contributors. (2019, March 19). scikit-learn. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Scikit-learn

[73] Meghna, K. (2018, May 14). Python | Introduction to Matplotlib - GeeksforGeeks. GeeksforGeeks. https://www.geeksforgeeks.org/python-introduction-matplotlib/

[74] Datasets & DataLoaders — PyTorch Tutorials 1.11.0+cu102 documentation. (2024). Pytorch.org. https://pytorch.org/tutorials/beginner/basics/data_tutorial.html

[75] Writing Custom Datasets, DataLoaders and Transforms — PyTorch Tutorials 1.10.1+cu102 documentation. (2024). Pytorch.org. https://pytorch.org/tutorials/beginner/data_loading_tutorial.html

[76] *Transforming and augmenting images — Torchvision main documentation*. (2024). Pytorch.org. https://pytorch.org/vision/main/transforms.html

[77] *Transforming and augmenting images — Torchvision 0.15 documentation*. (2024). Pytorch.org. https://pytorch.org/vision/0.15/transforms.html

[78] AWS. (2023). What is Data Augmentation? - Data Augmentation Techniques Explained - AWS. Amazon Web Services, Inc. https://aws.amazon.com/what-is/data-augmentation/

[79] Gad, A. F. (2020, November 13). Faster R-CNN Explained for Object Detection Tasks. Paperspace Blog. https://blog.paperspace.com/faster-r-cnn-explained-object-detection/

[80] GeeksforGeeks. (2020, February 27). Faster R-CNN | ML. GeeksforGeeks. https://www.geeksforgeeks.org/faster-r-cnn-ml/

[81] Gupta, A. (2021, October 7). *A Comprehensive Guide on Deep Learning Optimizers*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-deep-learning-optimizers/

[81] Gad, A. F. (2020a, October 12). *Accuracy, Precision, and Recall in Deep Learning*. Paperspace Blog. https://blog.paperspace.com/deep-learning-metrics-precision-recall-accuracy/

[82] amsten. (2020, September 1). *ML | Evaluation Metrics*. GeeksforGeeks. https://www.geeksforgeeks.org/metrics-for-machine-learning-model/

[83] Overview of Machine Learning Part 1 - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Visualizing-accuracy-recall-aka-sensitivity-and-precision-which-are-the-common_fig3_346129022 [accessed 27 Jul 2024]

[84] Shah, D. (2023, May 30). *Intersection over Union (IoU): Definition, Calculation, Code*. Www.v7labs.com. https://www.v7labs.com/blog/intersection-over-union-guide

[85] Anwar, A. (2022, May 13). *What is Average Precision in Object Detection & Localization Algorithms and how to calculate it?* Medium; Towards Data Science. https://towardsdatascience.com/what-is-average-precision-in-object-detection-localization-algorithms-and-how-to-calculate-it-3f330efe697b

[86] Narkhede, S. (2018). *Understanding Confusion Matrix*. Medium; Towards Data Science. https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62

[87] Wikipedia Contributors. (2019b, October 22). *Confusion matrix*. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Confusion_matrix