

AUTOMATED DENTAL CARIES DETECTION

Aayush Sengupta, Ansuman Kanungo, Mehul Singh, Saquib Usman
Student

Department of Computer Science and Engineering,
AMC Engineering College, Bengaluru, Karnataka, India.

Prof. Dr. Nirmala S
Professor

Department of Computer Science and Engineering,
AMC Engineering College, Bengaluru, Karnataka, India.

Abstract: Deep learning techniques are continuously being utilized in many fields. Healthcare is a field in which it can thrive. The study conducted focuses on using artificial models to detect dental cavities in an individual's mouth. These images taken from a camera will be fed live to the object detection model to discover the precise coordinates of dental caries if it happens to exist. Previous studies depict that X-rays were often used for detecting dental caries. This study wants to put emphasis on avoiding the use of X-rays since they have a chance of harming human tissue, as well as, and they cannot detect hidden caries. Thus, it is necessary to detect dental caries in an accurate manner, with the proper tools. Studies have also conducted dental caries prediction using the frontal view of the images only. Some have made use of different angles for the images in the dataset, however, there still lies the problem of capturing the posterior teeth. Roughly 300 images get used, as the dataset, for learning and testing of the object detection model. 80% for learning whereas 20% is for testing. Two process are made to evaluate dental cavities, the You Only Once (YOLO) V5 object detection model and the Faster Region-CNN object determination model. Our results show that the YOLO V5 model consists of an accuracy of 75%, while Faster Region-CNN had an accuracy of 80%. The sensitivity values of YOLO V5 and Faster R- CNN were 76% and 73% respectively. The model with good performance would be used in future development of the product, alongside hardware components.

Keywords: Faster Region-Convolution Neural Network(R-CNN), Convolutional Neural Network (CNN), You Only Once (YOLO) V5

I. INTRODUCTION

Many people often overlook dental care. Without going to the dentist, detecting any caries or issues connected to the teeth is very intricate. Dental caries concludes to tooth pain and tooth loss. It is significant to detect dental caries early

on, in way to prevent it from aggravating. According to WHO (World Health Organization), dental diseases are costly to get treated, they take up about 5-10% of the healthcare budget. The normal procedure for the diagnosis using an x-ray, visual inspection, a dental probe, and handheld mirrors. However, an X-ray is found to be harmful and should be avoided. Even though exposure to radiation is minimum, continuous exposure can have a long-term effect by damaging the molecular structure and causing harm. Another issue with the use of X-rays is that they cannot detect 'hidden carries' (invisible at the surface of teeth) in an accurate manner. Finally, much of this procedure is manual, making it laborious for dentists. Deep learning is a vast topic that has surpassed almost every field. It's a subset of machine learning, where the machines have to learn from the given data themselves, with the minimum human intervention. Deep learning takes inspiration from artificial neural networks. A neural network is described as a paradigm that teaches machines to learn and process information in a manner that is inspired by the human brain. This model will serve the purpose of taking as an input, an image of teeth, and then being able to detect where dental caries are located. We have various object determination models today, such as CNN, Region Based CNN(R-CNN), Fast R-CNN, Faster R-CNN, YOLO V5, Retina-net and etc. Region-based CNN models are the most popular algorithms to use for performing computer vision tasks. This algorithm was developed by Ross Girshick et al. He used the selective search approach to deduce 2000 areas from the specimen for classification. Faster R-CNN is said to be a more advanced version of the R-CNN model, in terms of better accuracy and speed, thus it is an idea that is being considered for use in our study.

II. LITERATURE SURVEY

Elyas Palantei, et al.,[1] Patients health records plays very important role in the medical field environment. The medical data recorded could be used as the doctor references for determining the appropriate medical procedures and provide the advanced prompt treatment information to



improve the patients' health situation. A common best practice in Indonesia that the medical data recording have been regularly stored using an inefficient paper notes systems.

Konstantinos Moutselos et al.,[2] Based on an image dataset of 88 in-vivo dental images taken with an intra oral camera, we show that a Deep Learning model (Mask R-CNN) can detect and classify dental caries on occlusal surfaces across the whole 7-class ICDAS (International Caries Detection and Assessment System) scale.

Yong-Keum Choi et al.,[3] A smart dental health-IoT system based on intelligent hardware, deep learning, and mobile terminal is proposed in this paper, aiming at exploring the feasibility of its application on in-home dental healthcare. Moreover, a smart dental device is designed and developed in this study to perform the image acquisition of teeth.

Jay A. Aldous, Gary Lowder et al.,[4] This study created a web-based application that has an automatic dental caries detection and assessment tool, along with an online storage for patient dental records. The dental caries detection tool enabled the user to upload dental images, and output the severity class level, based on the ICDAS II criteria.

Priyanca P. Gonsalves et al.,[5] SUS and UTAUT questionnaires were used to assess the usability. The SUS resulted with an average standard deviation of 0.7765, that interpreted the respondents' answers to be consistent and a mean score of 76.25 which shows that most of the users found the system usable. The Unified Theory of Acceptance and Use of Technology (UTAUT) raw scores revealed high marks towards the questions under way toward using technology, social influence, and behavioral intention.

Hu Chen, Kailai Zhang et al.,[6]The dentist identifies tooth decay in patients by visually inspecting their teeth and occasionally using dental x-rays. An automated system could assist dentists in detecting cavities by analyzing x-ray images. This paper introduces a model designed to detect cavities in x-ray images using a series of image processing techniques. These include converting the images from RGB to grayscale, generating binary images, identifying the regions of interest, removing background noise, segmenting the image into multiple blocks, and ultimately identifying cavities present in the x-ray image.

Marcelo A. Iruretagoyena et al.,[7] To enhance detection accuracy, we propose three post-processing techniques for faster R-CNN. Firstly, a filtering algorithm removes overlapping boxes assigned to the same tooth. Secondly, a rule-based module uses a teeth numbering system to correct labels violating logical sequencing rules. These techniques refine tooth detection, improving precision and adherence to expected tooth numbering conventions.

Babita Kaushal, Arvind K. Sharma et al.,[8] In dental and medical sciences, IoT has revolutionized disease prediction and analysis by teaming with devices like smartphones and smartwatches. In the past decade, these advancements have

democratized healthcare, enabling both medical professionals and individuals to monitor diseases regularly. Dentistry has particularly evolved into an Internet of Dental Things (IoDT) system, leveraging computerized diagnostics. This transformation has hugely impacted techniques for managing and preventing dental cavities, oral diseases, oral cancers, and other oral health conditions, allowing for more proactive and personal care.

III. LIMITATION IN THE DOMAIN OF DENTAL CARIES DETECTION

Although intraoral photos captured by smartphones aren't typically relied upon for diagnosing dental caries clinically, research indicates that visual inspection through photographs can be quite accurate in detecting them. Dentists' experience plays a significant role in this diagnosis. In our study, we used oral photos taken with smartphones to train deep learning models. However, there were various challenges in capturing these photos, including factors like saliva, poor lighting, varying camera angles, and the presence of soft tissues rather than just teeth, which could affect their quality.

- a. **Lack of Intraoral Datasets:** Even for seasoned dentists, recognizing early-stage caries can be challenging. Photos exhibited sensitivity and specificity rates of just 67% and 79% when used for this purpose, according to a research. We evaluated four deep learning models (YOLOv3, SSD, Retina Net, Faster R-CNN, and Very Early Caries) for early caries detection in our study. Sadly, the sensitivity of YOLOv3 (36.7%), Faster R-CNN (23.4%), Retina Net (26.5%), and SSD (0%), was all rather poor. These examples were trained to concentrate on the size, color, and position of known lesions; however, because early lesions are frequently small and difficult to spot, a larger dataset would be needed to increase accuracy.
- b. **Identifying distinct little objects in a smartphone:** Taking high-quality photos of individual teeth for dental analysis requires expensive equipment like professional cameras and macro lenses, along with skilled photographers. This makes it impractical for regular those who wish to check their teeth with an app. A new investigation by J. Kuhnisch achieved over 90% accuracy in detecting tooth decay using deep learning, but it relied on single-tooth photos taken with professional-grade equipment.
- c. **Poor neural network design:** In this context, "lesser prediction" may refer to instances in which the algorithm is unable to precisely figure out if a caries exists based on the input photos. This might occur because of the intricate link between visual cues and the signs of carries, a lack of training data, or poor neural network design. The solution to this problem might require utilizing larger datasets, investigating various



network topologies, and possibly including new data sources like medical history.

IV. DENTAL CARIES DETECTION TECHNIQUES

Imaging with NIR light at 1310nm has given considerable probability for the detection of early demineralization compared to dental X-rays. The technique was mainly described for proximal caries but it can additionally detect occlusion caries and cracks. Several studies have shown NIR to have higher sensitivity than BW to detect both proximal and occlusion caries and may be used for monitoring. Over the past ten years, NIRT imaging was further examined for multiple indications including primary and secondary caries detection on uneven surfaces, early caries monitoring, caries removal, cavity recognition and monitoring under sealants, and guided caries removal .

They are a subset of artificial intelligence that the workings of the human brain to process data and recognize patterns. These models leverage complex algorithms to learn from large datasets and improve their ability to detect carious lesions with high accuracy. Building an object detector using deep learning models involves training algorithms to recognize specific objects or patterns within images. In dental caries detection, this approach focuses on developing algorithms that can identify and localize carious lesions within dental images. By utilizing deep learning ways such as CNN, DNN, R-CNN, and YOLOv3, researchers aim to create efficient and accurate systems for automated caries detection in clinical settings. Advancements in caries detection ways include visual inspection, X-ray imaging, computer-aided diagnosis systems, digital imaging technology, and smartphone-based imaging with AI algorithms for automated detection.

Time Period	Methodology	Key Advancements
Pre-20th century	Visual inspection by dentists	Early recognition of carious lesions
20th century	X-ray imaging	Radiographic detection of caries Improved accuracy in diagnosis
Late 20th century	Computer-aided diagnosis (CAD) systems	Integration of image processing techniques Automated analysis
Early 21st century	Digital imaging technology	High-resolution images for detailed analysis Improved patient comfort
Present	Smartphone-based imaging coupled with AI algorithms	Accessibility to imaging technology AI-driven analysis for automated detection

Table 1. Detection techniques for Dental Caries evolution over time period.

V. SYSTEM ARCHITECTURE FOR DENTAL CARIES DETECTION

For this, a variety of deep learning techniques can be applied, including SSD and YOLOv3. The two-stage object detection, specifically Faster R-CNN, is chosen as it has

high localization and recognition accuracy. Faster R-CNN model are explained, and the Inception-Resnet-v2 network is used for feature extraction. The Faster R-CNN architecture uses a region-based network to identify bound boxes.

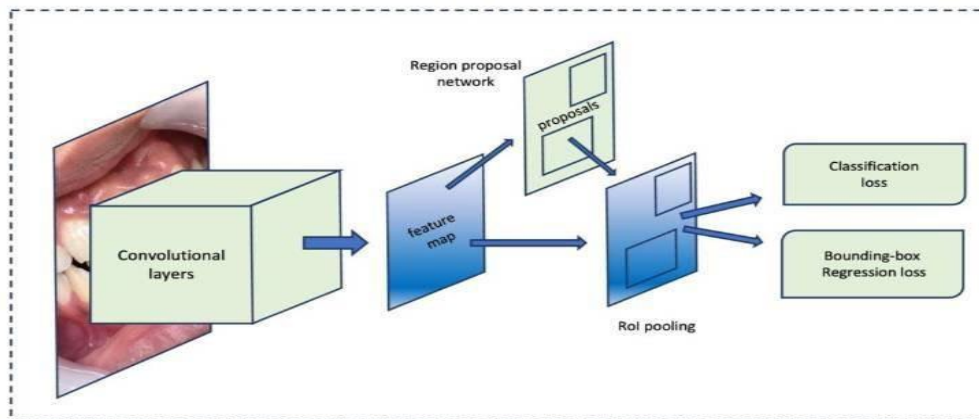


Figure 1: System Architecture for Dental Caries detection

The one-stage detectors, such as the YOLO, SSD, and Retina Net algorithms, are notable for their fast inference times. The YOLO detector series comprises versions 5 and 3, which were debuted in 2015 and 2018, respectively. YOLOv3 employs a single neural network to predict bounding boxes and item labels from an image. It also uses the "dimensional clustering proposal" approach to determine bounding boxes. Object detection methods employing SSD

and Retina Net. SSD creates bounding boxes on a feature map and employs non-maximum suppression to get the desired output. Using a focused loss function, RetinaNet attempts to distinguish between front and rear classes. Both models were implemented in Python using the Pytorch backend. The final output label for each photo included all diagnoses without detection of lesion location.

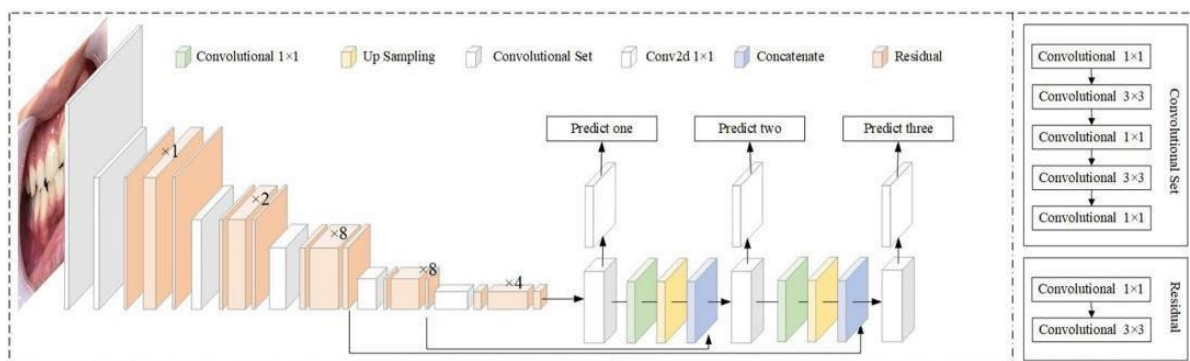


Figure 2: YOLO V3 Architecture

VI. COMPARATIVE ANALYSIS

Table 2: An implementation of the Dental Caries detection technology.

Ref No.	Framework	Type	Key Features	Pros	Cons
[1]	CNN	Convolutional Neural Network	Hierarchical feature extraction Suitable for image recognition	Effective for feature learning Robust to spatial variance	Requires large datasets for training Computationally intensive
[2]	DNN	Deep Neural Network	Multiple hidden layers Suitable for complex data	Versatile, applicable to various tasks Can learn intricate patterns	Prone to overfitting Training can be slow and resource-intensive

[3]	R-CNN	Region-based CNN	Region proposal network Two-stage detection process	High detection accuracy Handles objects of different scales	Slow inference speed High computational cost
[4]	Fast R-CNN	Faster Region-based CNN	Single-stage detection with region of interest pooling	Faster inference compared to R- CNN End-to-end training	Still slower than subsequent models Complex architecture
[5]	Faster R- CNN	Faster Region-based CNN	Region proposal network integrated into the model	Improved speed compared to R- CNN and Fast R- CNN Precise localization	Requires considerable computational resources Training can be complex
[6]	Mask R- CNN	Mask Region-based CNN	Extends Faster R- CNN with instance	Accurate instance segmentation	High computational cost More complex architecture
[7]	YOLOv3	You Only Look Once v3	Single-stage object detection Divides image into grid cells	Very fast inference speed Good for real-time applications	Less accurate localization Struggles with small objects
[8]	RetinaNet	Single Detector	Shot address imbalance Feature pyramid network	Focal loss to effectively address class imbalance in detection	High accuracy Slower inference compared to YOLOv3 Complexity in implementation
[9]	SSD	Single Detector	Shot Single-stage object detection Multi-scale feature maps	Good balance between speed and accuracy Handles objects of various sizes	May struggle with small objects Less accurate than some other models

VII. ALGORITHMS AND FRAMEWORKS USED

Dental caries, a prevalent chronic illness caused by bacteria, affects individuals of all ages . Restorative fillings are commonly used for treating dental cavities, but these treatments often fail. Researchers propose individualized caries treatment based on caries risk assessment, highlighting the need for personalized care . Deep learning techniques have given good results in diagnosing caries in radiology, prompting the study to explore the detection of caries lesions using deep learning algorithms on dental panoramic films.

a. YOLO V3: You Only Look Once Based Detection Framework: It is a popular object detection model known for its speed and accuracy. YOLOv3 is a widely used object detection framework, including in medical

imaging like dental caries detection, offering improvements in speed, accuracy, and usability over its predecessors. For dental caries detection with YOLOv3, steps involve dataset preparation with annotated images, model training using transfer learning, evaluation on a validation set, inference on new images, and post-processing for refining results. The model's performance in dental caries detection depends on factors like dataset quality, hyper-parameters, and addressing challenges like lighting variations during training and inference.

Intraoral images are preprocessed for quality enhancement before being annotated with bounding boxes indicating dental caries areas. YOLOv3 is trained on this annotated dataset using transfer learning with per-trained weights for

specific caries detection tasks. YOLOv3 architecture uses CNN to predict bounding boxes and class probabilities at different scales for detecting objects of varying sizes. Post-processing techniques like NMS are applied to refine detection's, and the model's performance is evaluated on a validation dataset before deployment for dental caries detection.

This algorithm is considered fast, because of the base it uses for feature extraction. The base is known as Darknet-53. This architecture is made up of 24 convolutional layers and 2 fully connected layers. The 20 convolutional layers are met by a pooling layer as well as a fully connected layer. This base has already been trained on an ImageNet dataset. The layers consist of 3 x 3 convolutional layers and 1 x 1 reduction layers.

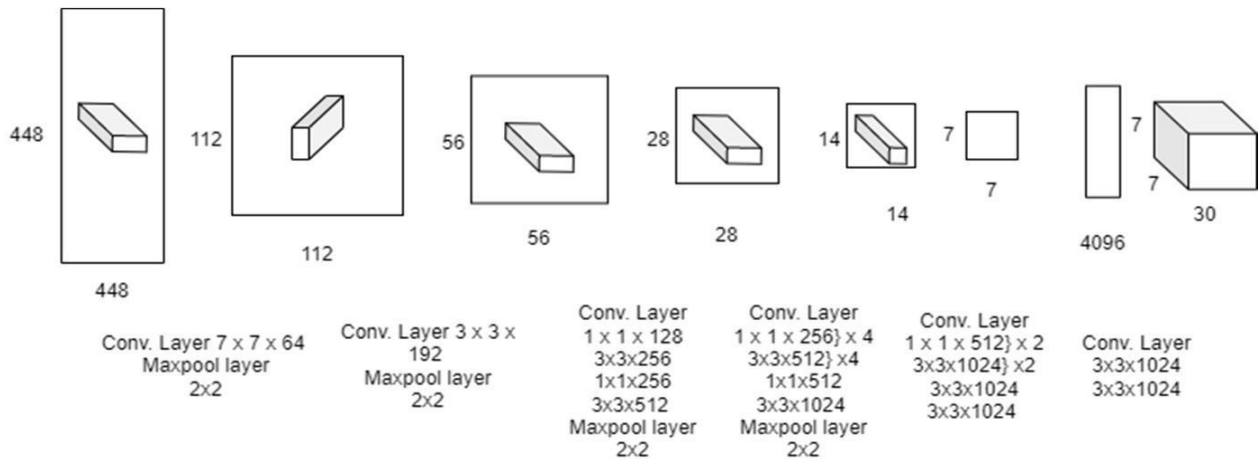


Figure 3: YOLO V3 base architecture

b. CNN : Faster R-CNN: Convolutional Layers in a CNN consist of Base and Head components for feature retrieval and classification. Feature extraction involves filtering, ReLu activation, and pooling to create feature maps. Convolution layers train filters to extract significant features like texture and shape from images. Region Proposal Network (RPN) in CNN verifies object presence and predicts bounding boxes. Fully connected neural network classifies objects and uses

Support Vector Machine classifier to draw bounding boxes.

Convolution layers serve the purpose of training filters to retrieve the significant features from the image, based on the object present in the image. Features could include texture, shape, edges, etc. For example, if one wishes to retrieve the important features of an animal, then the filters will be trained accordingly to learn the shapes, colors, and other features of that animal.

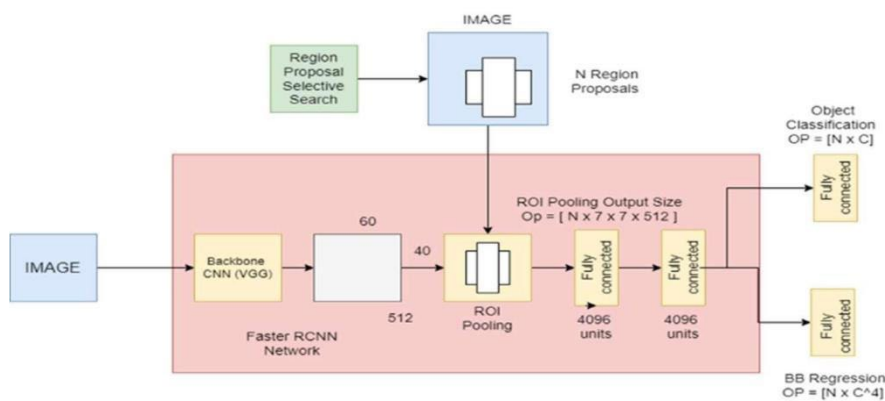


Fig. 4: Architecture of Faster R-CNN network using a backbone of Visual Geometry Group (VGG) with 16 layers

The architecture consists of layers, which extracts from intraoral images, such as texture, shape, and edges. The base

layer extracts visual features, while the head classifies these features and performs object detection tasks. The ReLU

activation function introduces non-linearity, while Max/Average Pooling reduces the dimensionality of feature maps. The RPN verifies the presence of its objects and proposes bounding boxes, while the Fully Connected Neural Network performs object classification and regression using a Vector classifier. The system is designed to enhance computational efficiency while preserving important features.

SSD (Single-Shot Detector) : Single-Shot Detector (SSD) is a type of object recognition algorithm used in image processing. It is made to classify objects out of images or videos in the neural network. The term "single-shot" in SSD classifies that it can detect objects in a single inference, making it faster and more efficient compared to other object recognition algorithms that may require multiple journey through the network. SSD achieves this by using a CNN to

predict the presence of multiple objects in an image at different scales. This is in contrast to two-stage detectors like Faster R-CNN, which have separate region proposal and object detection stages. In SSD, the network is divided into multiple convolutional layers that predict the class labels and bounding boxes for objects at different scales. These predictions are made at multiple feature maps of the network to capture objects of various sizes. By predicting objects at different scales within a single network, SSD is able to detect in proportional objects efficiently. This makes it particularly useful for real-time applications where speed and accuracy are crucial. SSD is known for its balance between speed and precision in object detection tasks. It has been widely used in applications such as autonomous driving, surveillance, and image recognition due to its ability to quickly and accurately detect objects in complex scenes.

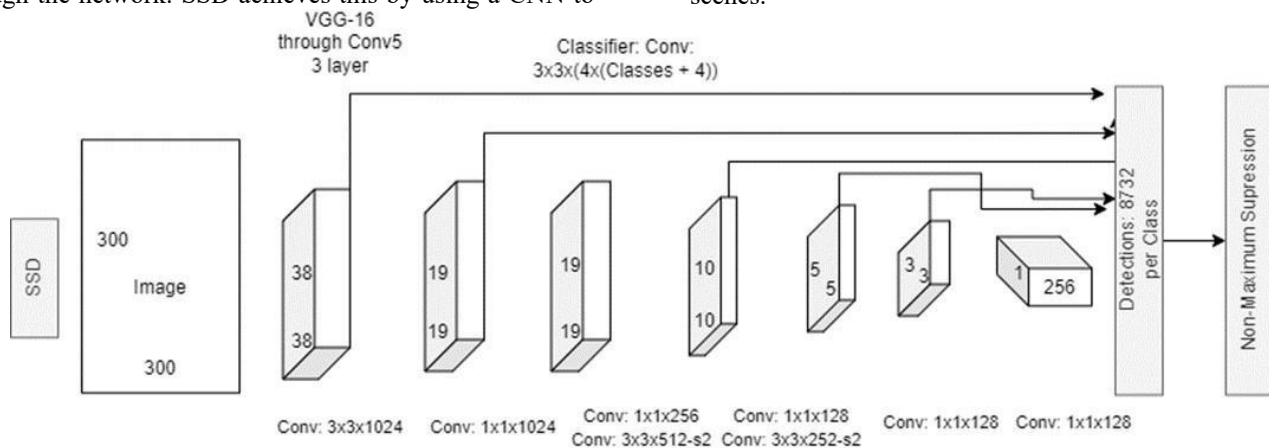


Fig 5: Single Shot Multi-box Detector network architecture multiple convolutional layers

c. RetinaNet: RetinaNet is a small object detection system developed by the Meta Artificial Intelligence research team. It consists of three parts: a backbone network, a sub-network for object classification, and a sub-network for object regression. The backbone network is divided into two components: the bottom-up pathway and the top-down pathway with lateral connections. The bottom-up pathway develops feature maps at various dimensions, while the top-down approach selects spatially sparse feature maps for higher pyramid levels. The sub-network for object classification attaches a fully convolutional network to each Feature Pyramid Network. Sigmoid activation is used for object classification. The sub-network for

object regression attaches a regression network to the Feature Pyramid Network, with a 3x3 layer with four filters. The regression sub-network creates four amounts for each anchor box, estimating difference between anchor box and the desired box. The architecture has four major components: the bottom-up pathway, which computes feature maps at multiple scales; the top-down pathway, which unsamples spaced-out coarser feature maps from pyramids at higher levels; the classification sub-network, which predicts the presence of object in a spatial location; and the regression sub-network, which reverts the distance for bounding boxes from the anchor boxes for each ground-truth item.

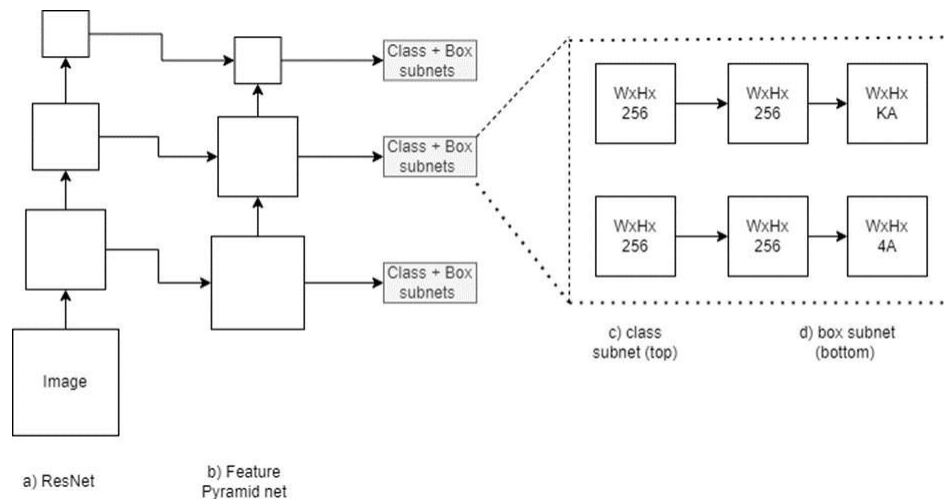


Fig 6: RetinaNet architecture

There are four major components in architecture : (a) Bottom-up pathway: Acts as the supporting backbone network which computes feature maps at multiple scales (b) Top-down pathway: This unsamples the spaced-out coarser feature maps from the pyramids at higher levels (c)

Classification sub-network: Making a prediction regarding the presence of object in a spatial location (d) Regression Sub-network: For each ground-truth item, it reverts the distance for the bounding boxes from the anchor boxes.

Table 3: Model Evaluation Results of C vs NC Classification.

Model Deep Learning	Sensitivity %	Specificity %	Accuracy %	Precision %
YOLOv3	74	86.6	83.4	65.3
Faster R-CNN	71.2	92.9	87.4	77.3
RetinaNet	63.2	89.8	83	67.7
SSD	26	99.7	81	97.1

C: cavitated; NSC: no surface change; VNC: visually non-cavitated.

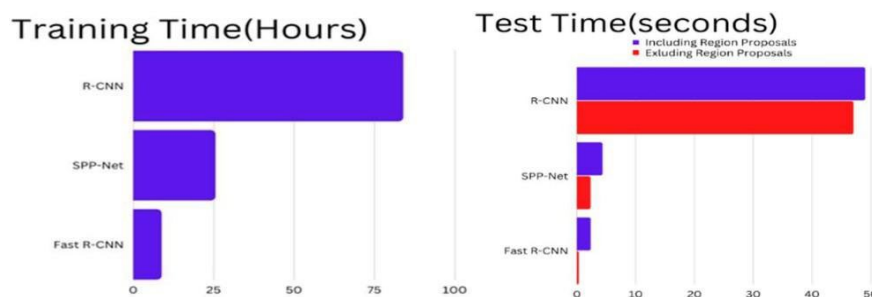


Fig 7: Graphical illustration of (a) training time for different deep learning models (b) testing time for different deep learning models

From the experiments conducted by the authors in depicts the comparison between the different models. Of those four models, Faster R-CNN had the highest accuracy of 87.4%, followed by YOLO V3, having an accuracy of 83.4%.Faster R-CNN's accuracy was recorded at 80%, slightly higher than the YOLO V3 algorithm's

accuracy of 75%. The responsiveness of the YOLO v3 algorithm was 76%, whereas the responsiveness of the Faster R-CNN was recorded as 73%. The precision of Faster R-CNN was higher, 78%, as compared to that of YOLO v3, 74%. Most of the performance metrics favoured the Faster R-CNN algorithm.

Figure 10 illustrated the live working of the Faster R-CNN algorithm to locate dental caries on one of the testing.

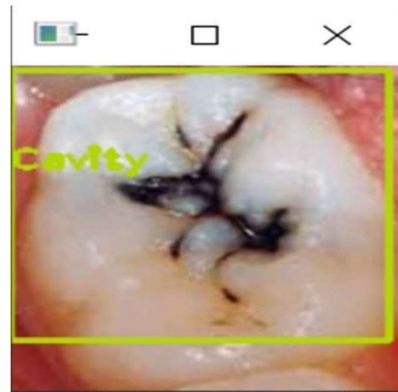


Fig 10: Dental caries detection using YOLO V3

Table 4: Comparison between the two proposed models using four different performance metrics

Methods	Accuracy	Specificity	Sensitivity	Precision
YOLO V3	75%	72%	76%	74%
Faster R-CNN	80%	78%	73%	78%

VIII. CONCLUSION AND FUTURE SCOPE

It seems like the study compared the effectiveness of Faster R-CNN and YOLOv3 algorithms for caries detection using a dataset consisting of approximately 300 images, with 250 images used to training and 50 for testing. The results indicated that Faster R-CNN outperformed YOLOv3 in most parameters, achieving higher accuracy scores. However, the accuracy achieved by both algorithms in this study was roughly similar to those reported in previous research. Two algorithms, Faster R-CNN and YOLO v3, were used to detect dental caries in a dataset of approximately 300 images. The Faster R-CNN performed better in most parameters compared to the YOLO v3 algorithm. The authors' algorithms achieved similar accuracy rates, with the Faster R-CNN model achieving an efficiency of 80%. This study suggests that future datasets could be expanded to include 1000 images from various dental hospitals and a larger variety of datasets to improve results. The study focuses on training machine algorithms to differentiate between different levels of dental caries in individuals. Future research could explore algorithms like Retinanet and Single Shot Detector to explore their potential applications. The study emphasizes the importance of expanding the range of data and exploring other algorithms for better results.

REFERENCES

- [1]. Elyas Palantei, Dewiani, Asrul Ramadhan, and Sigit Lukman, A smart card-based campus dental clinic services: experimental tests. In 2019 IEEE International Conference on Communication, Networks and Satellite (Comnetsat), August 1 - 3, 2019, Makassar, Indonesia. IEEE, New York, NY, 9-13.
- [2]. Konstantinos Moutselos, Elias Berdouses, Constantine Oulis, and Ilias Maglogiannis. 2019. Recognizing occlusal caries in dental intraoral images using deep learning. In 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), July 27, 2019, Berlin, Germany. IEEE, New York, NY, 16171620.
- [3]. Lizheng Liu, Jiawei Xu, Yuxiang Huan, Zhuo Zou, Shih-Ching Yeh, and Li-Rong Zheng. 2020. A smart dental health-IoT platform based on intelligent hardware, deep learning, and mobile terminal. IEEE Journal of Biomedical and Health Informatics, 24, 3, 898-906
- [4]. Yong-Keum Choi and Kyung-Oh Lee. 2018. A research on the developing smartphone application or automatic dental caries detection. Journal of Engineering and Applied Sciences, 13, 24, 10468-10473.



- [5]. Jay A. Aldous, Gary Lowder, and Jerald Boseman. 2017. Evaluation of a caries detecting software system. ODOVTOS-International Journal of Dental Sciences, 193, 61-67.
- [6]. Priyanca P. Gonsalves. 2017. Diagnosis of dental cavities using image processing. International Journal of Computer Applications, 180, 5, 28-32.
- [7]. Lee, J. H., Kim, D. H., Jeong, S. N., & Choi, S. H. (2018). Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. Journal of dentistry, 77, 106-111.
- [8]. Liu, C., Tao, Y., Liang, J., Li, K., & Chen, Y. (2018, December). Object detection based on YOLO network. In 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC) (pp. 799-803). IEEE.