DOI: 10.5937/bjdm2402099T

BALKAN JOURNAL OF DENTAL MEDICINE

ISSN 2335-0245

Automatic Detection of Degenerative Changes in the Temporomandibular Joint Region Using Deep Learning with Panoramic Radiographs

SUMMARY

Background/Aim: *The temporomandibular joint (TMJ) is a complex anatomical region composed of the mandibular condyle located in the glenoid fossa of the temporal bone and covered with fibrous connective tissue. Excessive and continuous forces lead to progressive degeneration of the bony surfaces of the TMJ. The aim of this study is to determine the success of automatic detection of degenerative changes detected on panoramic radiographs in the TMJ region with deep learning method.* ***Material and Methods:*** *Panoramic images of 1068 patients (1000 with normal TMJ appearance and 68 with TMJ degeneration) over 18 years of age were included in the study. CVAT, open-source annotation tool (https://www.cvat.ai/) was used for labeling image data. All images were resized using the bilinear interpolation method. With the using data augmentation techniques, the number of images data reached 1480. BSRGAN model was applied to the data to increase the resolution of the data. YOLOv5, YOLOv7 and YOLOv8 algorithms were used for TMJ degeneration detection. TP, FP, TN, FN, accuracy, precision, recall, F1-score and AUC (Area Under the Curve) metrics were used for statistical analysis.* ***Results:*** *YOLOv5s training resulted in 94.40% accuracy, 81.63% precision, 86.96% sensitivity, 84.21% F1 score and 91.45% AUC. YOLOv7 training resulted in 99.63% accuracy, 97.87% precision, 100% sensitivity, 98.92% F1 Score and 99.77% AUC. YOLOv8 training resulted 96.64% accuracy, 91.11% precision, 89.13% sensitivity, 90.11% F1 Score and 93.66% AUC.* ***Conclusions:*** *All three algorithms have high success rates, with the best results obtained in YOLOv7.*

Keywords: Deep Learning, Degeneration, Panoramic Radiography, Temporomandibular Joint

Melek Tassoker[1], Huseyin Hakli[2], Metin Furkan Yaman[2], Sema Nur Ekmekcı[2], Senanur Incekara[2], Serhat Kamaci[2], Busra Ozturk[1]

1Department of Oral and Maxillofacial Radiology, Necmettin Erbakan University, Faculty of Dentistry, Konya, Türkiye
2 Department of Computer Engineering, Necmettin Erbakan University, Konya, Türkiye

ORIGINAL PAPER (OP)
Balk J Dent Med, 2024;99-116

Introduction

The temporomandibular joint (TMJ) is a complex anatomical region composed of the mandibular condyle located in the glenoid fossa of the temporal bone and covered with fibrous connective tissue. It is classically considered to be a condylar joint, connected to the skull on both sides, moving together, capable of complex rotational and translational movements, and is the only joint that does not function on the other side without movement in the contralateral joint. Due to these structural features, it is quite different and complex from other joints in the human body[1].

Excessive and continuous forces lead to progressive degeneration of the bony surfaces of the TMJ, which eventually results in radiographic signs of osteoarthritis such as flattening, formation of osteophytes, erosion, reduction in joint space, and subcortical sclerosis. Once degenerative changes begin in the TMJ, this pathology can be progressive and lead to various morphologic and functional deformities[2]. Osteoarthritis is diagnosed by clinical examination of TMJ symptoms, detection of degenerative changes on radiography, or a combination of both[3]. Radiographic findings provide important information about the presence and severity of bone changes in the TMJ. Since some of these changes may be

part of the age-related remodeling process or related to a physiologic response, they may occur in asymptomatic patients. Therefore, it is not always possible to correlate radiographic data with clinical status[4].

Although computed tomography (CT) is considered to be the gold standard for detecting morphologic changes in the TMJ region, panoramic radiography is a useful imaging method for TMJ examinations in dental practice due to its considerably lower radiation dose compared to CT and acceptable observer compliance[5]. In panoramic radiography, the lateral and central parts of the condyle can be observed because the rays are oblique to the long axis of the condyle. Since the skull base and zygomatic arch are superimposed on the articular tubercle and fossa, only obvious structural changes can be diagnosed. Significant erosion, sclerosis, osteophytes, asymmetry of the condyles, fractures and large deformations can be identified[6]. As a result of its advantages such as low radiation dose, low cost and widespread use, panoramic radiography is a clinician-assisted imaging modality in routine use in the examination of such degenerative changes[7].

Panoramic radiographs are the most commonly used images in dentistry routine and the first evaluation of the jaw and teeth. Changes in the bone tissue in the TMJ region can be examined with panoramic radiographs; however, since the TMJ region is surrounded by the bones of the skull, it may be difficult for non-specialists in the field of dentomaxillofacial radiology to recognize the changes in two-dimensional images; this may lead to diagnostic inadequacies and delays in treatment[8]. Due to its complex structure, there may be inter-observer differences in TMJ examinations[9,10]. It will be possible to overcome this problem with trained artificial intelligence models. The aim of this study is to determine the success of automatic detection of degenerative changes detected on panoramic radiographs in the TMJ region with deep learning method.

Material and Methods

This study was conducted by retrospectively reviewing the panoramic radiography records of individuals who applied to Necmettin Erbakan University Faculty of Dentistry, Department of Oral, Dental and Maxillofacial Radiology for dental examination between January 2020 and March 2023. The ethics committee approval for the study was obtained from Necmettin Erbakan University, Faculty of Dentistry, Drug and Non-Medical Device Research Ethics Committee with decision number 2023/274.

Sample

- No pathologies related to bone and cartilage tissue (tumor, cyst, osteomyelitis, bone dysplasia) in the right-left temporomandibular joint (TMJ) region,
- No evidence of trauma (fracture line),
- Artifact-free images that are diagnostically adequate for examination of the right-left TMJ region

Panoramic images of 1068 (patients with normal TMJ appearance, patients with TMJ degeneration) individuals over 18 years of age who provided the specified criteria were included.

Radiological evaluation

Degenerative changes in the TMJ condyle structure are defined in Figure 1: erosion (loss of cortical density), sclerosis (increase in cortical density), osteophytes (bone process anterior to the condyle), flattening (loss of condyle convexity) and other types of degeneration (e.g. subchondral cyst, cavity due to degeneration).

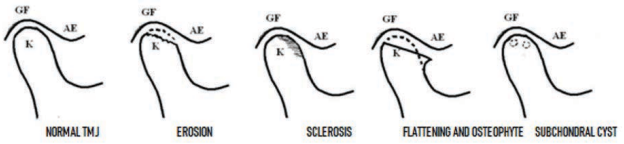


Figure 1. Findings of healthy TMJ and degenerative changes [10]

The TMJ will be labeled as 'degenerated' if any of the findings of erosion, sclerosis, flattening, osteophytes, subchondral cysts are detected in the digital panoramic images included in the study. In Figure 2, the TMJ region with erosion in the right and left TMJ region is labeled as 'degenerated'. Figure 3 shows an example labeling process of a patient with healthy TMJ condyle morphology. CVAT, an open-source, web-based tool (https://www.cvat.ai/) was used for labeling image data. Interobserver agreement was assessed by labeling 20% of the data by two different observers; all labeling was performed by an observer (MT) with 12 years of experience in oral radiology.

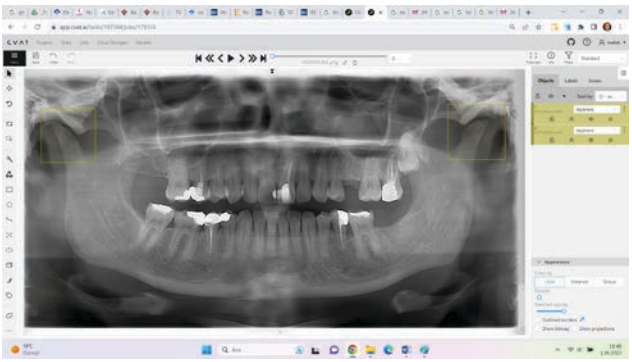


Figure 2. Labeling of TMJ regions with erosion in the right and left TMJ regions

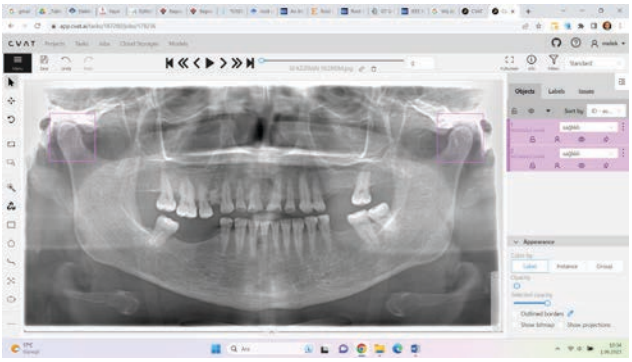


Figure 3. Labeling of the right-left healthy TMJ condyle region

Images obtained with the MORITA (2D Veraviewepocs®, MFG Corp., Kyoto, Japan) panoramic X-ray machine were used in the study.

2.1 Data Augmentation

In order to perform the detection process with deep learning algorithms, the radiological images must be in the same format. All the images in different formats (*.png, *tif and *jpg) were converted to *.jpg format.

The panoramic radiographs collected for use in the study have different row and column sizes. In order for the algorithms to make appropriate and accurate evaluations, all images were resized using the bilinear interpolation method so that the row and column sizes were the same.

A dataset is considered unbalanced when the data samples are not evenly distributed among different classes/categories, which is quite common in real-world datasets. In such a dataset, the class with more data samples is defined as the majority class, while those with far fewer data samples are called minority classes. Most classifiers are modeled by exploring data statistics, so majority classes can be correctly biased classifiers and hence show very low classification accuracy in minority classes[11]. Of the total 1068 data analyzed in the study, 1000 of them are healthy individuals, while the remaining 68 data are ill individuals. The distribution of healthy individuals over ill individuals is unbalanced. This problem can lead to bias, overlearning and erroneous results in datasets. Data augmentation techniques overcome this negative situation. There are several data augmentation techniques in the literature. Studies have found rotate, mirror, mirror-rotate, and shift as data augmentation techniques and these techniques were used in our study[12]. Below are the data augmentation methods used in this study (Figure 4).

Rotation: The process of rotating the data 180° counterclockwise.

Mirror: The reflection image of the data in a mirror.

Mirror-Rotate: The process of rotating the reflected data 180° counterclockwise.

Shift Method: It is the method of shifting the data in +-Y directions. The previous coordinate where the data is located is changed with this method.

Mirror-Rotate and Rotate data were used only in the training set. In line with the results obtained using data augmentation techniques, the number of patient data reached 480. Our total data set used in the study was increased to 1480 data.

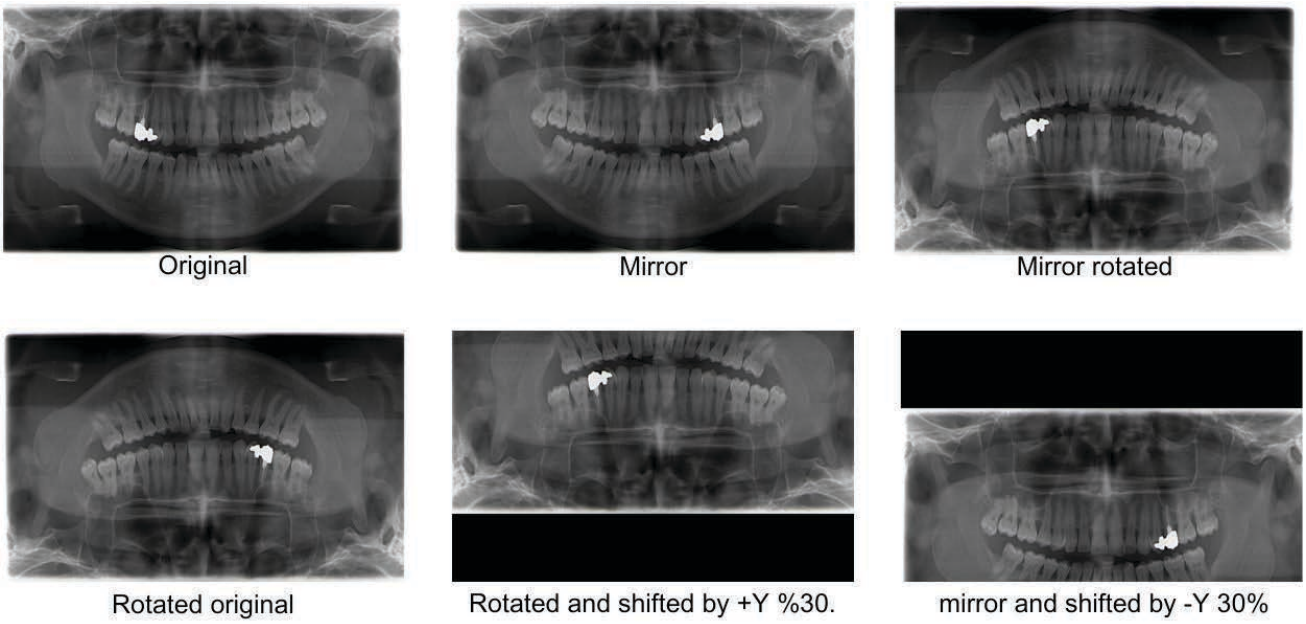


Figure 4. The data augmentation methods used

Image pre-processing

When detecting degenerative changes in the TMJ region, deformities on the medial and superior surfaces of the condyle head are examined. In order to increase the success rate of detection, morphologic transformations have been performed to remove noise, enlarge the shape or perform gradient operations in this region.

First, gray level image transformation was applied to the images and then morphological operations were performed and the morphological operation method that gives the best result according to the deep learning algorithm was selected[13]. The morphological transformations (Figure 5) used are given below[14].

Erosion: It removes the noise around the object by performing a comparison process with the kernel structure it uses, and at the same time, it causes the size of the object to decrease while performing this process.

Dilation: Increases the size of the white region or foreground object in the image.

Opening: Dilation, which is an expansion method, is applied to the image that has been eroded first, creating an image that is noise-free and increases our object area.

Closing: Expansion followed by erosion. It is useful for covering small holes in foreground objects or small black spots on the object.

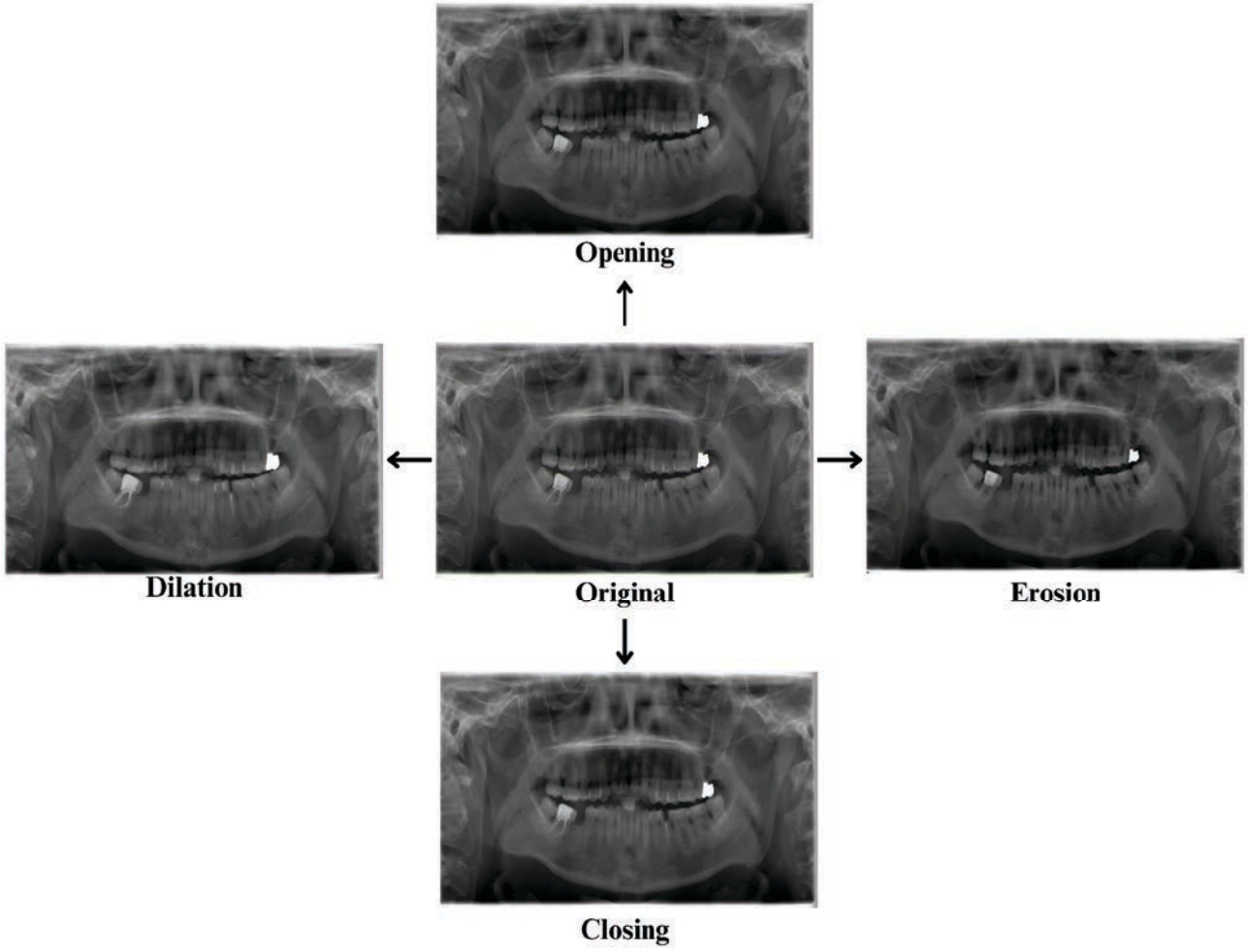


Figure 5. Morphological transformation methods used

After these processes are completed, the BSRGAN (Bilateral Self-Attention Generative Adversarial Network) model is applied to the data to increase the resolution of the existing data. All data were resized to prevent the negative effects of size inequality and size increase on model training. In this process, bilinear interpolation method was used to preserve the proportional integrity of row and column dimensions. Bilinear interpolation uses only the 4 closest pixel values in diagonal directions from a given pixel to find the appropriate color intensity values of the pixel. It then takes a weighted average of these 4 pixels to arrive at the final interpolation value. The weight of each of the 4 pixel values is based on its distance from each of the known points of the calculated pixel[15].

Blind image super-resolution (BSRGAN) is a super-resolution method that takes a low-resolution version of an image and tries to estimate a high-resolution version based on this low-resolution input. The goal of BSRGAN is to solve the problem of general-purpose blind image super-resolution. It also allows to remove blurred images of an image depending on its Laplace variance[16]. As a result of the BSRGAN method, bilinear interpolation method was applied again to the data whose dimensions changed and the dimensions were arranged in 1024x640 format.

YOLO Deep Learning Models

The YOLO algorithm used in this study is a convolutional neural network-based single-stage deep learning model that has been widely applied in the literature to detect and localize an object in an image. The YOLO algorithm was first introduced in 2016, followed by YOLOv2, YOLOv3, YOLOv4 and YOLOv5 [17-20].

In the deep learning studies conducted in the literature for the TMJ region, the diagnostic success of different algorithms such as ResNet, VGG16, InceptionV3, EfficientNet-B7 has been investigated, but the success of the YOLO algorithm, which stands out with its ability to detect objects at high speed, in the detection of degenerative changes in the TMJ region is effectively investigated in this study. In addition to being fast, YOLO does not require a complex data pipeline and has high accuracy[19].

Two-stage algorithms (Mask-RCNN, R-CNN and Faster R-CNN) are generally more accurate, but slower than single-stage (YOLO) algorithms[21]. Considering the busy working conditions of dentists, there is a need for solutions that offer a combination of speed and performance. Therefore, this study investigates the success of the YOLO algorithm in detecting degenerative changes in the TMJ region.

In the YOLO algorithm, a single convolutional network simultaneously predicts multiple bounding boxes and class probabilities for these boxes. This unified model has several benefits over traditional object detection methods. First, YOLO is extremely fast. Since we frame the object to be detected as a regression problem, there is no need for a complex data pipeline. Second, YOLO reason universally about the image when making predictions. Third, YOLO learns generalizable representations of objects[22].

The aim of the research is to compare YOLOv5, YOLOv7 and YOLOv8 algorithms to determine the most effective deep learning algorithm and the most suitable morphological transformation for this dataset.

YOLOv5

YOLOv5 is compiled under a new training environment in PyTorch, making the training process easier compared to Darknet. This has significantly improved the detection accuracy and speed of the model compared to previous versions [23].

YOLOv5 consists of four versions, YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. This is classified according to memory storage size, but the principle is the same. YOLOv5x has the largest storage size and YOLOv5s has the smallest storage size[24]. It has been reported that YOLO v5 achieves higher accuracy values than YOLOv3 and YOLO v4 [25].

The structure of YOLOv5, whose general structure is given in Figure 6, consists of three main parts. First, the backbone forms the main body of the network. The backbone for YOLOv5 is designed with the New CSP-Darknet53 structure, which is a modification of the Darknet architecture used in previous versions. The CSPNet module solves recurring gradient problems in large-scale backbones and reduces model parameters by integrating gradient changes into the feature map. This process increases the inference speed and accuracy and reduces the model size[23]. Next comes the neck section, which connects the backbone to the hand. YOLOv5, SPPF and New CSP-PAN structures are used here. The purpose of SPPF is to provide a multi-scale representation of the input feature maps. By pooling at different scales, SPPF allows the model to capture features at different levels of abstraction[26]. PANet improves the use of accurate localization signals in the lower layers, which can improve object location accuracy. PANet is used in YOLOv5 to preserve feature levels[23]. Finally, the head section is responsible for producing the final output. In this part, YOLOv5 uses the YOLOV3 Head structure[27].

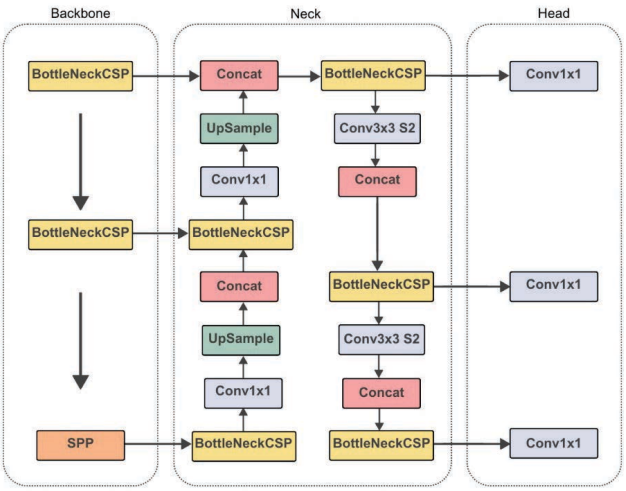


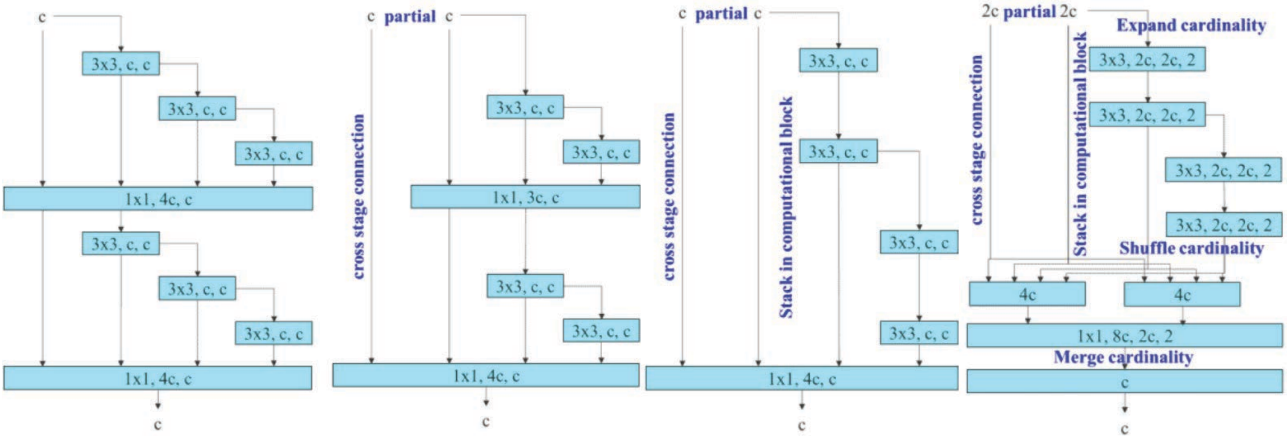
Figure 6. General structure of the YOLOv5 architecture 23

YOLOv7

The basic structure of the YOLOv7 algorithm model bears great similarities to its predecessor, YOLOv5. This structure is usually divided into four basic components:

Input, Backbone, Neck and Head[28]. YOLOv7 uses a concept known as " trainable bag-of-freebies", which includes modules and optimization methods designed to improve object detection accuracy[27].

The backbone framework used by YOLOv7 makes use of the E-ELAN module, which has a final stack module containing four branches (Figure 7). This configuration allows for a denser residual structure resulting from the use of multiple stacks, which allows for increased accuracy by increasing network depth while increasing simplicity in optimization[28].

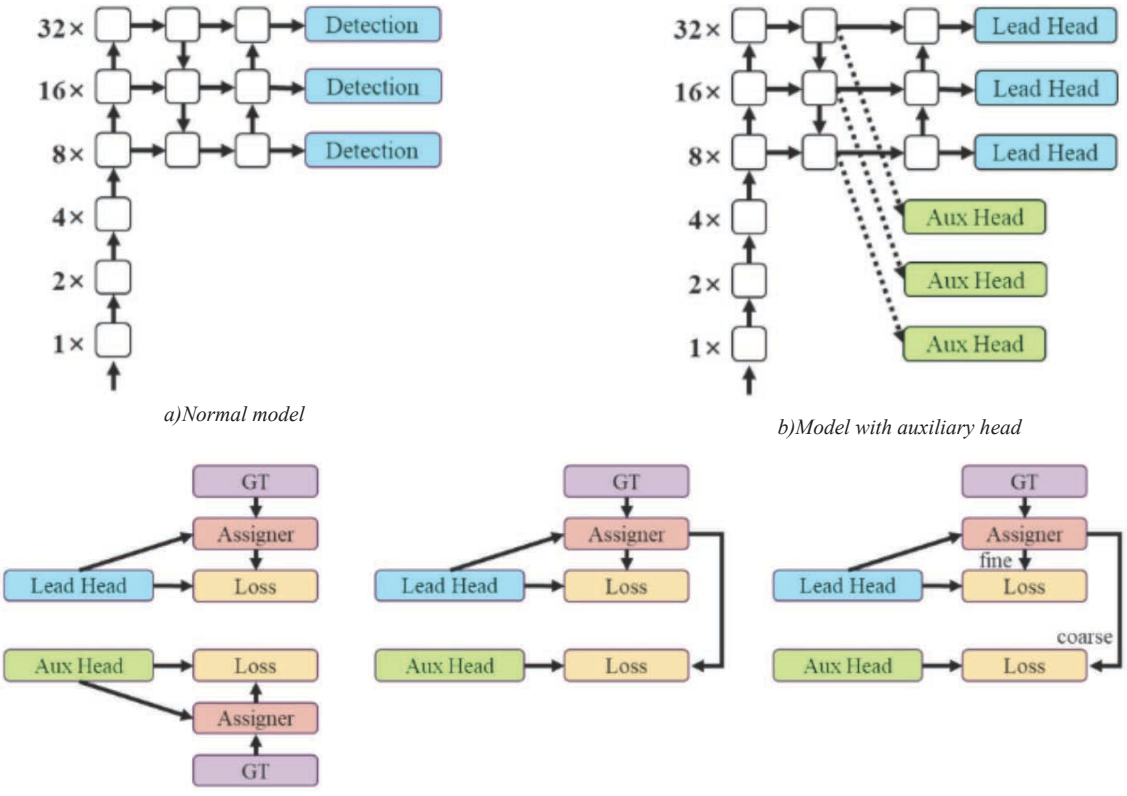


a)VovNet *b)CSPVovNet* *c)ELAN* *d)E-ELAN*

Figure 7. The evolution of layer aggregation strategies in YOLOv7 29

The neck structure retains the Panet structure inherited from the previous series. This not only increases the original features of the architecture, but also includes an additional round of feature sampling to realize a more comprehensive feature fusion [28].

In the head structure, it uses the "Coarse-to-fine lead head guided label assigner" structure (Figures 8 and 9). Coarse-to-fine lead head guided label assigner generates a soft label using the result predicted by the lead head and reality [29].



a)Normal model *b)Model with auxiliary head*

c)Independent assigner *d)Lead guided assigne* *e) Coarse-to-fine lead Guided assigner*

Figure 8. Coarse-to-fine auxiliary head supervision in the YOLOv7 network [30]

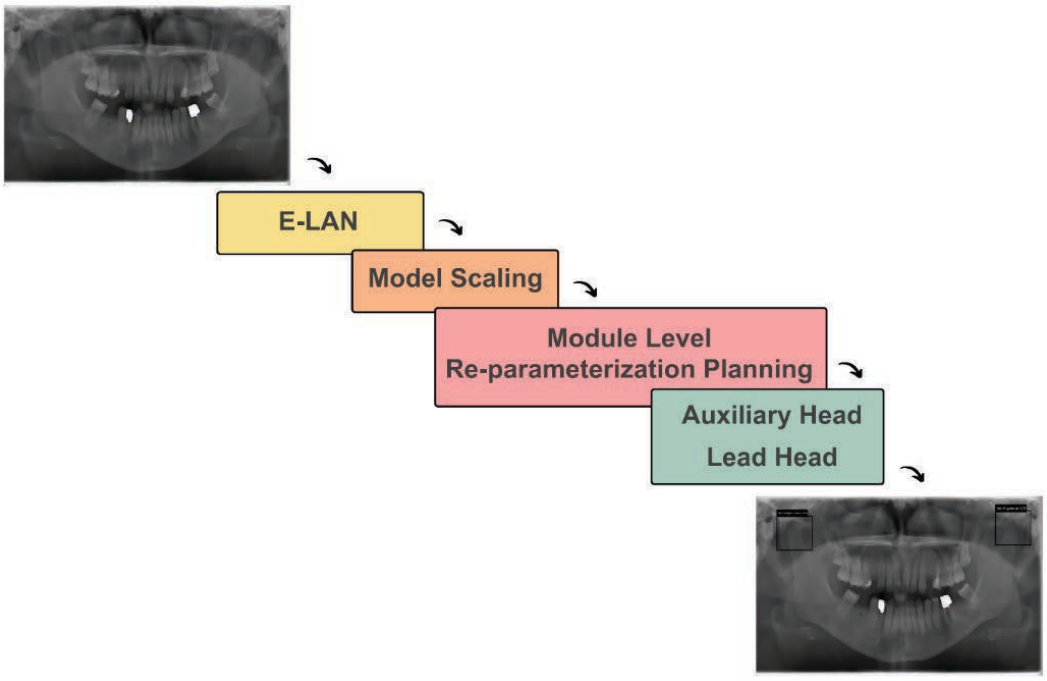


Figure 9. Architecture of the YOLOv7 model

YOLOv8

The YOLOv8 architecture uses a modified CSPDarknet53 backbone. The C2f module replaces the CSPLayer used in YOLOv5. The spatial pyramid pooling fast (SPPF) layer speeds up the computation by aggregating features into a fixed-size map. Each convolution has batch normalization and SiLU activation. The head is separated to process objectivity, classification and regression tasks independently. YOLOv8 uses a backbone similar to YOLOV5 with some modifications to the CSPLayer, called the C2f module. The C2f module (partial bottleneck with two convolutional cross-over stages) combines high-level features with contextual information to improve detection accuracy.

In the output layer of YOLOv8, the sigmoid function is used as the activation function for the objectivity score, which represents the probability that the bounding box contains an object [31].

The sigmoid activation function is a continuous and differentiable function that produces a value between 0 and 1 for each of the input values[32]. The formula for the sigmoid function is shown below in Equation 1. NET is the total stimulus entering a nerve cell. It is a sum of signals from other neurons and a bias term. e is the Euler constant and is approximately equal to 2.71828.

$$F(NET) = \frac{1}{1 + e^{-NET}}$$

YOLOv8 uses the Softmax function for class probabilities, which represents the probability of objects belonging to each possible class[31].

The Softmax layer consists of exponential and division functions, which require complex modules during hardware implementation. The exponential function is modeled in an optimized way using the CORDIC algorithm. This algorithm gives the cosh and sinh values that need to be added using a floating-point adder to obtain the exponential value. It uses the same floating-point adder to perform the summation required in the denominator of the Softmax function. The division operation in the Softmax function is solved by applying the binary division algorithm[33]. The formula for the Softmax function is shown in Equation 2.

$$\sigma^{-1}(z)_j = 1 + \sum_{i=1}^{K} e^{z_i - z_j} : i \neq j$$

In Equation 2, z is an arbitrary vector with real values zi, where i=1,2,3,...K is generated in the penultimate network layer and K is the vector size. In the Softmax case, the maximum likelihood state or the output of the layer with the maximum value is selected as the class predicted by the model, while in the Inverse Softmax case, the class predictions are reversed and the output with the smallest value is the predicted class. Although there is only reordering, no information is lost during the training process by replacing the Softmax function with the Inverse Softmax function[33].

Metrics

After the model training is completed and tested, the results of the model are evaluated using various parameters. As a result of the comparison of the YOLO deep learning models used in the study, TP, FP, TN, FN, accuracy, precision, recall, F1-score and AUC (Area Under the Curve) metrics were used for statistical analysis of the models. The evaluation metrics used and their formulations are shown in Table 1.

Table 1. Metrics used and their formulations

Metric	Formulation	
Accuracy: Determines how accurately the model works. It is calculated as the ratio of correct predictions to the total dataset. The accuracy score is between 0 and 1 and the model is considered successful if the score approaches 1.	(TP + TN) / (TP + TN + FP + FN)	(3)
Precision: True Positive indicates how much of the predicted data is positive. It is calculated as the ratio of true positive (TP) predictions to the entire positive (TP + FP) prediction dataset.	TP / (TP + FP)	(4)
Recall: The ability of the model to predict positive cases. True positives indicate how much of the samples are correctly predicted as positive. It is calculated as the ratio of true positive prediction (TP) to the sum of true positive and false negative values.	TP / (TP + FN)	(5)
F1-Score: It is the harmonic mean between precision and recall. It is between 0 and 1, with scores approaching 1 indicating that the model is considered successful.	2 * (Precision * Recall) / (Precision + Recall)	(6)
AUC Value (Area Under the Curve): It measures how well a model discriminates between different classes. A higher AUC value indicates better predictive ability of the model.	It is the area under the ROC curve.	(7)

TP (True positive): Predicting degenerated data as degenerated, **FP (False positive):** Predicting normal data as degenerated, **TN (True negative):** Predicting normal data as normal, **FN (False negative):** Predicting degenerated data as normal.

Results

The model training required for the study was performed on Google Colab. Colab is a hosted Jupyter Notebook service that provides free access to computing resources including GPUs and TPUs and does not require any installation[34]. The hyperparameters of the YOLO deep learning algorithms were configured as 150 epochs and 8 batches. Epoch determines the number of iterations of the dataset on training, while Batch-size refers to the number of data processed in each epoch and is the size of the subset of the dataset fed into the model at each iteration within an epoch[27]. In our dataset, which contains 1480 data in total, 134 (9.05%) data were used for testing, 1215 (82.09%) for training and 131 (8.85%) for validation. The flowchart of the data splitting, training and testing process used in the study is shown in Figure 10.

Experimental Results of YOLO Models with BSRGAN and Bilinear Interpolation

In this part of the study, model training was performed using YOLOv5, YOLOv7 and YOLOv8 algorithms with a dataset containing BSRGAN and bilinear interpolated panoramic X-ray images. In order to fully understand the actual performance of each model, the model results were evaluated and compared with the performance metrics given in Table 1 and the best model was selected based on these evaluations.

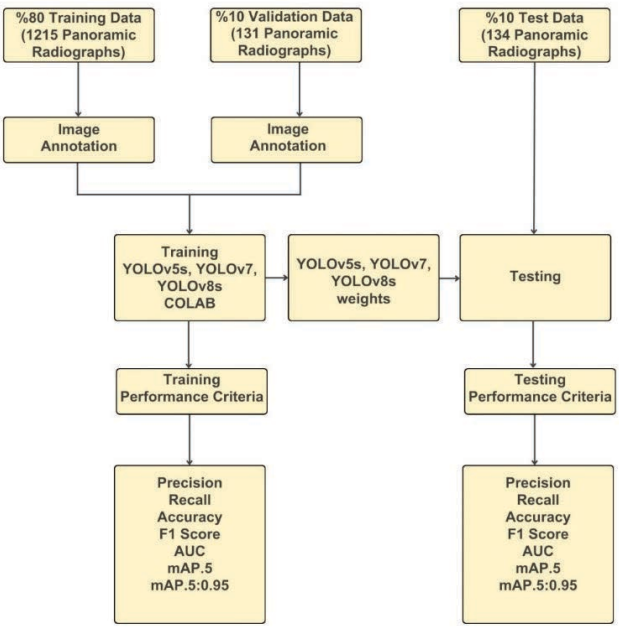


Figure 10. *The flow chart of diving the data, training, and testing process*

YOLOv5s Results

In the study, YOLOv5s, one of the sub-architectures of YOLOv5, was preferred. The graphs obtained as a result of YOLOv5s training are shown in Figures 11 and 12. It was observed that the trained model achieved 94.40% accuracy, 81.63% precision, 86.96% recall, 84.21% F1 score and 91.45% AUC with the test dataset.

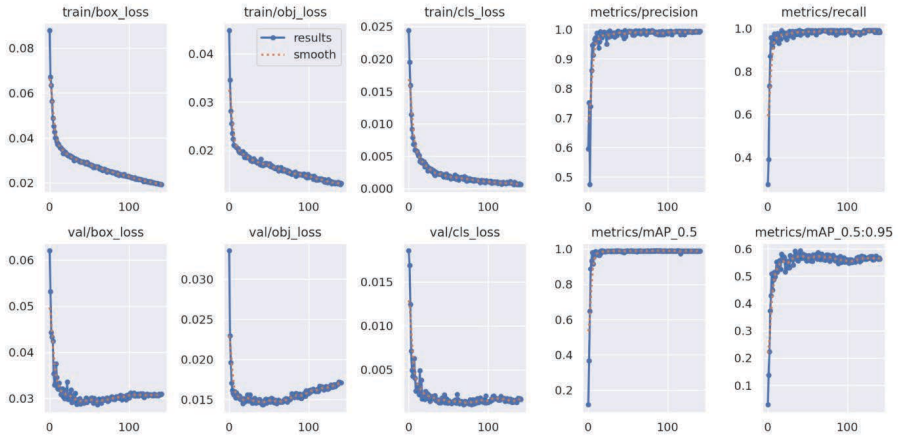


Figure 11. YOLOv5s performance graphs

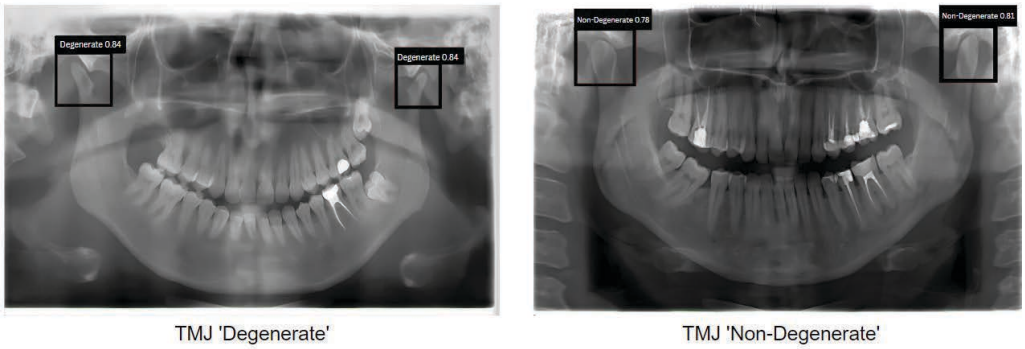


Figure 12. Examples of YOLOv5 test results

YOLOv7 Results

The graphs obtained as a result of YOLOv7 training are shown in Figure 13. The trained model achieved 99.63% Accuracy, 97.87% Precision, 100% Recall, 98.92% F1 Score and 99.77% AUC on the test dataset.

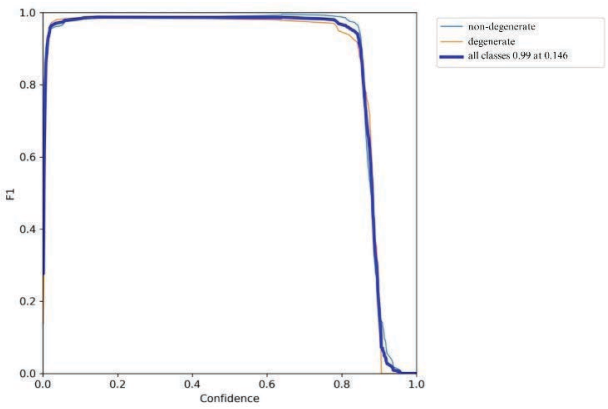


Figure 13. YOLOv7 F1-score graphs

When we examine the F1 score on the graph in Figure 13, it shows a high performance for the healthy class and a consistent performance throughout the confidence interval. The F1 score for the patient class is also high, but less consistent than the healthy class. As a result, the model is able to distinguish between healthy and patient classes with high accuracy.

In the Precision graph, it starts with a precision of approximately 0.8. This shows that the model quickly identifies the targets correctly at the beginning of the training process. As the training progresses, the precision increases rapidly, indicating that the model is successful in correctly identifying the targets without generating false positives. These results show that the model performs well and correctly identifies the targets.

The training result shows that the model achieves a high recall rate. Recall shows how many positive examples are correctly identified. The model is able to correctly identify the majority of positive examples.

The mAP@0.5 value in Figure 14 is around 0.8-0.9. This shows that the model can detect objects with 80%-90% accuracy. The Figure 15. shows the results of YOLOv7 test.

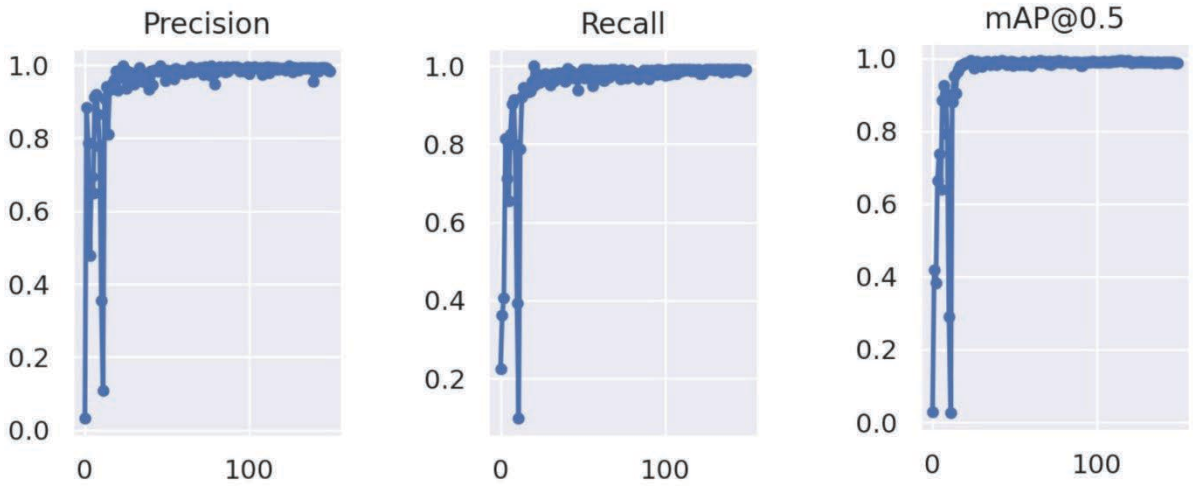


Figure 14. YOLOv7 performance graphs

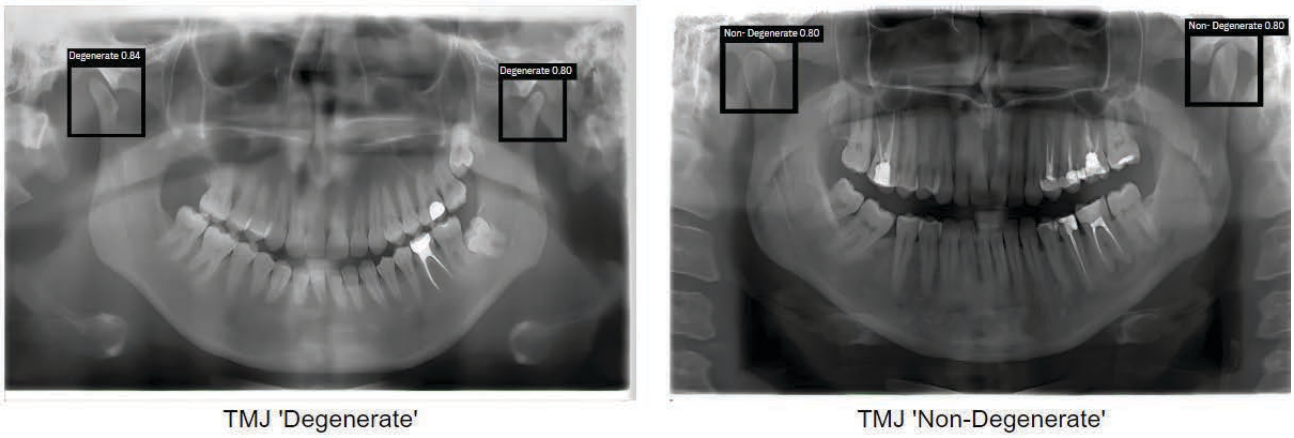


Figure 15. Examples of YOLOv7 test results

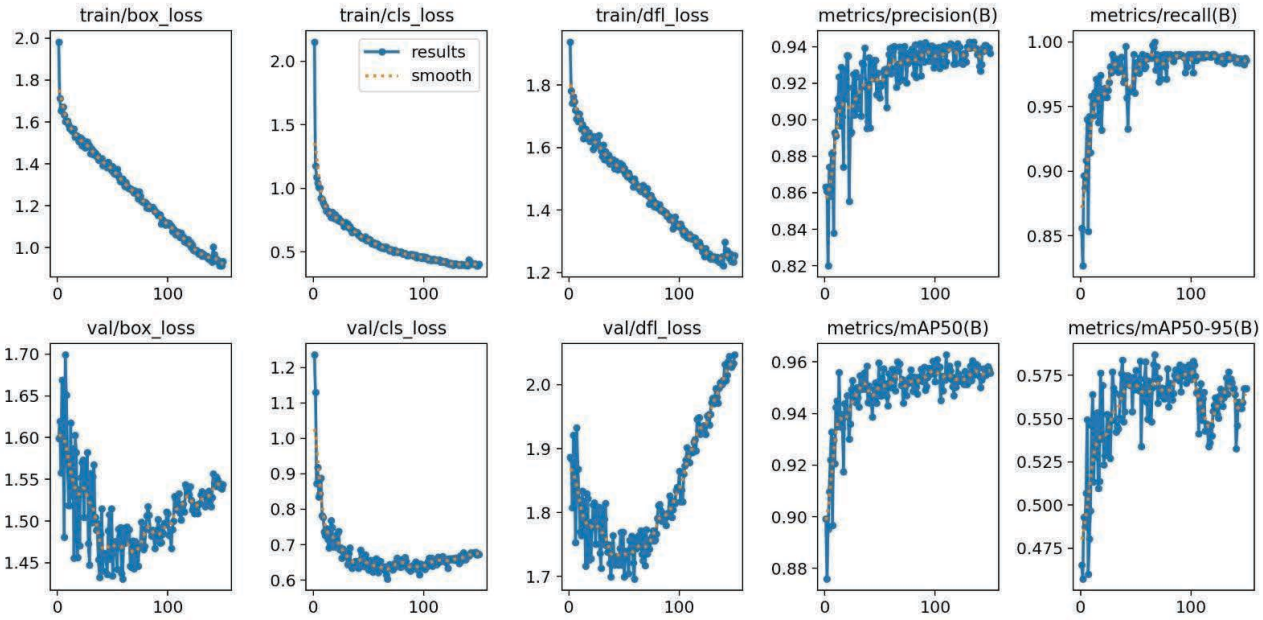


Figure 16. YOLOv8 performance graphs

YOLOv8 Results

The graphs obtained as a result of YOLOv8 training are shown in Figure 16. The trained model achieved 96.64% accuracy, 91.11% precision, 89.13% recall, 90.11% F1 Score and 93.66% AUC on the test dataset. The Figure 17. shows the results of YOLOv8 test.

The metrics obtained as a result of the studies on the datasets with BSRGAN and bilinear interpolation are compared in Table 2. As a result of this comparison, it was decided to use the YOLOV7 deep learning algorithm, which stands out among the models, for the next training with the morphologically processed dataset.

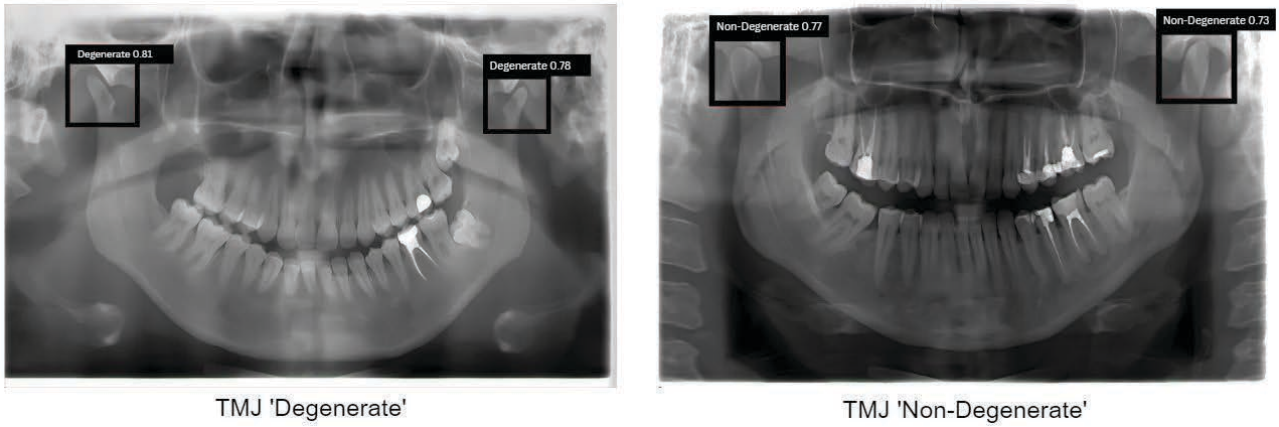


Figure 17. Examples of YOLOv8 test results

Table 2. Metrics of YOLO Models with BSRGAN and Bilinear Interpolation

MODEL	ACCURACY	PRECISION	RECALL	F1 SCORE	AUC	MAP@.5	MAP@.5:.95
YOLOV5	0.944	0.816	0.869	0.842	0.914	0.989	0.593
YOLOV7	0.996	0.978	1.0	0.989	0.997	0.988	0.552
YOLOV8	0.966	0.911	0.891	0.901	0.936	0.955	0.586

Table 3. Confusion Matrix of YOLO Models with BSRGAN and Bilinear Interpolation

MODEL	TN (LABEL)	FP (LABEL)	FN (LABEL)	TP (LABEL)
YOLOV5	79.48% (213)	3.36% (9)	2.24% (6)	14.93% (40)
YOLOV7	82.46% (221)	0.37% (1)	0.00% (0)	17.16% (46)
YOLOV8	81.34% (218)	1.49% (4)	1.87% (5)	15.30% (1)

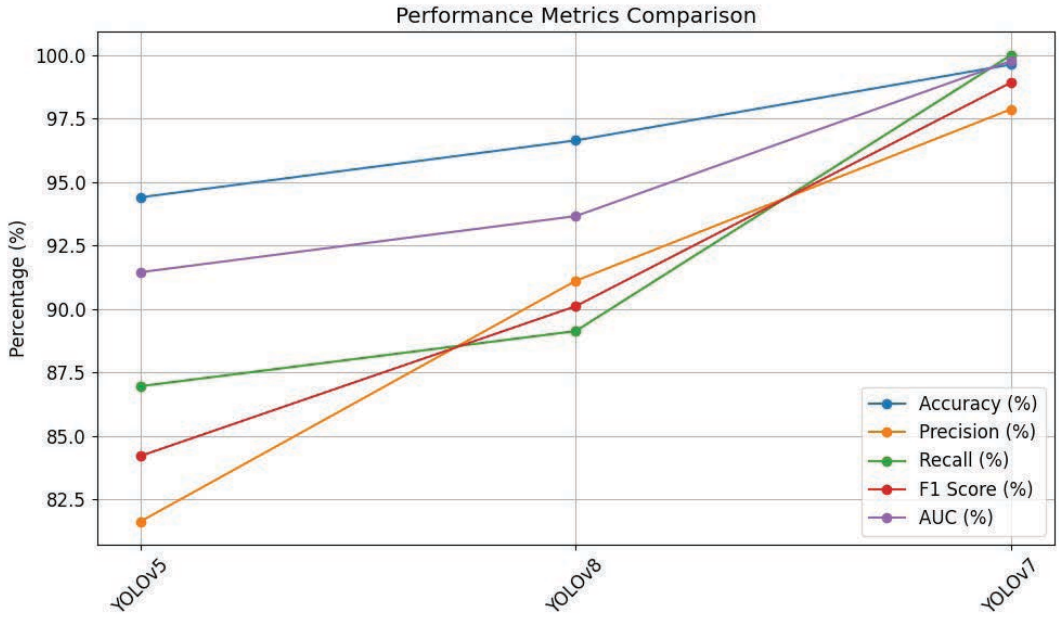


Figure 18. Performance Metrics Comparison of YOLO Models with BSRGAN and Bilinear Interpolation

Experimental Results of YOLOv7 with Morphological Transformations, BSRGAN and Bilinear Interpolation

In this section, morphological transformations, BSRGAN and bilinear interpolation are applied to the dataset respectively. The application of morphological transformation to the dataset aims to observe how the result performance metrics of the YOLO algorithms will change (Table 3 and Figure 18).

Erosion

Morphological processing, which is the erosion method, was performed on the data by removing the white noise in the image and reducing the thickness and size of the foreground object. The training result of the YOLOV7 algorithm with erosion processed data is shown in Figure 19 and examples on Figure 20. As a result of the evaluations made with the test data, 95.90% accuracy, 84.31% precision, 93.48% recall, 88.66% F1 score and 99.94% AUC value were obtained.

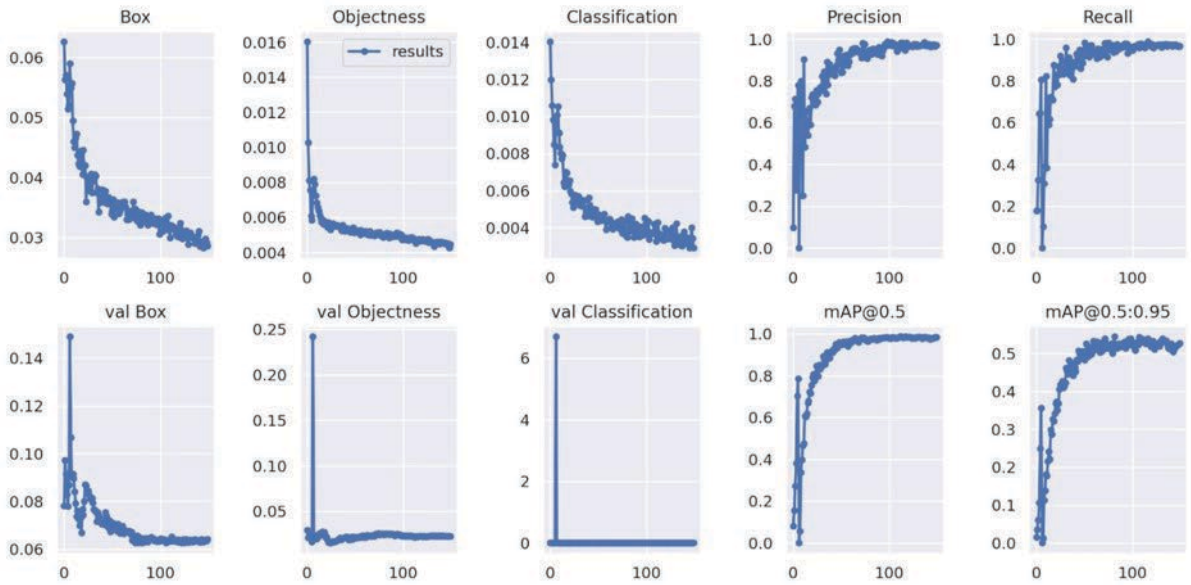


Figure 19. YOLOv7 with erosion performance graph

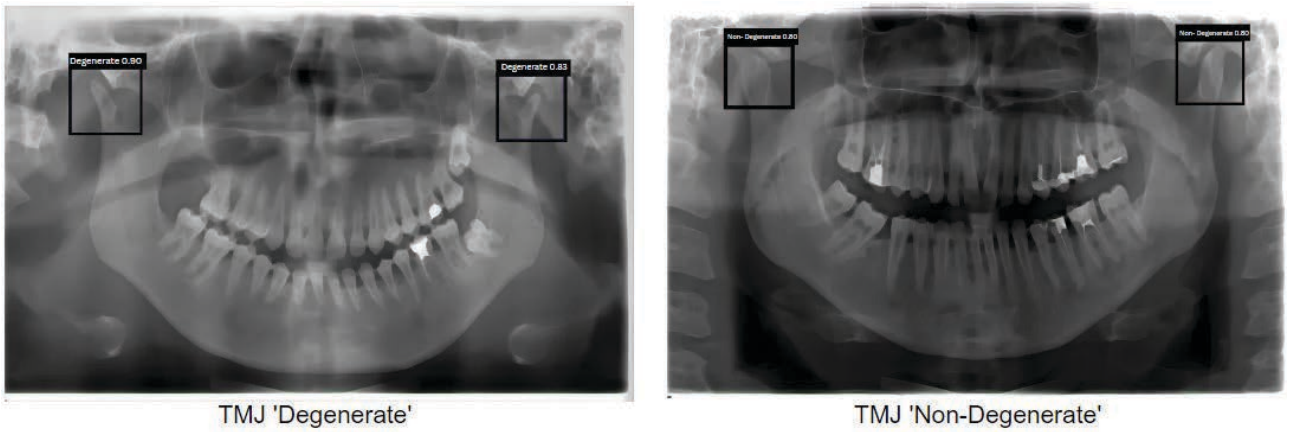


Figure 20. Examples of YOLOv7 erosion test results

Dilation

The dilation method, in contrast to the erosion method, performs a process that increases the white region in the image and the size of the foreground object. The model training result is shown in Figure 21 and examples on Figure 22. The test result of the training resulted in 96.64% accuracy, 93.02% precision, 86.96% recall, 89.89% F1 score and 92.80% AUC.

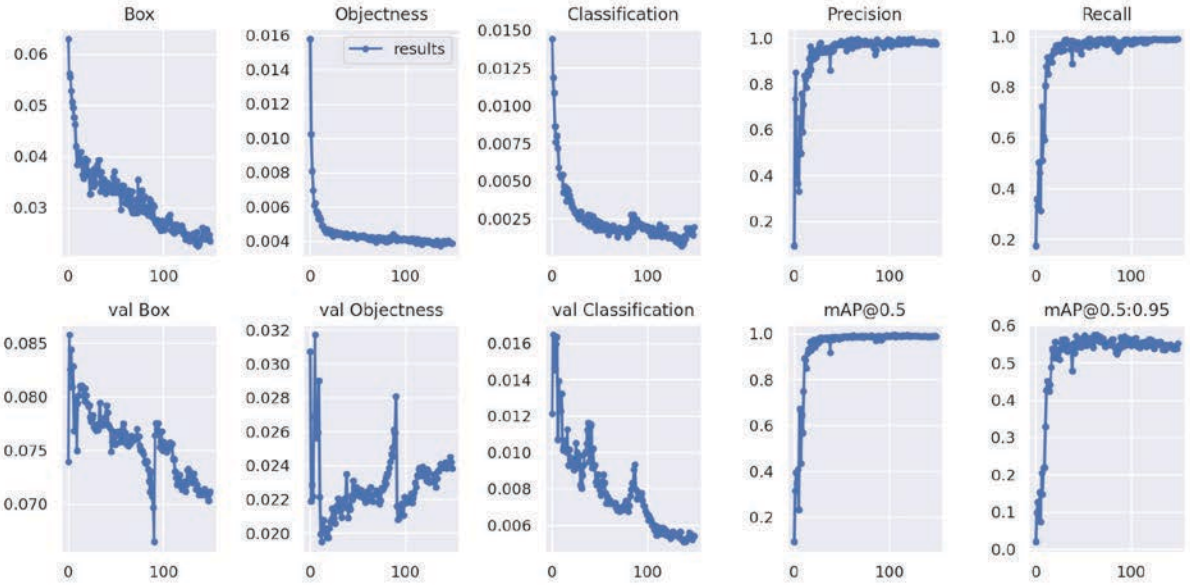


Figure 21. YOLOv7 with dilation performance graph

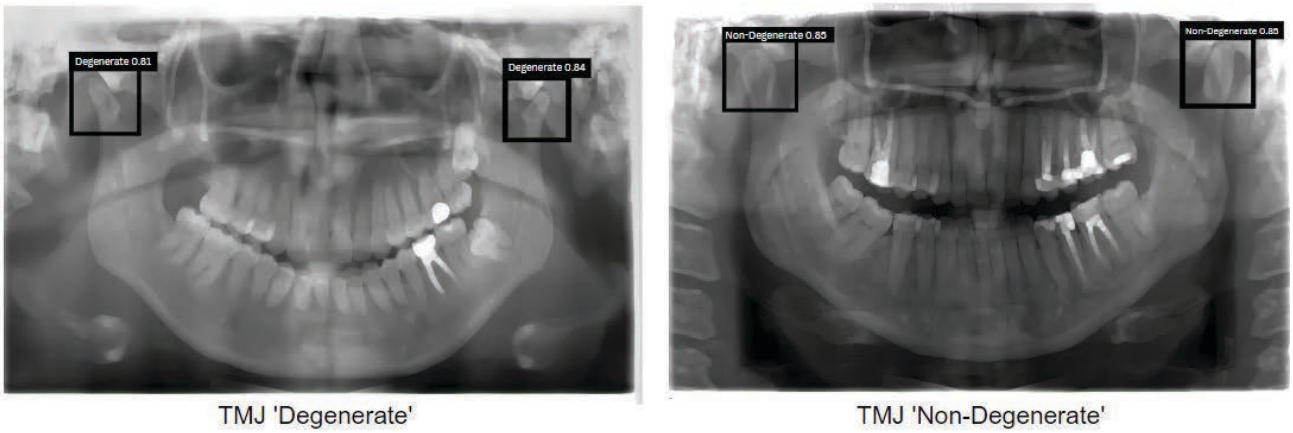


Figure 22. Examples of YOLOv7 dilation test results

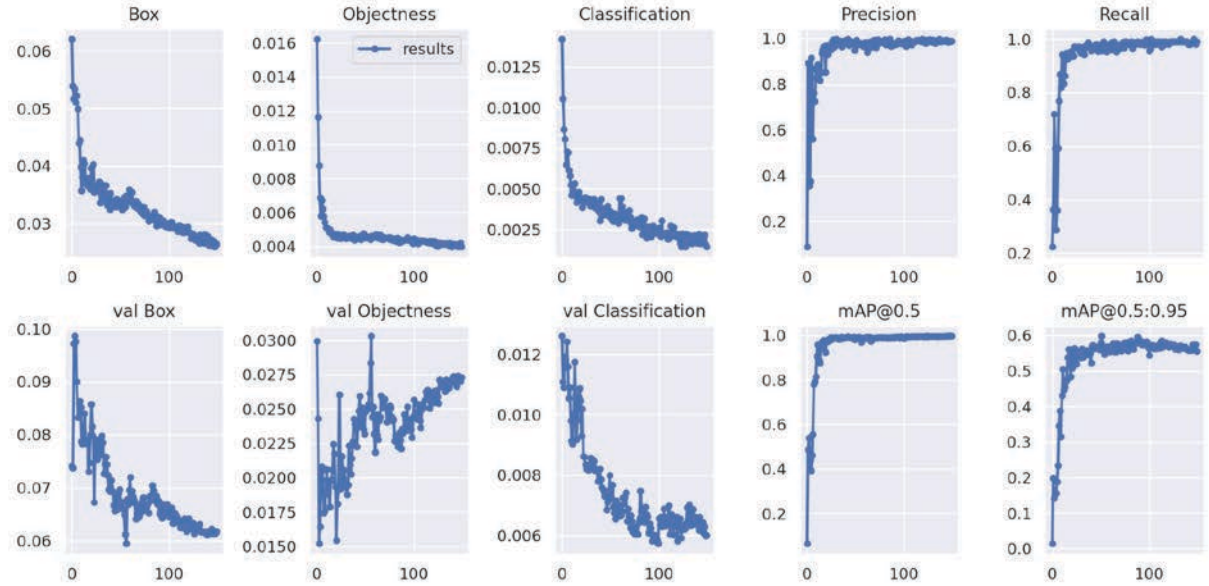


Figure 23. YOLOv7 with opening performance graph

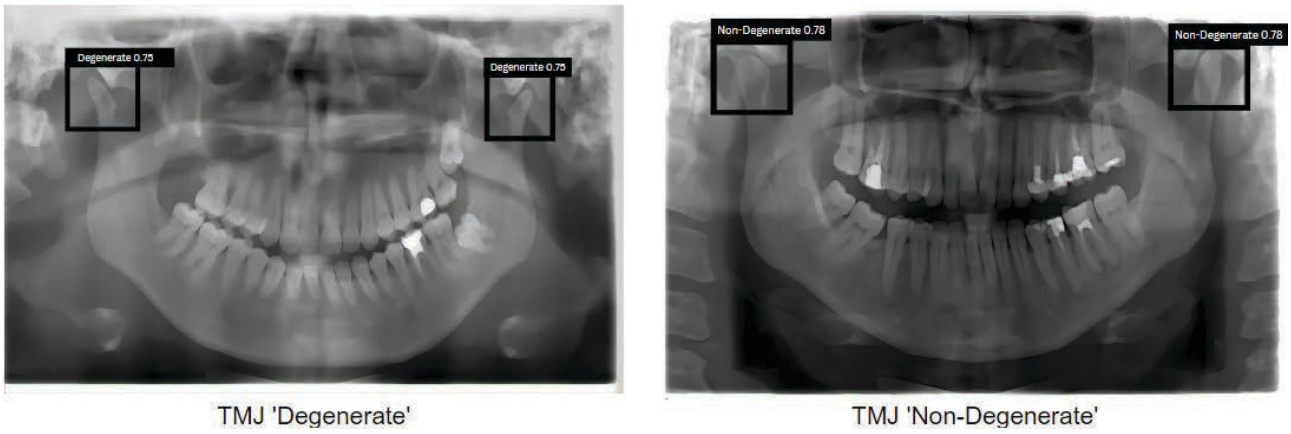


Figure 24. YOLOv7 opening test result examples

Opening

As a result of erosion, noises are removed, but the size of the object is reduced. In order to prevent this, dilation is applied to the eroded image to enlarge the area of the reduced object. This process is called opening. As a result of the morphological transformation, the performance metrics in the test data were measured as accuracy rate 97.76%, precision rate 93.48%, recall rate 93.48%, F1 score rate 93.48% and AUC value 96.06%. The performance graphs of the model as a result of training are shown in Figure 23 and the results in Figure 24.

Closing

In the Closing method, the erosion method was performed after dilation. This method is useful for covering small holes in foreground objects or small black spots on the object. At the end of training, the metrics of 98.88% accuracy, 95.74% precision, 97.83% recall, 96.77% F1 score and 98.46% AUC value were obtained. The performance graphs of the training results are given in Figures 25 and 26.

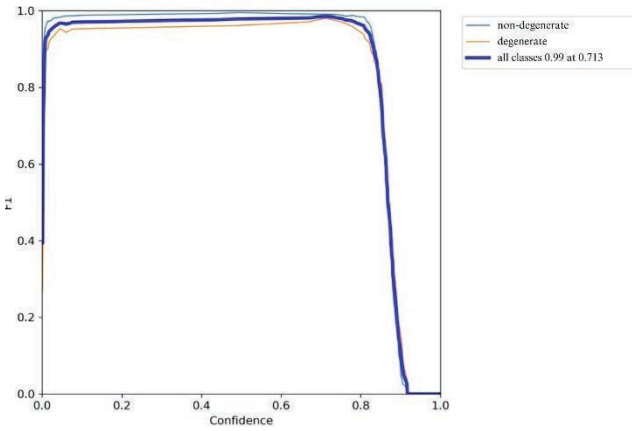


Figure 25. YOLOv7-Closing F1-score graphs

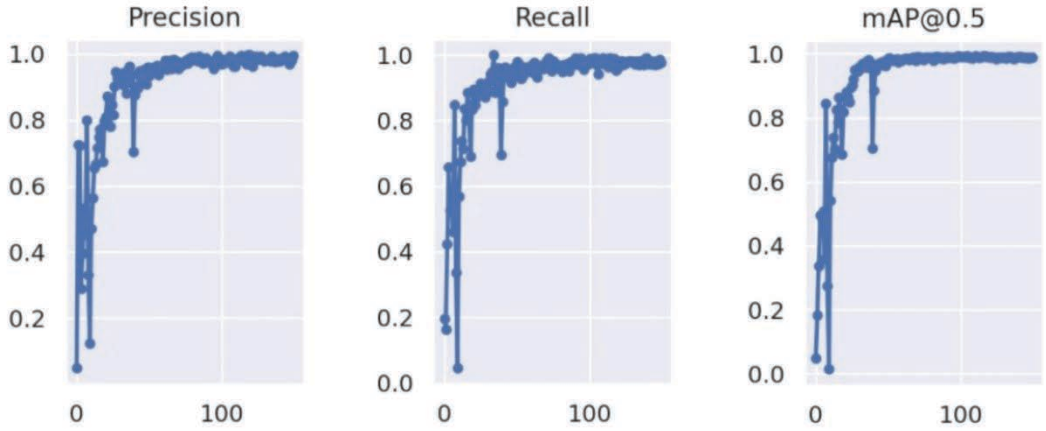


Figure 26. YOLOv7-Closing performance graphs

Looking at the curve in the graph in Figure 25, we see that the F1 score is high for the healthy class and remains consistent across the confidence interval. The F1 score is also high for the degenerated class, but less consistent than for the normal class. As the confidence interval increases, a decrease in the F1 score is observed. In summary, the model is able to distinguish between normal and degenerated classes with high accuracy. The F1 score for all classes is also acceptable.

In Figure 26, the precision plot shows that the precision is generally quite high. The curve is above 0.8, which means that the model correctly classifies more than 80% of the positive samples. However, we can see that the curve does not represent perfect precision (1). This means that the model incorrectly classifies some positive samples. Overall, this graph shows that the model has a high precision. However, it should be noted that the model still makes some mistakes.

Looking at the Recall graph, the Yolov7 algorithm generally recalls objects with high accuracy. However, it has some limiting cases with recall values close to 0. This means that the algorithm has difficulty in detecting some objects.

This graph shows the change over time of a measure of mAP@0.5 (average mean precision). mAP@0.5 is a metric used to evaluate the performance of an object detection model. It takes a value between 0 and 1, with values closer to 1 indicating better performance. The slope of the curve in the graph shows how mAP@0.5 has changed over time:

If the slope is up: mAP@0.5 is increasing over time, which means that the performance of the model is improving.

If the slope is down: mAP@0.5 is decreasing over time, which means that the model's performance is decreasing.

If the slope is horizontal: mAP@0.5 does not change over time, which means that the model's performance remains constant. Example is shown on Figure 27.

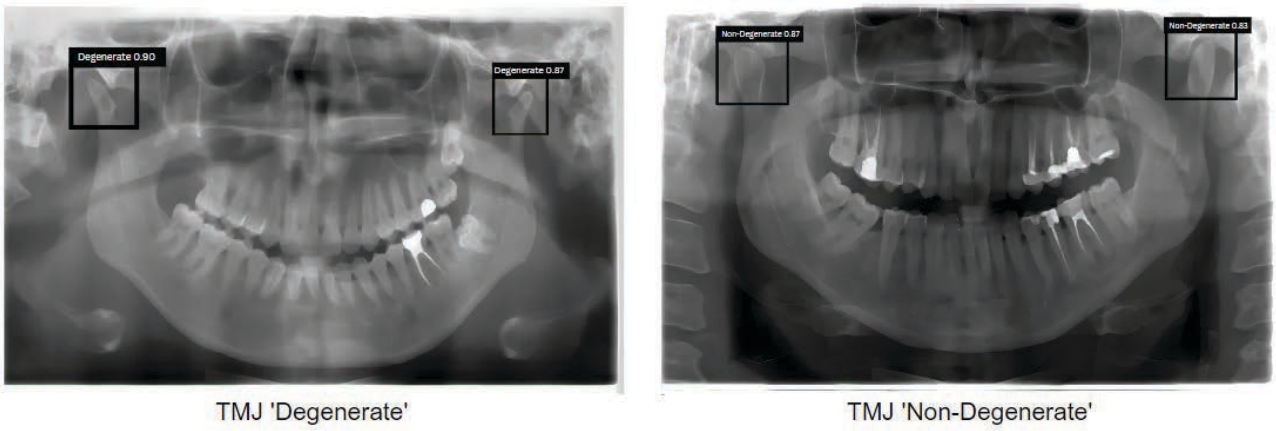


Figure 27. YOLOv7 closing test result examples

Table 4. Metrics of YOLOv7 with Morphological Transformations, BSRGAN and Bilinear Interpolation

MODEL	ACCURACY	PRECISION	RECALL	F1 SCORE	AUC	MAP@.5	MAP@.5:.95
YOLOV7 EROSION	0.959	0.843	0.935	0.887	0.949	0.984	0.528
YOLOV7 DILATION	0.966	0.930	0.870	0.899	0.928	0.989	0.552
YOLOV7 OPENING	0.978	0.935	0.935	0.935	0.961	0.996	0.555
YOLOV7 CLOSING	0.989	0.957	0.978	0.968	0.985	0.99	0.558

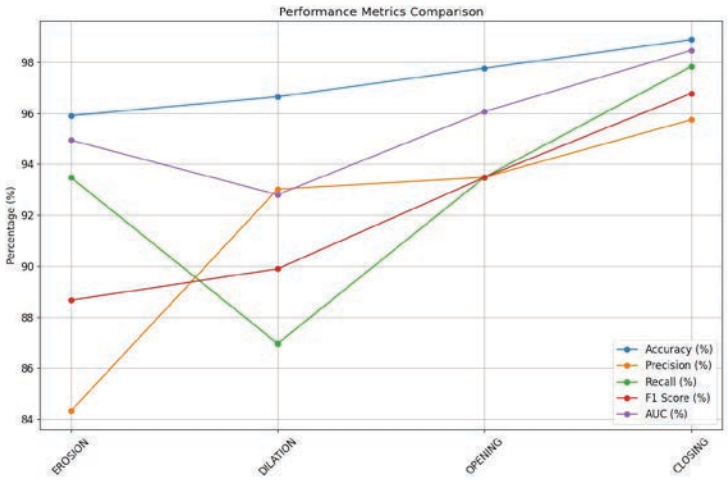


Figure 28. Performance Metrics Comparison of YOLOv7 with Morphological Transformations, BSRGAN and Bilinear Interpolation

The erosion operation reduces the object area in the data, preventing the best results. On the other hand, dilation increases the scope of the object by expanding the object size. Opening can smooth out the boundaries and make them smoother by expanding the area of the object while preserving small details on the object. However, these operations may not be as effective as closing in preserving the sharp boundaries of the object and emphasizing small details. Closing is used to fill small holes in foreground objects and to cover small black dots on the object. This tightens and sharpens the object boundaries, resulting in a clearer and more defined object. In addition, the closing process increases the homogeneity of the object by filling small gaps within the object (Table 4, Figure 28).

Discussion

TMJ osteoarthritis, an important subtype of temporomandibular disorders, is a joint disease caused by a series of degenerative processes. It leads to dysfunction, joint pain, malocclusion and decreased quality of life[35]. Its presence is confirmed by structural bone changes observed on radiographic examination. Orthopantomogram (OPG), one of the routine imaging modalities, is an important two-dimensional diagnostic tool in the primary diagnosis of TMJ osteoarthritis, but it is less successful than other advanced imaging modalities in identifying small-sized bone changes in the TMJ structure. Its widespread use, low cost and easy accessibility are the most important factors in the widespread use of panoramic radiographs in primary diagnosis compared to other methods[36].

The use of artificial intelligence (AI) in visual tasks has generated significant interest in the field of dentistry. Machine learning (ML), an important subset of artificial intelligence, is widely used in many fields today to solve complex problems. Deep learning (DL) is a type of ML created by simulating the connectivity of human brain neurons in processing complex images[37]. ML algorithms are used on dental radiographs, and studies on TMJ diseases are increasing. In our study, the YOLO algorithm, a single-stage deep learning model based on a convolutional neural network, was used. YOLO is faster than two-stage sensors, has high sensitivity, and does not require a complex data[19]. In this study, the performance of YOLO v5, YOLOv7 and YOLOv8 was compared in detecting degenerative changes that may be associated with osteoarthritis.

There are few studies in the literature on the evaluation of degenerations in the TMJ region. Choi *et al.* [38] compared the performance of a pre-trained ResNet artificial intelligence model with a maxillofacial radiologist. The 1189 panoramic radiographs used in this study were categorized into three groups: normal, indeterminate osteoarthritis, and osteoarthritis. The diagnostic accuracy of the model was 0.78, sensitivity: 0.73 and specificity: 0.82. Jung *et al.* [39] compared the diagnostic performance of the pre-trained Resnet152 and EfficientNet-B7 algorithms in the detection of TMJ degenerations (osteophytes, subchondral cysts, erosions) with an expert physician using panoramic radiographs of 858 individuals. A region of interest (ROI) of 224x224 pixels was used, segmented from the condyle region. ResNet152 yielded a sensitivity of 0.94, specificity of 0.79, and accuracy of 0.87, while EfficientNet-B7 yielded a sensitivity of 0.86, specificity of 0.91, and accuracy of 0.88. In a study comparing different algorithms for the classification of TMJ diseases, the highest specificity was 94% using fine-tuned VGG16 [40]. The highest sensitivity was 100% using SVM, random forest, logistic regression [41] and YOLOv5 [42]. In our study; YOLOv5, YOLOv7 and YOLOv8 algorithms were used and their results were compared. Although all three versions have high success rates, the best results were obtained in YOLOv7 with 99.63% accuracy, 97.87% precision, 100% sensitivity,

98.92% F1 Score and 99.77% AUC values. The results showed higher detection performance than previous studies reported in the literature. It is very important to detect degenerative changes that may be associated with osteoarthritis at the earliest possible stage and start treatment. In clinics that do not have an oral radiologist, it is possible that only the teeth are examined, and the general evaluation of the jaws is neglected, and degenerations in the TMJ bone tissue are overlooked. Thanks to the successful results obtained in this study, creating a clinical decision support system promises hope for the dentist to successfully perform the TMJ evaluation of the patient without disregarding it.

The limitation of this study is that it was conducted with images obtained from a single center with a limited data set. In addition, only radiographic findings were evaluated, and clinical findings related to the possible diagnosis of TMJ osteoarthritis were not evaluated.

Conclusions

YOLOv7 algorithm showed the highest performance which resulted in 99.63% accuracy, 97.87% precision, 100% sensitivity, 98.92% F1 Score and 99.77% AUC values in detecting degenerative changes in the TMJ region. In future deep learning studies, more comprehensive studies with multi-center collaboration are needed to combine clinical data regarding the TMJ region with radiological findings.

Acknowledgement. This study was supported by Scientific and Technological Research Council of Turkey (TUBITAK) under the Grant Number 123E431. The authors thank to TUBITAK for their supports.

References

1. Okeson J. Management of temporomandibular disorders and occlusion. 8th ed. Mosbey 2019, USA.
2. Dsouza VD, Rao PK, Kini R (2020). „Assessment of degenerative changes in the temporomandibular joint: A retrospective cone-beam computed tomography study". *SRM J Res Dent Sci*, **11** (4): 195-198.
3. Shetty US, Burde KN, Naikmasur VG, Sattur AP (2014). "Assessment of condylar changes in patients with temporomandibular joint pain using digital volumetric tomography". *Radiol Res Pract*. **2014**: 106059. doi: 10.1155/2014/106059. PMID: 25332835
4. Koç N (2020). Evaluation of osteoarthritic changes in the temporomandibular joint and their correlations with age: A retrospective CBCT study. *Dent Med Probl.* **57** (1): 67-72. doi: 10.17219/dmp/112392. PMID: 31997586
5. Bäck K, Ahlqwist M, Hakeberg M, Dahlström L (2017). „Occurrence of signs of osteoarthritis/arthrosis in the temporomandibular joint on panoramic radiographs in Swedish women". *Community Dent Oral Epidemiol*. **45** (5): 478-484. doi: 10.1111/cdoe.12312. PMID: 28699681.
6. Yıldırım D, Alkış Ü (2016). "Imaging methods used in the evaluation of temporomandibular joint disorders". *Med J SDU*, **7** (2).
7. Serindere G, Belgin CA (2019). "Examination of condylar changes with panoramic radiography". *Selcuk Dent J*, **6** (4): 173-177.
8. Krishnamoorthy VK, Baskaran S (2022). "Optimized adversarial network with faster residual deep learning for osteoarthritis classification in panoramic radiography". *Int J Intell Eng Syst*. 2022; **15** (6): 191-200. doi: 10.22266/ijies2022.1231
9. Li M, Punithakumar K, Major PW, Le LH, Nguyen KCT, Pereira CP, et al. (2022). "Temporomandibular joint segmentation in MRI images using deep learning". *J Dent.* **127**: 104345. doi: 10.1016/j.jdent.2022.104345. PMID: 36368120
10. Erzurumlu ZÜ, Çelenk P (2022). "Effects of aging on temporomandibular joint". *Turk Dent Res J*, **1** (2): 75-78.
11. Yan Y, Chen M, Shyu ML, Chen SC (2015). "Deep learning for imbalanced multimedia data classification". *IEEE International Symposium on Multimedia (ISM)*, Miami, FL, USA, pp. 483-488, doi: 10.1109/ISM.2015.126.
12. Göceri E (2023). "Medical image data augmentation: techniques, comparisons and interpretations". *Artif Intell Rev.* **56** (11): 12561-12605. doi: 10.1007/s10462-023-10453-z. PMID: 37362888
13. Yıldız G, Yıldız D (2018). "Brain tumour location and tumour area calculation by means of morphological processes and edge detection methods". *IJMSIT*, **2** (2): 39-42.
14. (March 1,2024) https://docs.opencv.org/
15. Titus J, Geroge S (2013). "A comparison study on different interpolation methods based on satellite images". *Int J Eng Res Technol.* **2** (6): 82-85.
16. Zhang K, Liang J, Gool L, Timofte R (2021). "Designing a practical degradation model for deep blind image super-resolution". *IEEE Comput. Soc. Conf. Comput. Vis*. pp: 4771-4780. doi: 10.1109/ICCV48922.2021.00475.
17. Redmon J, Farhadi A (2017)."YOLO9000: Better, Faster, Stronger". *IEEE Conf Comput Vis Pattern Recognit.* pp. 6517-6525. doi: 10.1109/CVPR.2017.690.
18. Redmon J, Farhadi A. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767. 2018.
19. Bochkovskiy A, Wang CY, Liao HYM (2020). "Yolov4: Optimal speed and accuracy of object detection". arXiv preprint arXiv:2004.10934. doi: 10.48550/arXiv.2004.10934.
20. (March 3,2024) Jocher G, Nishimura K, Mineeva T, Vilariño R. Yolov5" Code repository https://github.com/ultralytics/yolov5. 2020.
21. Kaya E, Gunec HG, Gokyay SS, Kutal S, Gulum S, Ates HF. Proposing a CNN method for primary and permanent tooth detection and enumeration on pediatric dental radiographs. J Clin Pediatr Dent 2022.

22. Redmon J, Divvala S, Girshick R, Farhadi A (2016). "You Only Look Once: Unified, Real-Time Object Detection". *IEEE Conf Comput Vis Pattern Recognit.* pp. 779-788, doi: 10.1109/CVPR.2016.91.
23. Güney E (2021). "Mobile gpu based real-time situation analysis and detection applications for driver assistance systems". Sakarya University.
24. Wang H, Zhang S, Zhao S, Wang Q, Li D, Zhao R (2022). "Real-time detection and tracking of fish abnormal behavior based on improved YOLOV5 and SiamRPN++". *Comput Electron Agric*. **192** : 106512. doi: 10.1016/j.compag.2021.106512.
25. Kivrak O, Gürbüz MZ (2022). "Performance comparison of YOLOv3, YOLOv4 and YOLOv5 algorithms: a case study for poultry recognition". *EJOSAT*. 38: 392-397. doi: 10.31590/ejosat.1111288
26. (March 4, 2024). https://github.com/ultralytics/yolov5/issues/8785
27. (March 22, 2024). https://docs.ultralytics.com/tr/models/yolov7/
28. Li K, Wang Y, Hu Z (2023). „Improved YOLOv7 for small object detection algorithm based on attention and dynamic convolution". *Appl Sci*. **13** (16): 9316. doi: 10.3390/app13169316
29. Wang C, Bochkovskiy YA, Liao HYM (2022). "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors". *Proc IEEE Conf Comput Vis Pattern Recognit*. doi: 10.48550/arXiv.2207.02696.
30. (March 22, 2024). What is YOLOv7? A Complete Guide. Roboflow Blog: https://blog.roboflow.com/yolov7-breakdown/
31. Terven J, Córdova-Esparza DM, Romero-González JAA (2023). "Comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS". *Mach Learn Knowl Extr.* **5** (4): 1680-1716. doi: 10.3390/make5040083
32. Çelik Ş (2022). "Application of Artificial Neural Networks in Different Activation Functions: Prediction of Duck Population in Turkey". *OKU J Inst Sci Technol.* **5** (2): 800-811. doi: 10.47495/okufbed.990995
33. Raghuram S, Bharadwaj AS, Deepika SK, Khadabadi MS, A. Jayaprakash A (2022). "Digital Implementation of the Softmax Activation Function and the Inverse Softmax Function". *4th International Conference on Circuits, Control, Communication and Computing.* pp. 64-67. doi: 10.1109/I4C57141.2022.10057747.
34. (March 22, 2024). https://colab.google/
35. Wang X, Zhang JN, Gan YH, Zhou YH (2015). „Current understanding of pathogenesis and treatment of TMJ osteoarthritis". *J Dent Res*. **94** (5): 666-673. doi: 10.1177/0022034515574770. PMID: 25744069
36. Buckwalter JA, Mankin HJ, Grodzinsky AJ (2005). „Articular cartilage and osteoarthritis". *Instr Course Lect*. **54**: 465-480. PMID: 15952258
37. Kulkarni S, Seneviratne N, Baig MS, Khan AHA (2020). "Artificial intelligence in medicine: where are we now?" *Acad Radiol*, **27** (1): 62–70. doi: 10.1016/j.acra.2019.10.001.
38. Choi E, Kim D, Lee JY, Park HK (2021). „Artificial intelligence in detecting temporomandibular joint osteoarthritis on orthopantomogram". *Sci Rep*. **11** (1): 10246. doi: 10.1038/s41598-021-89742-y. PMID: 33986459
39. Jung W, Lee KE, Suh BJ, Seok H, Lee DW (2023). "Deep learning for osteoarthritis classification in temporomandibular joint". *Oral Dis*. **29** (3): 1050-1059. doi: 10.1111/odi.14056. PMID: 34689379
40. Kim D, Choi E, Jeong HG, Chang J, Youm S (2020). "Expert system for mandibular condyle detection and osteoarthritis classification in panoramic imaging using R-CNN and CNN". *Appl Sci*, **10**: 7464. doi: 10.3390/app10217464.
41. Orhan K, Driesen L, Shujaat S, Jacobs R, Chai X (2021). "Development and validation of a magnetic resonance imaging-based machine learning model for TMJ pathologies". *Biomed Res Int*. **2021**:6656773. doi: 10.1155/2021/6656773. PMID: 34327235
42. Eşer G, Duman ŞB, Bayrakdar İ, Çelik Ö (2023). "Classification of temporomandibular joint osteoarthritis on cone beam computed tomography images using artificial intelligence system". *J Oral Rehabil*. **50** (9): 758-766. doi: 10.1111/joor.13481. PMID: 37186400

Received on March 14, 2024.
Revised on April 10, 2024.
Accepted on May 29, 2024.

Conflict of Interests: Nothing to declare.
Financial Disclosure Statement: Nothing to declare.
Human Right Statement: All the procedures on humans were conducted in accordance with the Helsinki Declaration of 1975, as revised 2000. Consent was obtained from the patient/s and approved for the current study by national ethical committee.
Animal Rights Statement: None required.

Correspondence

Busra Ozturk
Department of Oral and Maxillofacial Radiology
Necmettin Erbakan University, Faculty of Dentistry, Konya, Turkey
e-mail: dtbusraozturk1@gmail.com