Deep Learning Approach for Detection of Dental Caries in X -Ray Images

[1]Jyothi G. C., [2] Dr. Chetana Prakash, [3] Dr. Babitha G. A., [4] Kiran Kumar G. H.

[1] Department of Information Science and Engineering
[2]Department of Computer Science and Engineering
[4] Department of Electronics and Communication, Bapuji
[1,2,4] Institute of Engineering and Technology, Davangere-577004, Karnataka,
[3]Department of Periodontics College of Dental Science, Davangere-577004, Karnataka

***Abstract*- Dental caries is the most prevalent disease in the world, affecting more than 3.5 billion people. Dental caries must be treated, which costs money and takes time in every nation's healthcare system. Early disease detection can improve a patient's prognosis and lower the cost of treatment. X-ray imaging is the method used to detect dental caries most frequently after the visual method. A proximal and early-stage carious lesion can be easily missed by the visual examination, so X-ray imaging is very beneficial for early detection and a chance of healing without the need for additional dental care. Using M-FCM and Level Set Techniques, this paper addresses the problems of segmenting dental X-ray images and detecting caries using Faster R- CNN and YOLO V5 Deep Learning algorithms. A dataset of 1200 X-ray images with 800 dental caries annotations was generated. We used it to compare the performance of various architectures that we trained for object detection.**

***Keywords*- Dental caries detection, M-FCM, Level Set, Faster R-CNN, YOLO V5, Deep Learning, X-ray images.**

I. INTRODUCTION

Dental caries, which affects more than 3.5 billion people, is the most common disease in the world [2]. The frequency has not decreased despite advancements in medical technology, placing strain on the global health care systems. In [10], dental expenses made up more than 6% of all healthcare costs. Over the past ten years, machine learning, particularly neural networks, has advanced significantly, outperforming humans in the ImageNet classification task [12]. Since then, the ImageNet dataset's deep learning model error rates have decreased by a factor of four[14]. Deep learning is now widely used in numerous fields, including medical imaging, as a result of this dramatic increase. On some tasks, like the detection of caries, deep learning models performed better than humans [11]. The goal of this study is to create a deep-learning model that can identify dental cavities in X-ray images. Each carious lesion has a basic bounding box around it to identify its location. With the aid of this model, dentists will be able to validate the accuracy of their initial determination that carious lesions were present in the X-ray image. Additionally, dentists gain from the ability to save information about dental caries in a digital format without their involvement by instantly identifying the location of caries from the image. By superimposing an X-ray image with the location of dental caries, dentists can better help patients understand the issue. Additionally, they could decide to save the data for later use in observing the lesion's growth. Additionally, such software would be helpful in the classroom because it would allow dentistry students to practice spotting dental cavities without the aid of a tutor.

II. LITERATURE REVIEW

Medical background

Human teeth:
The primary and permanent tooth sets make up the human dentition. At six months of age, the primary, or deciduous, starts to erupt. It has twenty teeth. Around the age of 13, a permanent set of teeth with 32 teeth completely replaces this dentition. On the basis of their shape and intended use, they can be categorized into four groups. These classes include:

Incisors:
Both the primary and permanent dentitions have eight incisors in total. They are situated at the mouth's opening and primarily serve to cut and shear food. They affect a smile's phonetics as well as its appearance.

Canines:
Dental arches have four canines overall, which divide them into a frontal and lateral portion at their corners. They are triangular in shape, and the incisal edge has a single cusp point. They can grasp, pierce, shred, and cut food thanks to the structure. They are necessary for aesthetics along with the incisors.

Premolars:
All eight of the primary molars are followed by the premolars, which are only present in the permanent dentition. Premolars can grasp and grind food because of their anatomical structure, which is similar to that of canines and molars.

Molars:
There are twelve permanent molars in humans, and none of them had deciduous predecessors. They are primarily responsible for chopping food into manageable bites. Their wide occlusal surfaces enable them to perform this function. Due to the extensive area of contact between neighboring molars and the deep grooves that run along the occlusal surface of the teeth, molars are particularly susceptible to dental caries. Because these areas are difficult to clean, bacteria frequently congregate there.

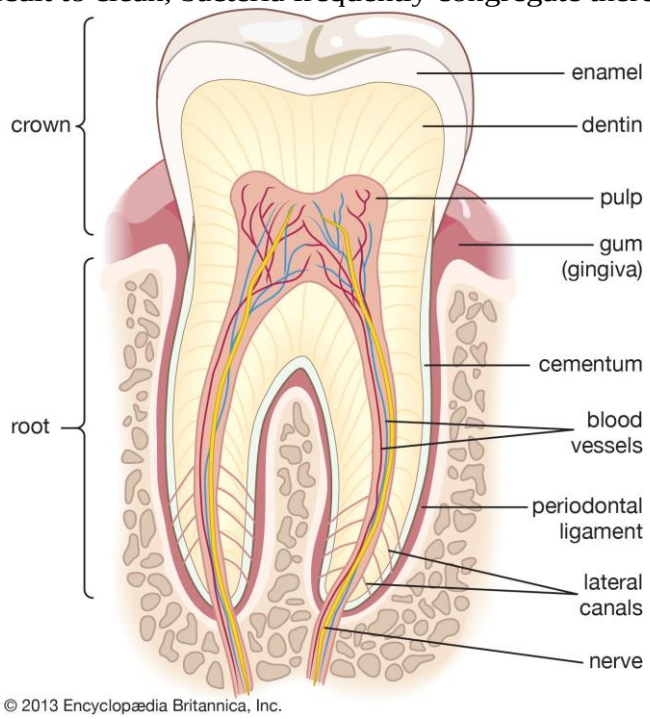

Figure 1: Structure of teeth

The cementum, pulp-dentin complex, and enamel make up the three structural components of teeth. Figure 1 depicts the tooth's internal architecture. The enamel, a highly mineralized crystalline structure, covers the surface of the anatomical crown of a tooth. Enamel is the hardest material found in teeth and even the human body because minerals (hydroxyapatite) make up more than 90% of its volume.

Its thickness typically ranges between 2 and 3mm, though it differs from tooth class to tooth class. Enamel cannot regenerate because it is made by cells that are only present during the amelogenesis process' developmental stage. Acidic conditions can cause demineralization, which puts the enamel most at risk. Enamel can remineralize, but if the underlying cause of the issue is not addressed, the enamel is irreparably damaged and a cavity develops.

Pulp-Dentin complex:
Two specialised connective tissues are pulp and dentin. However, some sources claim that they are just one tissue in a complex [27]. Four purposes are carried out by dental pulp: reparative, nutritive, formative, and sensory. It is situated in the pulp chamber of the tooth. Dentin, which is produced by particular cells during the dentino genesis process, surrounds the pulp. Although their cytoplasmic cell processes, which are housed in dentinal tubules, extend into the mineralized dentin, their cell bodies are found in the pulp chamber.

These processes are thought to make dentin a living tissue. Its purpose is to help the body heal and respond to pathological stimuli, such as causing minerals to precipitate in the affected area to stop the spread of carious lesions, and to help the body do both. The most important part of a tooth is the dentin. The coronal portion of the tooth is covered by enamel, and the root is covered by cementum. There are various varieties of dentin. The dentin layer that is closest to the enamel is called primary dentin, and it is also the outermost and most noticeable layer. It is created when teeth are developing. Secondary dentin develops after root development is finished. The production of tertiary (reactive) dentin is increased in response to pathological

stimuli like damage or caries. To protect the pulp, it is created at the pulp-dentin interface. The presence of mineral precipitates in the dentinal tubules as a result of damage or ageing allows transparent dentin to be distinguished from other types of dentin.

Dental caries

Cause

Dental caries is an infectious condition that causes the hard dental tissues to become demineralized. Plaque on the teeth is the main offender (also called a biofilm). Plaque is a substance that can adhere to tooth structure and is made up of bacteria, their waste products, and saliva. Bacteria that create byproducts of organic acids from refined dietary carbohydrates reside in the plaque. The pH of the biofilm may drop below a crucial level (5.5 for enamel and 6.2 for dentin) if these acids are present for an extended period of time [16].

In an effort to restore equilibrium, low pH drives phosphate and calcium from the tooth into the biofilm. Demineralization, which occurs when minerals in a tooth are lost, can lead to a caries lesion if it is not stopped. If the pH returns to neutral and the relative concentration of soluble calcium and phosphate in the biofilm is higher than in the tooth, this process can be controlled and eventually stopped. The cycle of demineralization and remineralization occurs several times each day and is regulated by a number of highly personal and tooth-specific factors.

Epidemiology

The most prevalent medical condition is dental caries in permanent teeth, which is treated poorly [20]. 35% of the world's population felt the effects in 2010. People over 25 were found to have the highest prevalence of the condition.

Since the prevalence did not significantly change between 1990 and 2010 [20] [18], it is obvious that advancements in dental technology had no impact. By watching 100 people, Kassebaum et al. predict that 42 new cases of tooth decay in both primary and permanent teeth will appear each year. Health care systems are overworked as a result. The cost of dental care alone in the United States in 2016 was 0.1 trillion dollars, out of a total of 1.62 trillion dollars spent on healthcare, according to Huang et al. [15].

Diagnosis

The most effective method for checking for caries when inspecting teeth is visual-tactile diagnosis. The examination is carried out by dentists using a mouth mirror and pointed probe. The refractive index difference between healthy and carious enamel is greater when the tissue is dehydrated, so drying the teeth is crucial.

This increases the chance of finding a carious lesion early on, before it worsens and the tooth cavitates. In addition to the visual examination, dental X-rays are the second most popular method used by clinicians. Both intraoral (where the X-ray film is placed inside the mouth) and extraoral (where the film is placed outside the mouth) X-ray imaging are used in dentistry examinations (the X-ray film is outside the mouth).

The most common kinds of images are intraoral ones. Periapical and bitewing X-rays, which each display different dental features, fall under this category. Extraoral imaging is primarily used to identify dental problems near the jaw and skull. The most popular type of radiograph is a panoramic one [20].

Figure 2: Bitewing X-ray image and Periapical X-ray image

III. DATASET

Collection and annotation of data sets to the best of our knowledge, there isn't any publicly accessible dataset that includes X-ray radiographs of dental caries. The raw datasets with dental carious regions are also scarce and of poor quality. The research into dental cavity localization techniques is restricted by the available datasets. We created a database of 936 manually annotated X-ray radiographs that were primarily acquired from a hospital.

Labels

The Xml label file of each sample was created using the Labeling software after the doctors marked the caries locations. The size and coordinates of the caries in the sample are contained in the XML file. After that, the Xml files were used to create the sample set in VOC format.

III. PROPOSED METHOD

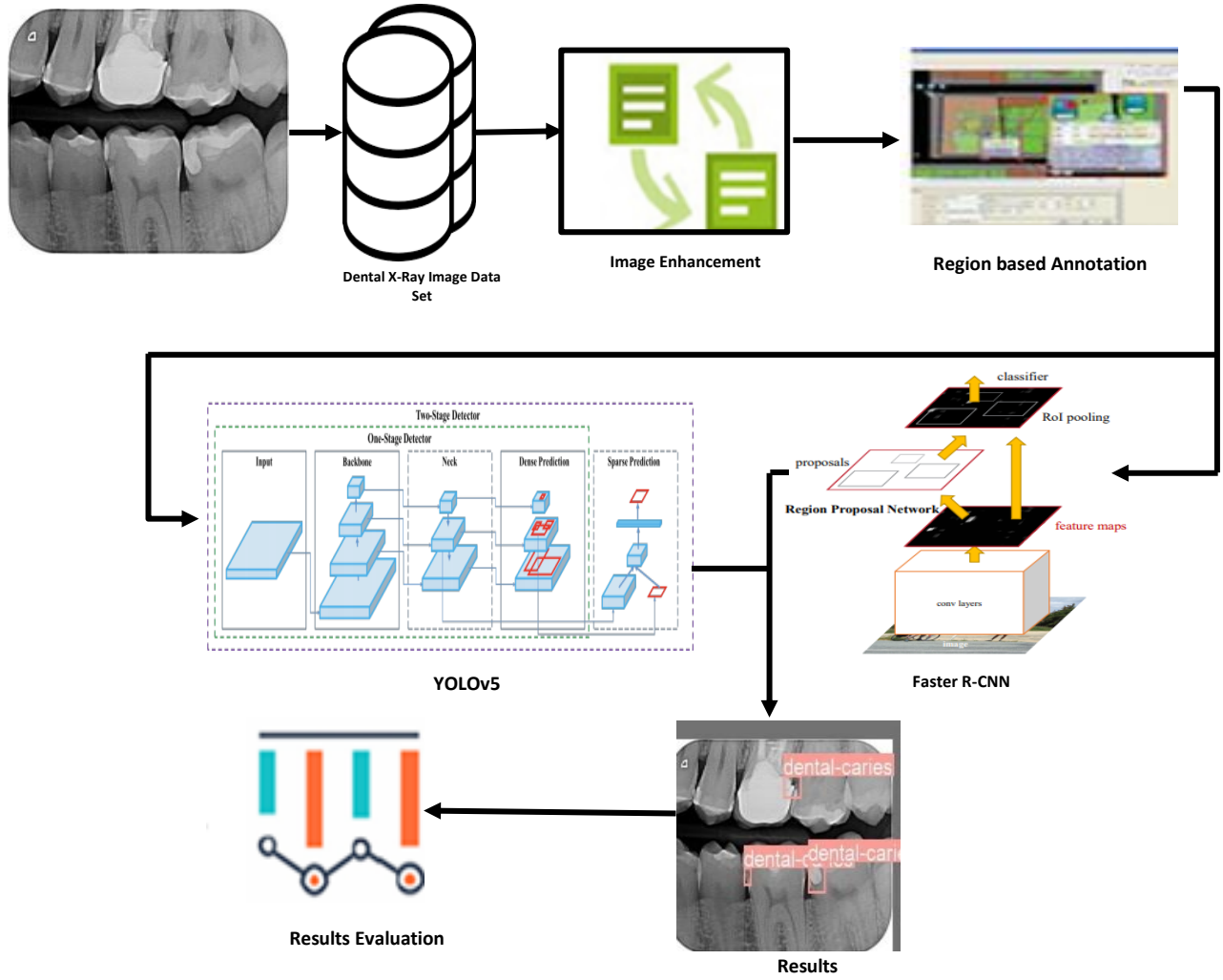

Figure 3: Overview of proposed dental caries regions detection and localization in X – Ray images.

The Figure 3 shows the overview of the proposed dental caries regions detection and localization in X– Ray images. Faster-RCNN is a two-stage object detection network algorithm. Compared to one-stage algorithms like YOLO, two-stage algorithms have a more complex network architecture. However, they also have relatively high detection precision at the expense of slower sample detection speeds. When used with a GPU server, the suggested method. enables the detection of a single sample in about 0.2 seconds. The requirements for clinical use are met by this detection time.

Faster RCNN

Faster RCNN is an object detection network created on the RCNN architecture. Additional backbone feature extraction networks that are frequently used include ResNet50, Xception, and VGG16. We used ResNet50, a residual network that can address the problem of network degradation, for this study. Figure 4 depicts the architecture of the Faster-RCNN network. The residual network architecture can effectively address the problem of network degradation [17]. When the number of network layers increases, which happens when the training accuracy starts to get close to saturation, the result is known as "network degradation," which is a decrease in training accuracy. This decline in training precision is unrelated to overfitting. The residual network's function is to provide the network with the capability of identical mapping in order to maintain the network's performance level as the number of network layers increases.

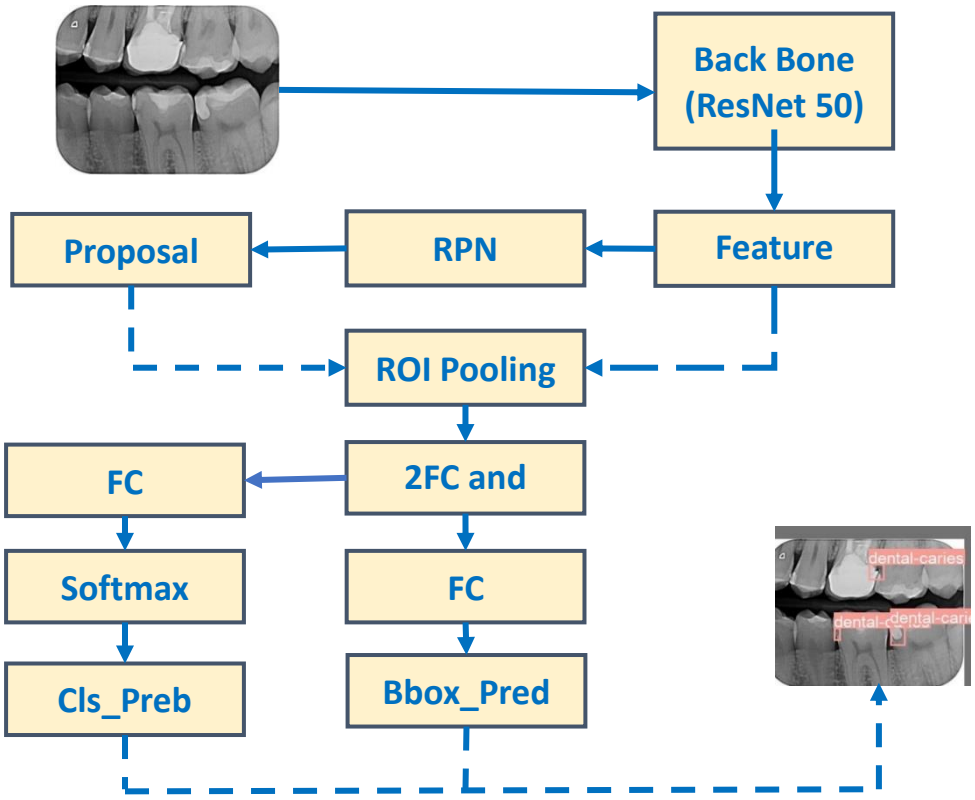

Figure 4: Network Architecture of Faster-RCNN.

The network design is shown in Figure 5. The left section displays the basic network architecture, while the right section displays the residual architecture. The network optimization effect f(X) after the network depth is increased shows that the network performance does not improve with an increase in network depth. As a result, the mapping f(X) X is identical, preventing any degradation as the network depth increases. As a result, ResNet50 was selected as the main feature extraction network in this study. Network architecture of RPN shown in Figure 6. Its main architecture consists of the RPN (regional proposal network), prediction & localization network, and backbone feature extraction network.

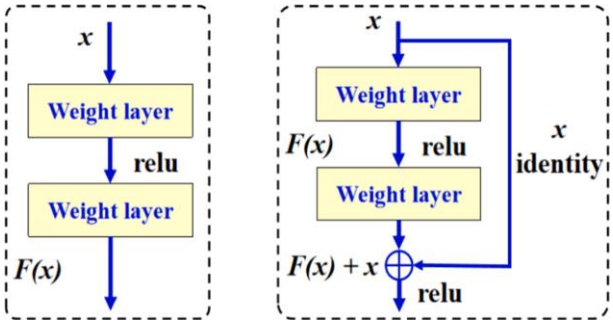

Figure 5: Normal network architecture and residual network architecture.

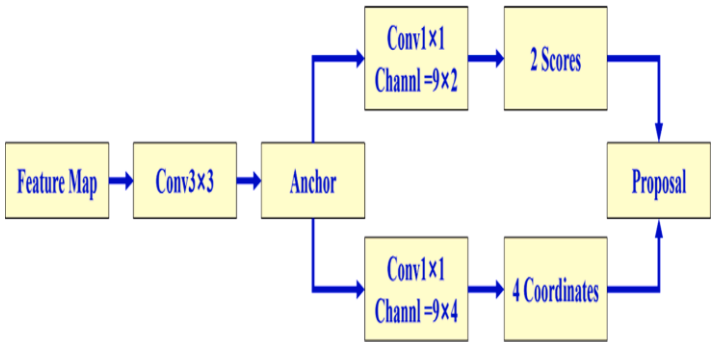

Figure 6: Network architecture of RPN

YOLOv5

Single-stage and two-stage detectors are the two types of object detection models that exist. The Backbone, the Neck, and the Head are the three components that single-stage object detectors (like the YOLO architecture) are made up of in order to make dense

predictions. The backbone is a pre-trained network that extracts rich feature representations for images. The resultant image has a lower spatial resolution but a higher feature (channel) resolution. The model neck is utilised for the extraction of pyramidal features. This helps the model generalise well to objects of different sizes and scales. CSP-Darknet53 serves as the framework for YOLOv5. The authors simply used the Darknet53 convolutional network, which served as the basis for YOLOv3, and the Cross Stage Partial (CSP) network strategy.

Information can flow to the deepest layers and the vanishing gradient problem is avoided by the YOLO deep network, which uses residual and dense blocks. However, using dense and residual blocks has the advantage of avoiding the problem of redundant gradients. CSPNet helps to resolve this issue by truncating the gradient flow. YOLOv5's CSPNet strategy divides the feature map of the base layer into two parts, which are then combined using a cross-stage hierarchy. The use of this method has significant advantages for YOLOv5, as it reduces the number of parameters and calls for less processing power (less FLOPS), which accelerates inference speed, a crucial aspect of real-time object detection models. YOLOv5 made two significant changes to the model neck. After initially using a variation of Spatial Pyramid Pooling (SPP), the BottleNeckCSP has been added to the architecture of the Path Aggregation Network (PANet) in order to modify it. After the SPP block [4] aggregates the data it receives from the inputs, a fixed length output is generated.
The most crucial context features are thus separated, and this has the advantage of greatly expanding the receptive field without slowing down the network. Earlier versions of YOLO (YOLOv3 and YOLOv4) used this block to separate the most important features from the backbone, but in YOLOv5 (6.0/6.1), SPPF—just another variation of the SPP block—was used to boost network speed.

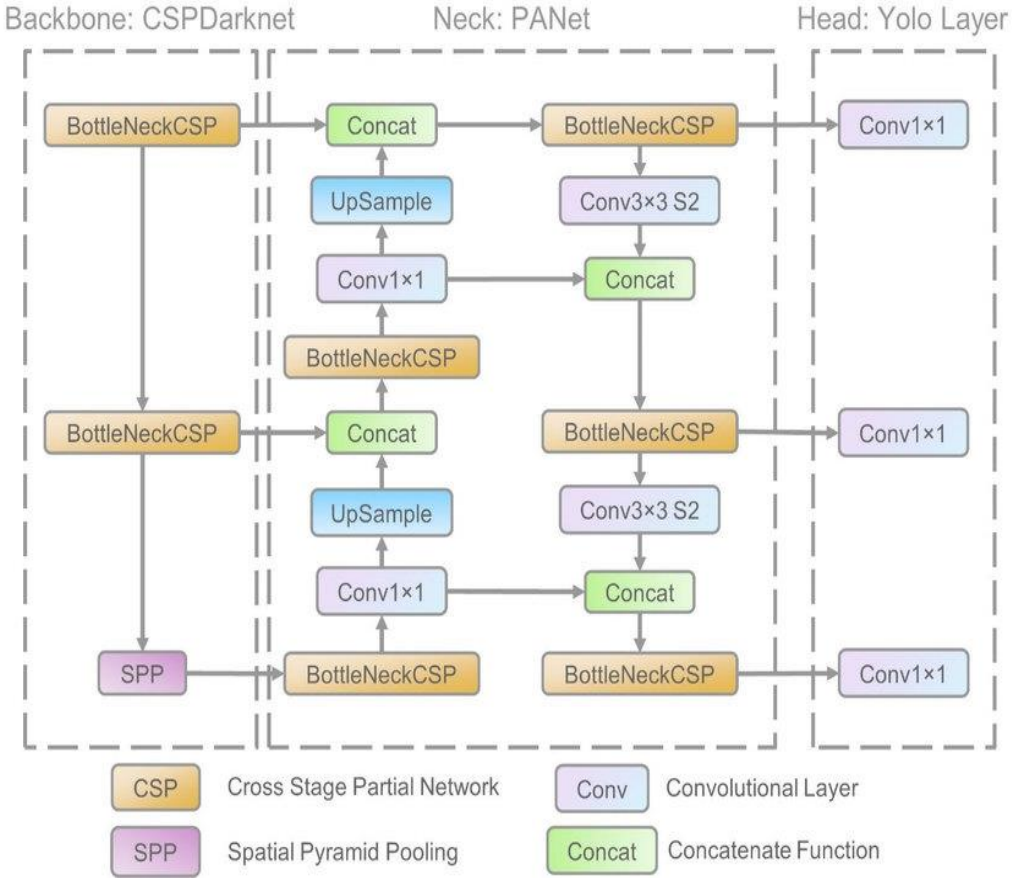

Figure 7: Network Architecture of YOLO V5.

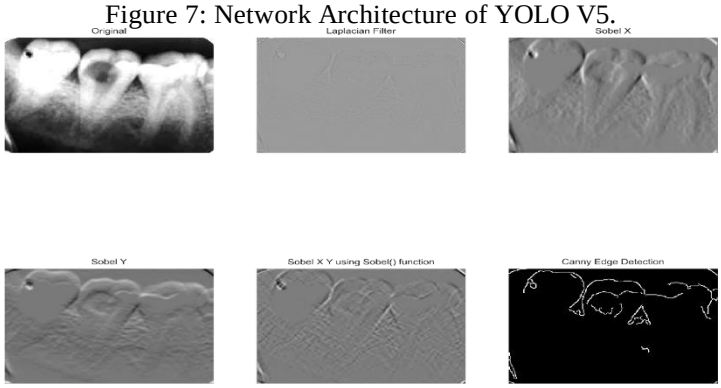

Figure 8: Image Enhancement

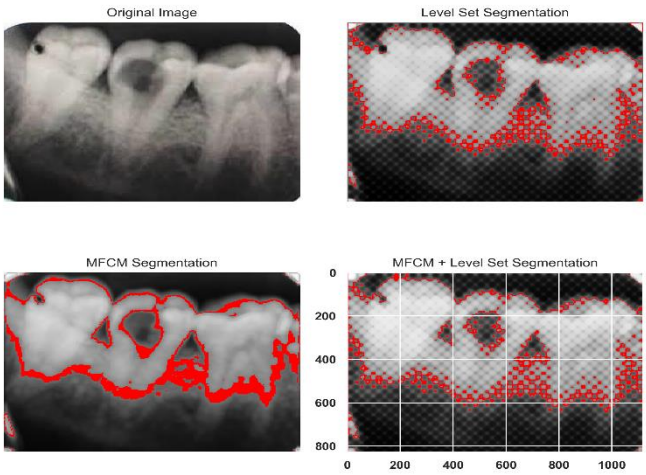

Figure 9: Dental X -Ray Image Segmentation using MFCM and Level Set Algorithms

IV. PERFORMANCE EVALUATION

Table 1: Data set for Batch 1 and 2

	Total Images -1000	
	Batch -1 (80-20)	Batch -2 (70-30)
No. of Training Images	800	700
No. of Testing Images	200	300

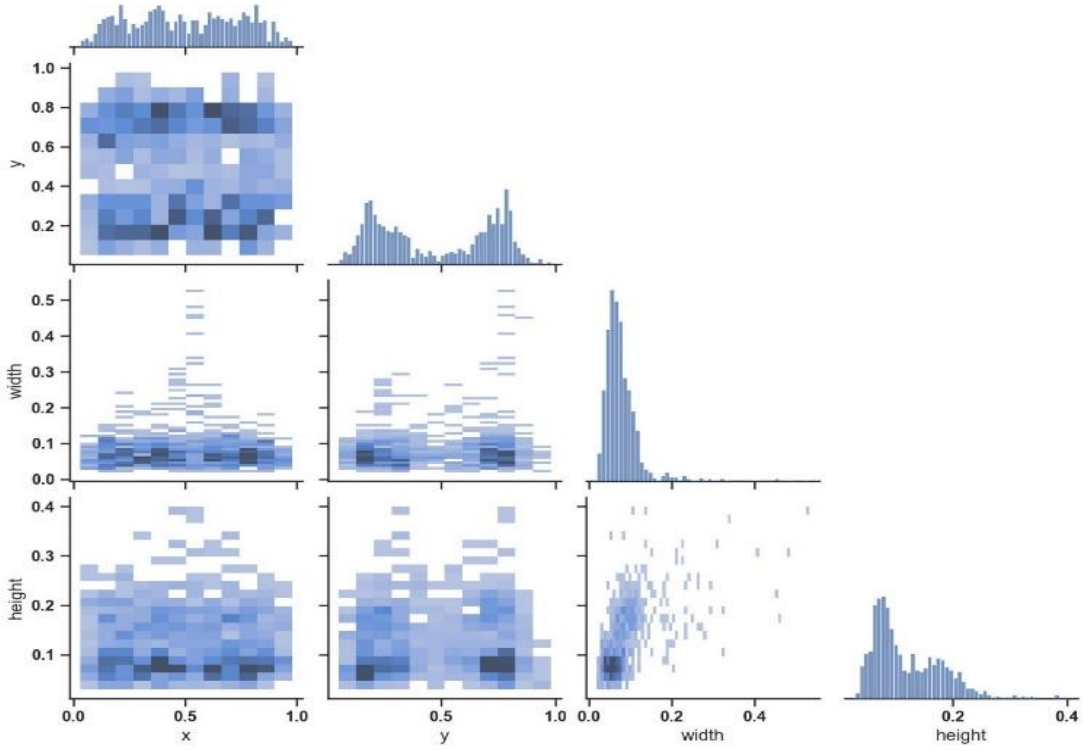

Figure 10: Labels Correlogram

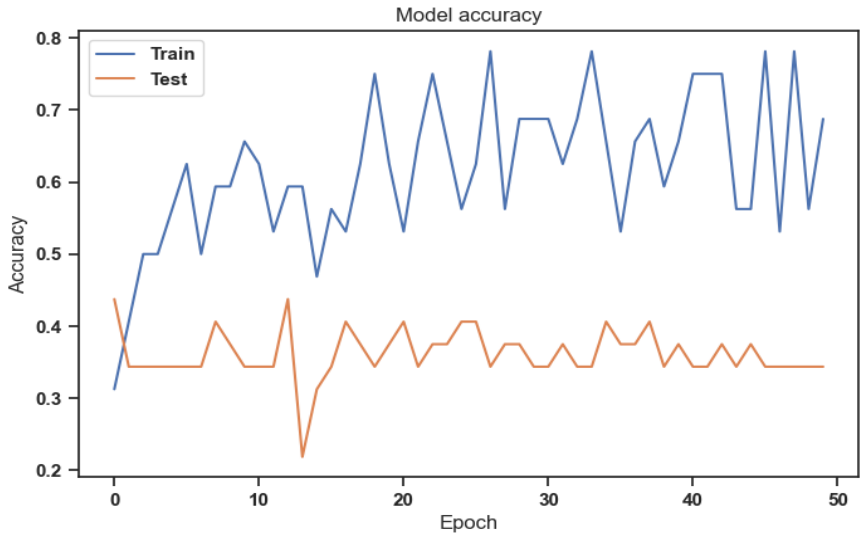

Figure 11: Model Accuracy

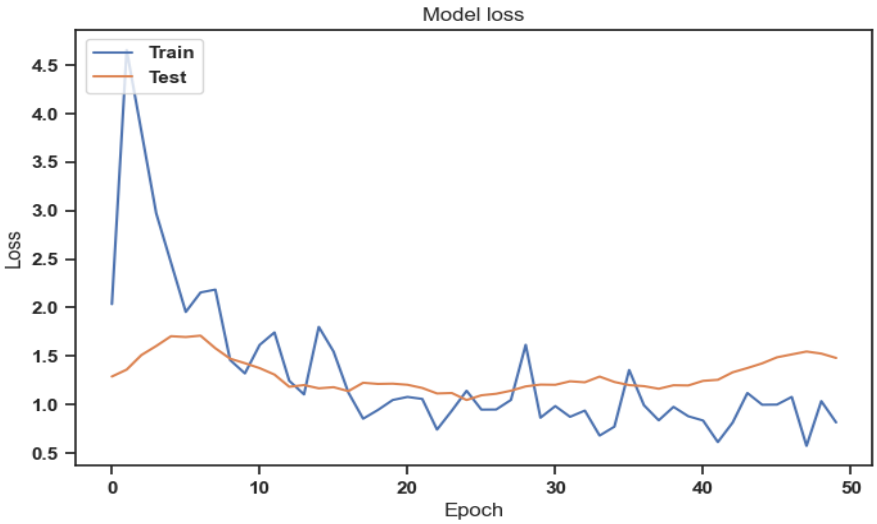

Figure 12: Model Loss

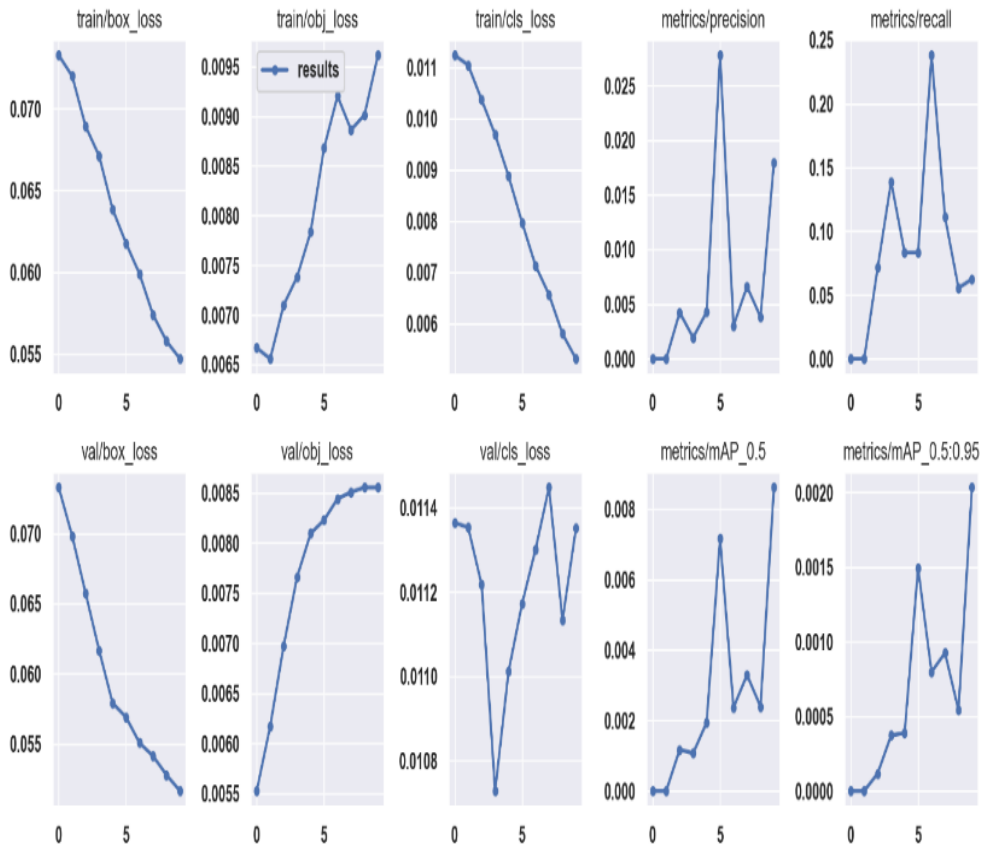

Figure 13: Metric curves

Table 2: Performance analysis for Batch -1 Dataset

Models/ Metrics	Accuracy	Precision	Recall	F1-Score
Faster R-CNN model	0.75	0.89	0.84	0.85
YOLO Model	0.88	0.94	0.86	0.92

Table 3: Performance analysis for Batch -2 Dataset

Models/ Metrics	Accuracy	Precision	Recall	F1-Score
Faster R-CNN model	0.73	0.73	0.82	0.77
YOLO Model	0.83	0.80	0.92	0.85

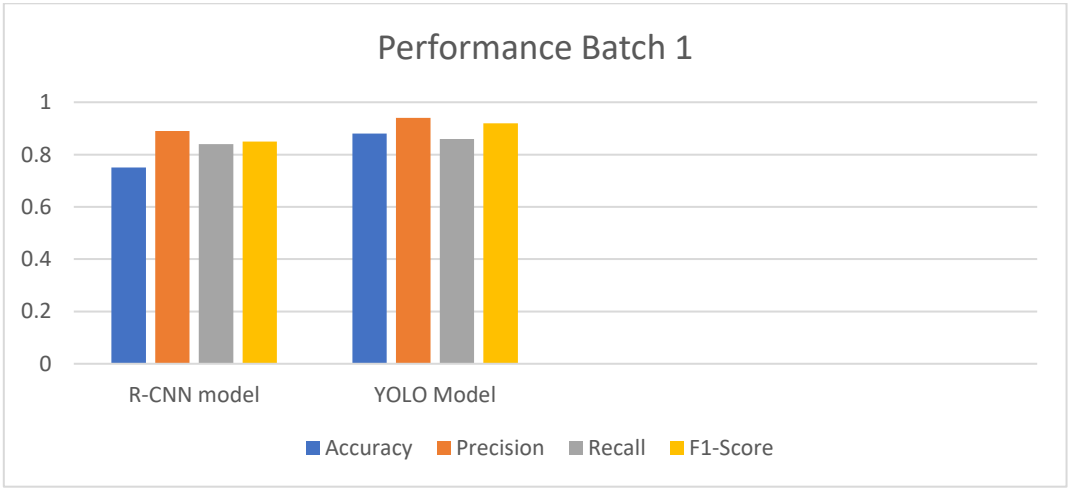

Figure 14: Performance graph for Batch 1 Dataset

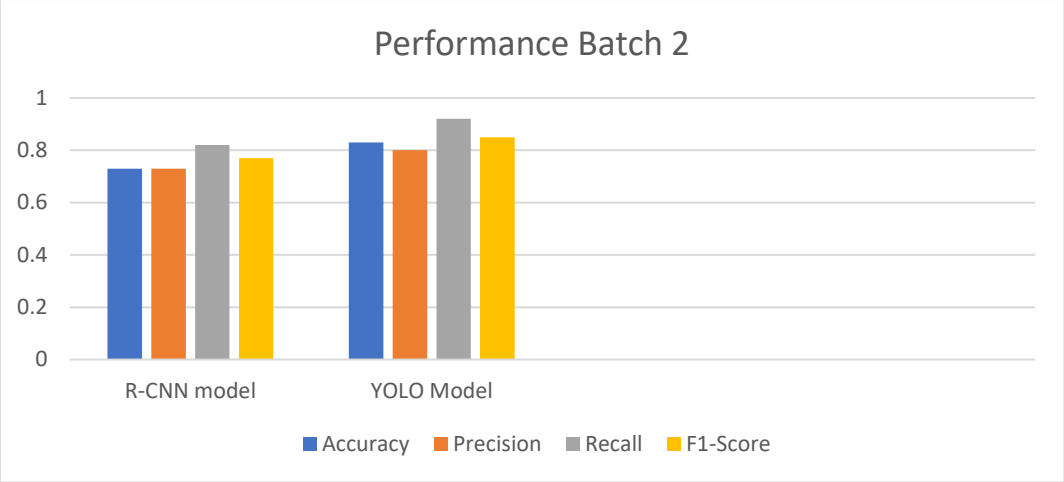

Figure 15: Performance graph for Batch 2 Dataset

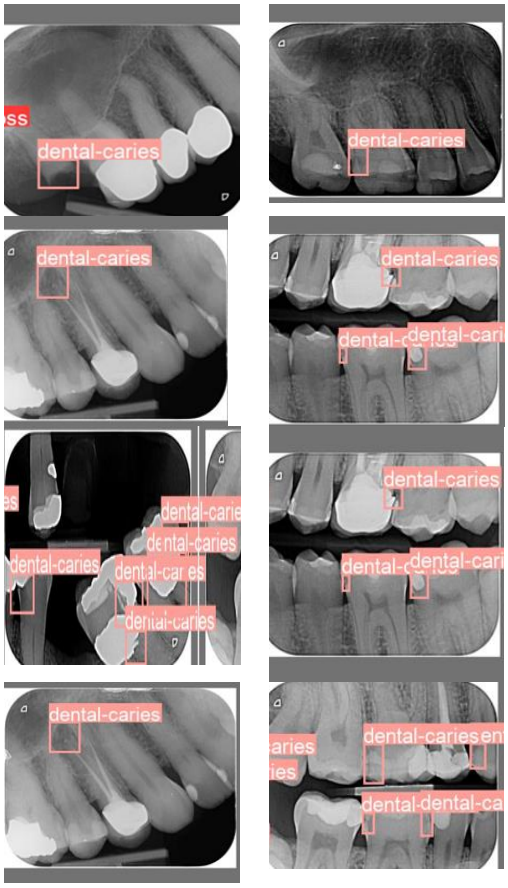

Figure 16: Caries detected X- Ray Images using Faster R-CNN and YOLOv5

Performance evaluations and comparisons of dentists were carried out using the Faster R-CNN and YOLO models (prior to and following revision). To find out if the created deep learning model can help dentists diagnose patients, we used the evaluation dataset, which contained 50 radiographs. We focused on how dentists' diagnoses changed both before and after they referred to the predictions of dental caries made by the deep learning model.

This will be accomplished using the deep learning model, which was previously trained using 400 training images. Thus, there were a total of seven sets of dental caries regions on each radiograph: six were from the dentists and one was from the model. We evaluated the differences between each dentist's first and second diagnoses in order to determine the validity of the model as a diagnostic support system by measuring the Accuracy, Precession, Recall, and F1-score. Based on the severity of the dental caries, 50 radiographs from the evaluation dataset were used to divide the cases into three subgroups (initial, moderate, and extensive). The radiographic presentation of the proximal caries was categorised as follows: Initial: may reach the dentin-enamel junction or the outer third of the dentin; moderate: may reach the middle third of the dentin; and extensive: may reach the inner third of the dentin. Sound: no radiolucency. The same dental caries was regarded as having been detected regardless of the overlap ratio on the

caries lesion if the same area was included. statistical analysis. Accuracy, Precision, Recall, and F1-score were used to calculate the diagnostic performance for readers and the U-Net CNN model.

V. RESULTS

Performance of the YOLO and Faster R-CNN models for diagnostic purposes. The faster R- CNN model's diagnostic performance on Batch 1 of the test dataset. 75% accuracy, 89% precision For dataset Batch-2 Accuracy 73%, Precision 73%, Recall 82%, and F1-score 77% were achieved.

The YOLO model's diagnostic performance on the test dataset Batch-1. For dataset Batch-2, the results were Accuracy 83%, Precision, 80%, Recall, 92%, and F1-score 85%. For dataset Batch-1, the results were Accuracy 88%, Precision, 94%, Recall, 86%, and F1-score 92%.

The deep learning program's performance in detecting dental caries was quite precise and showed a consistent pattern. All types of dental caries, including root caries, secondary dental caries, and gaps under restoration, could be discovered in addition to proximal dental caries. However, the false detection rate of dental caries was slightly higher when the radiograph quality was poor, there was significant dental overlap, and the third molar was present in the X-Ray images. They found more dental caries compared to the CNN model's findings. When the results were analysed based on the level of dental caries, the clinicians tended to be less adept at identifying early caries lesions.

VI. DISCUSSION

In this study, we developed a Faster R- CNN model and YOLOV5 midel for dental caries detection on dental radiographs using U-Net with quite accurate detection performance, confirming that this model could significantly aid clinicians in detecting caries in clinical situations. Dental caries were identified in the evaluation dataset by, and it was found that initial caries had the lowest sensitivity ratio, followed by moderate and extensive caries, respectively. Clinicians were less accurate on bitewing and Periapical radiography than they were for extensive caries, and they were more likely to overlook early caries. However, when clinicians used the results of the deep learning model to make diagnoses, the sensitivity ratio increased in every group of caries severity. This change was especially notable because the initial and moderate groups had statistically significant differences. As a result, the guidance provided by the deep learning model was helpful for helping clinicians spot early caries that might otherwise go unnoticed. Although over-detection of caries increased the false-positive rate and decreased PPV, the Faster R- CNN model and YOLO V5 model helped clinicians diagnose dental caries more accurately and sensitively.

When the model's findings were taken into account, it was more common for the observer to discover new dental caries or to increase the extent of already-existing caries than it was for the observer to get rid of the caries or to reduce the affected area. Therefore, identifying dental caries that had not been discovered initially due to delays or errors was possible by using the results of deep learning as a second opinion. The Faster R-CNN model and YOLOV5model can encourage clinicians to take a second look at the important areas because of its capacity to recognise candidate carious areas that might otherwise go unnoticed in a busy clinical setting, decreasing the likelihood that dental caries won't be discovered in time for treatment. In order to improve diagnostic accuracy and prevent missing the chance to provide preventive treatments, particularly for early dental caries lesions, a second opinion is advised when using bitewing radiography to detect caries.

It should be noted that bitewing and Periapical radiography, as used in this study, might not be sufficient for diagnosing dental caries in images where proximal surfaces overlap or distortion is severe, keeping in mind that it is prone to false-positive and false-negative diagnoses.

VII. CONCLUSION

The findings of the current study suggested that clinicians may be able to more precisely identify early dental caries by using the results of a deep learning model for caries detection as a second opinion. More training data must be added in order to obtain results that are more precise and reliable.

REFERENCES:

1. A. Creanga, H. Geha, V. Sankar, F. Teixeira, C. McMahan, and M. Noujeim, “Accuracy of digital periapical radiography and cone-beam computed tomography in detecting external root resorption,” Imaging science in dentistry, vol. 45, pp. 153–8, Sep. 2015.
2. “What is a Panoramic X-Ray,” Sep. 2017. [Online]. Available: https: //www.minthilldentistry.com/panoramic-x-ray
3. J. Cowton, I. Kyriazakis, and J. Bacardit, “Automated individual pig localisation, tracking and behaviour metric extraction using deep learning,” IEEE Access, vol. 7, pp. 108 049– 108 060, 2019.
4. R. Padilla, S. L. Netto, and E. A. B. da Silva, “A survey on performance metrics for object-detection algorithms,” in 2020 International Conference on Systems, Signals and Image Processing (IWSSIP). IEEE, jul 2020.
5. G. Zhang and H. Li, Effectiveness of Scaled Exponentially-Regularized Linear Units (SERLUs), Jul. 2018.
6. K. Yin, Sign Language Translation with Transformers, Apr. 2020.
7. A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, “An image is worth 16x16 words: Transformers for image recognition at scale,” 2020.
8. M. Tan, R. Pang, and Q. V. Le, “Efficientdet: Scalable and efficient object detection,” CoRR, vol. abs/1911.09070, 2019. [Online]. Available: http://arxiv.org/abs/1911.09070

9. K. Chen, J. Wang, J. Pang, Y. Cao, Y. Xiong, X. Li, S. Sun, W. Feng, Z. Liu, J. Xu, Z. Zhang, D. Cheng, C. Zhu, T. Cheng, Q. Zhao, B. Li, X. Lu, R. Zhu, Y. Wu, J. Dai, J. Wang, J. Shi, W. Ouyang, C. C. Loy, and D. Lin, "Mmdetection: Open mmlab detection toolbox and benchmark," 2019.
10. H. Li, Z. Xu, G. Taylor, C. Studer, and T. Goldstein, "Visualizing the loss landscape of neural nets," 2017.
11. Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), October 2021, pp. 10 012–10 022.
12. T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," 2017.
13. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015.
14. N. Bodla, B. Singh, R. Chellappa, and L. S. Davis, "Soft-nms – improving object detection with one line of code," 2017.
15. S. Lee, S. il Oh, J. Jo, S. Kang, Y. Shin, and J. won Park, "Deep learning for early dental caries detection in bitewing radiographs," Scientific Reports, vol. 11, no. 1, aug 2021.
16. M. M. Srivastava, P. Kumar, L. Pradhan, and S. Varadarajan, "Detection of tooth caries in bitewing radiographs using deep learning," CoRR, vol. abs/1711.07312, 2017. [Online]. Available: http://arxiv.org/abs/1711.07312.
17. T. Yeshua, Y. Mandelbaum, R. Abdalla-Aslan, C. Nadler, L. Cohen, L. Zemour, D. Kabla, O. Gleisner, and I. Leichter, "Automatic detection and classification of dentalrestorations in panoramic radiographs," Issues in Informing Science and Information Technology, vol. 16, pp. 221–234, 2019.
18. R. Abdalla-Aslan, T. Yeshua, D. Kabla, I.Leichter, and C. Nadler, "An artificial intelligence system using machine-learning for automatic detection and classification of dental restorations in panoramic radiography," Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology, vol. 130, no. 5, pp. 593–602, nov 2020.
19. P. Kumar and M. M. Srivastava, "Example mining for incremental learning in medical imaging," 2018.
20. N. J. Kassebaum, E. Bernabé, M. Dahiya, B. Bhandari, C. J. L. Murray, and W. Marcenes, "Global burden of untreated caries: A systematic review and metaregression,"