Alveolar Bone Quality Classification from Dental Cone Beam Computed Tomography Images using YOLOv4-tiny

Monica Widiasri[1,2], Nanik Suciati[1], Chastine Fatichah[1], Eha Renwi Astuti[3], Ramadhan Hardani Putra[3], and Agus Zainal Arifin[1]

[1] Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[2] Department of Informatics, Universitas Surabaya, Surabaya, Indonesia
[3] Department of Dentomaxillofacial Radiology, Universitas Airlangga, Surabaya, Indonesia
monicawidiasri.207025@mhs.its.ac.id,@monica@staff.ubaya.ac.id

Abstract. Bone quality is essential in dental implant planning for successful implant placement. Bone quality can be determined based on bone density observed from Beam Computed Tomography (CBCT) images which are commonly used in dental implant planning. The most accepted classification of alveolar bone quality is that proposed by Lekholm and Zarb (1985), classifying bone into four types based on the density of cortical and trabecular bone observed from CBCT images. Currently, determining the type of alveolar bone in the implant area depends on the clinician's subjectivity. This study uses deep learning to propose an alveolar bone quality classification system from CBCT images. The YOLOv4-tiny method, a detection and classification method with excellent performance and fast training time, was used to detect and classify alveolar bone from 2D dental CBCT images of mandibular coronal slices. The results of bone quality classification yielded a mean precision value of 99.91%. The study findings indicate that YOLOv4-tiny can accurately classify alveolar bone density. This precision is essential for proper dental implant placement and implant planning.

Keywords: Alveolar Bone, Bone Quality, Classification, CBCT images, Dental Implant, Detection, YOLO.

1 Introduction

Dental implants are a reliable treatment option to replace missing teeth. Before implant surgery, dental implant planning is carried out by analyzing dental images resulting from radiographic processes. CBCT imaging is a well-established radiographic modality in dental implant treatment planning that is becoming more popular and widely employed in oral health care [1].

The success of dental implants depends on the quantity and quality of the jawbone [2]. An assessment of bone quality is available by looking at the bone density at the implant site. However, due to the limited availability of CT-scan modalities in dental

M. Hartono et al. (eds.), *Proceedings of the 4th International Conference on Informatics, Technology and Engineering 2023 (InCITE 2023)*, Atlantis Highlights in Engineering 21,
https://doi.org/10.2991/978-94-6463-288-0_48

clinics, bone quality classification from Lekholm and Zarb (1985) has been commonly used in daily practice. Lekholm and Zarb classified bone quality into four types according to the cortical and trabecular bone ratio, which can be assessed using CBCT images [3]. Poor bone quality has been consistently reported as one of the main risk factors for implant failure. Therefore, an appropriate assessment of the quality of alveolar bone is needed in planning dental implants to ensure the success of dental implant therapy.

Studies regarding the classification of bone quality from CBCT images using deep learning are still not widely carried out. The first study to evaluate bone density was performed using a 3-D deep convolutional neural network (CNN) on CBCT images. The trabecular pattern of the bone was recognized and classified into four classes according to Misch (D1, D2, D3, and D4). D1 is the most dense bone, and D4 is the least dense. The result of classification accuracy is 95.2% [4]. Studies for bone classification based on Lekholm and Zarb using deep learning on CBCT images have yet to be carried out.

Object detection is identifying objects in an image, drawing bounding boxes around them, and classifying them. A state-of-the-art object detection system, You Only Look Once (YOLO), uses deep learning to detect and classify objects simultaneously [5]. Research on detecting alveolar bone using the YOLO method from CBCT images has been successfully carried out with excellent performance. The alveolar bone detection results from CBCT 2D grayscale images using the YOLOv3-tiny method yield a mean average precision (mAP) of 98.6% [5]. Meanwhile, the Dental-YOLO method, an efficient version of YOLOv4 specifically developed to detect alveolar bone and mandibular canal, produces an average precision of 99.37% in detecting alveolar bone [6].

This study proposes a classification of alveolar bone quality from dental CBCT images using the YOLOv4-tiny method. YOLOv4-tiny is one of the YOLOv4-based lightweight YOLO series methods proven to produce excellent and fast detection performance [7]. The efficient version of YOLOv4 has been shown to detect alveolar bone with impressive performance [6]. Therefore, the YOLOv4-tiny method is suitable for use in this study which performs the alveolar bone detection process. The YOLOv4-tiny model will detect alveolar bone from the 2D dental CBCT image input, then the bone type and the classification confidence value are displayed as the result of alveolar bone classification. The results of alveolar bone classification can be used to help clinicians in dental implant treatment planning according to the type of alveolar bone.

2 Methodologies

2.1 Alveolar bone quality classification

The alveolar bone quality is an essential factor that must be considered for a successful dental implant [3]. A grading method devised by Lekholm and Zarb to define the relationship between cortical and trabecular bone can be used to determine the quality of alveolar bone [2]. There are two systems of bone quality classification proposed by

Lekholm and Zarb (1985). The first classification, the type of bone quality, is classified into four types, as depicted in Fig. 1 [8]. The second classification, three new classes, is added to the first classification, subclasses of types 2 and 3 [9].

This study uses the first bone quality classification system. The classification of bone quality consists of four types, namely [9]:

- Type 1: Entirely homogenous cortical bone
- Type 2: Thick layer of cortical bone surrounding a core of dense trabecular bone
- Type 3: Thin layer of cortical bone surrounding a core of dense trabecular bone
- Type 4: Thin cortical bone layer surrounding a low-density trabecular bone core.

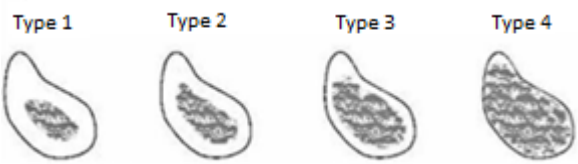


Fig. 1. The classification of bone quality according to Lekholm and Zarb [8].

The quantity of cortical bone is essential for the implant's main stability, whereas trabecular bone is crucial for long-term stability [8]. Although type 1 ensures optimal implant stability, research has indicated that types 2 and 3 have the most excellent-long-term outcomes, whereas type 4 results in the most frequent premature implant loss.

2.2 YOLOv4-tiny

Object detection involves the task of localizing and classifying some of the objects that may be present in an image. Two steps are taken in object detection; the first is to find the object's location marked with a bounding box, and the second is to classify the object in the bounding box into the appropriate class [10].

You Only Look Once (YOLO) is a high-accuracy and high-performance detection and classification technique. YOLO's architecture comprises three parts: the backbone, the neck, and the head [6]. The YOLOv4-tiny technique is built on the YOLOv4 approach and has a smaller architectural size, resulting in faster training time and object detection speed. It uses the CSPDarknet53-tiny backbone network instead of the YOLOv4 CSPDarknet53 backbone network [11]. The CSPDarknet53-tiny network employs the CSPBlock module in a cross-stage partial network. The CSPBlock module divides the feature map into two parts and joins them using a cross-stage residual edge. The CSP-Block module can improve the convolution network's learning capability to improve accuracy.

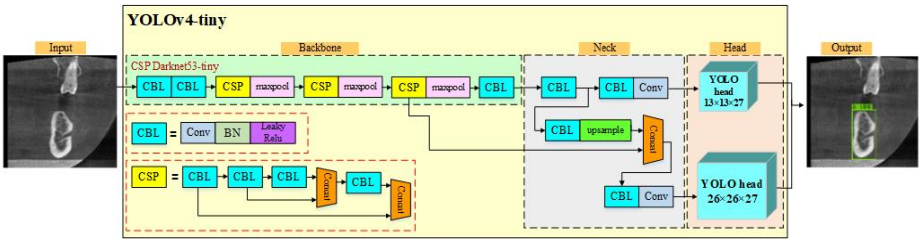


Fig. 2. YOLOv4-tiny architecture.

To speed up processing, the YOLOv4-tiny approach makes use of the CSPDarknet53-tiny network's LeakyReLU activation function. The YOLOv4-tiny technique uses a feature pyramid network to extract feature maps with two scaled feature maps, 13 × 13 and 26 × 26. Fig. 2 depicts the YOLOv4-tiny architecture for recognizing and classifying alveolar bone types using 2D dental CBCT image input from coronal slices.

3 Result of Experiment

3.1 Dataset

This study used 2D CBCT dental images of mandibular coronal slices from 26 patients. Image obtained from Airlangga University Surabaya Dental and Oral Hospital Education, with ethical clearance certificate number 617/HRECC.FODM/V/2023. The 626 images representing the alveolar bone in the implant area are used for system training and testing. Two dental radiologists from Airlangga University Surabaya Dental and Oral Hospital Education determined the ground truth of the alveolar bone type for each image. After that, the ground truth image is annotated with the labelImg annotation tool. The alveolar bone was annotated with bounding boxes and the quality type (type 1/2/3/4). The total number of alveolar bone annotations created was 767, including 102 type 1, 381 type 2, 196 type 3, and 88 type 4 annotations.

Table 1. The number of annotations for each alveolar bone type

	Train set				Test set			
Dataset	Type 1	Type 2	Type 3	Type 4	Type 1	Type 2	Type 3	Type 4
1	82	301	159	73	20	80	37	15
2	71	269	138	62	31	112	58	26

Two tests were performed to evaluate the performance of the YOLOv4-tiny model by splitting the dataset into two distinct proportions. The first dataset (dataset1) is split into 80% train set and 20% test sets, while the second dataset (dataset2) is split into 70% and 30%. Table 1 shows the number of annotations for each alveolar bone type in each dataset.

3.2 Metrics for Evaluation

The classification performance was evaluated using precision, recall, average precision, mean average precision, and F1-score.

The precision (P) is the ratio of positive samples (TP) among predicted positive samples (TP + FP). as shown in Equation (1) [12].

$$P = \frac{TP}{TP+FP} \times 100\% \tag{1}$$

The recall (R) ratio is a ratio of correctly predicted positive samples (TP) to labeled positive samples (TP + FN), as shown in Equation (2).

$$R = \frac{TP}{TP+FN} \times 100\% \tag{2}$$

The average precision (AP) summarizes a precision-recall curve as the weighted mean of precisions at each threshold, with the recall increase from the preceding threshold used as the weight, as shown in Equation (3) [13], where R_n and P_n are the precision and recall at the *n*th threshold.

$$AP = \sum_n (R_n - R_{n-1}) \times P_n \tag{3}$$

The mean Average Precision (mAP) is the mean of AP of each class, as shown in Equation (4) [5]. The mAP is calculated by finding AP for each class and then averaging over some classes (N).

$$mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i \tag{4}$$

Equation (5) indicates that the F1-score is the harmonic mean of precision and recall. The F1-score is a measure of the precision-to-recall ratio. When the F1-score is high, it indicates that precision and recall are both strong. A lower F1-score indicates more significant imbalance in precision and recall.

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{5}$$

3.3 Alveolar Bone Classification Performance

The detected alveolar bone is marked with a bounding box. Class type classification results are marked by different colors on the bounding box, and class type labels (1/2/3/4) are written on the top of the bounding box, followed by the confidence score of the alveolar bones are classified to that type. An example of the results of alveolar bone quality classification for four types of alveolar bone quality can be seen in Fig. 3.

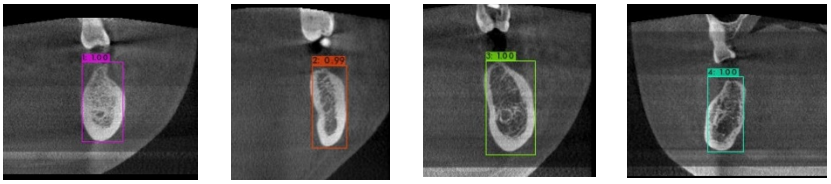


Fig. 3. Classification results of four types of alveolar bone quality.

The YOLOv4-tiny detection model is implemented using the Darknet framework. The settings in the YOLOv4-tiny training configuration file are batch size 64, subdivision 64, number of classes 4, so max_batches is 8000, step 6400 and 7200, and filter 27 all YOLO layers. YOLOv4-tiny trained weights were used to make inferences on the test images.

YOLO only displays detected objects with a confidence factor equal to or greater than a specific threshold value. In this study, experiments were conducted for two thresholds: 0.25 (the default threshold utilized by YOLO) and 0.5.

The test results using the best weight for the classification of each type of alveolar bone quality can be seen in Table 2. From the test results, alveolar bone type 1 achieved the best classification results with an AP of 100% in all datasets and thresholds. The worst classification results were in alveolar bone type 3, where the number of FPs was the largest in both the dataset and threshold.

Table 2. Classification test results for each type of alveolar bone quality

	Dataset1						Dataset2					
	Threshold 0.25			Threshold 0.50			Threshold 0.25			Threshold 0.50		
Type	AP (%)	TP	FP	AP (%)	TP	FP	AP (%)	TP	FP	AP (%)	TP	FP
1	**100**	20	1	**100**	20	1	**100**	31	0	**100**	31	0
2	99.69	78	1	99.69	76	0	99.98	112	1	99.98	111	1
3	99.72	37	8	99.72	37	6	99.80	58	6	99.80	57	3
4	**100**	15	0	**100**	13	0	99.86	25	0	99.86	25	0

Table 3 shows the classification performance of YOLOv4-tiny on datasets 1 and 2. Dataset2, which splits training and testing data by 70% and 30%, produces a slightly better classification performance than dataset1, which divides data by 80% and 20%. The classification performance of dataset1 produces the best mAP is 99.91%, whereas dataset2 is 99.85%.

Table 3. The alveolar bone quality classification performance in dataset1 and dataset2

	Threshold 0.25				Threshold 0.5			
Dataset	mAP (%)	F1-score	P	R	mAP (%)	F1-score	P	R
1	99.85	0.96	0.94	0.99	99.85	0.96	0.95	0.96
2	**99.91**	**0.98**	**0.97**	**1**	**99.91**	**0.98**	**0.98**	**0.99**

4 Discussion

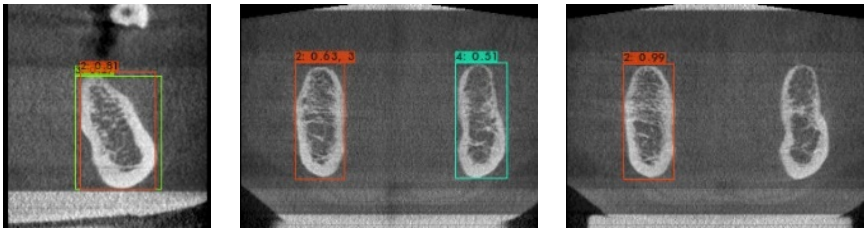
Alveolar bone from 2D dental CBCT image input coronal slices of the mandible can be detected very accurately using YOLOv4-tiny. In addition, the YOLOv4-tiny method also produces excellent performance for classifying alveolar bone quality into four types, according to Lekholm and Zarb. The confidence score for most classification findings is 100% or close to 100%. The mAP is generated for all experiments by changing the proportion of data, and the threshold reaches 99%.

YOLOv4-tiny predicts detection results using two feature maps of various scales, 13 × 13 and 26 × 26. Because alveolar bone objects can be seen clearly and are relatively big, these two feature maps are suitable for classifying alveolar bone and produce excellent classification results. As a result, a three-scale feature map, as utilized in YOLOv4, is unneeded, because a third feature, 52 × 52, is better suited for recognizing tiny objects.

The YOLOv4-tiny method can classify alveolar bone with the best mAP of 99.91%. As shown in Table 3, the resultant precision is lower than the recall. It is because the number of false positives (FP) is higher than the false negatives (FN), especially for the detected type 3 alveolar bone. FP mainly occurs in the alveolar bone, which is classified into multiple types. As shown in Table 3, the higher the threshold, the higher the precision value. This implies that the FP is decreasing. However, the recall is decreasing as the FN grows larger; that is, there is an alveolar bone that cannot be detected.

Figures 4a and 4b show examples of FP. The left alveolar bone in Fig. 4a was classified as type 2 bone with an 81% confidence factor and type 3 bone with a 27% confidence factor. Because the confidence factor type 2 and type 3 are more than 0.25, two left alveolar bone bounding box outputs are produced if the threshold value is 0.25, as shown in Fig. 4a. However, if the threshold value is 0.5, just one output, bone type 2, is created. The left alveolar bone in Fig. 4a is type 2 bone, resulting in a false positive for type 3 bone classification. In Fig. 4b, the left alveolar bone has a confidence factor of 63% for type 2 and 55% for type 3. Two bounding boxes will be presented with a threshold of 0.25 and 0.5 since the confidence factor value for both classes is larger than the threshold. However, Fig. 4b displays just one bounding box on the left alveolar bone since the location of the bounding box for type 2 and type 3 is almost similar. There is also a false positive for type 3 bone in this case since the ground truth of the left alveolar bone is type 2 bone. Figure 4c depicts a false negative. The right alveolar bone was not identified as an alveolar bone. The ground truth alveolar bone on the right is type 4. The alveolar bone on the right has a confidence

factor of type 4 of 15%. Because the confidence factor is less than the threshold (both 0.25 and 0.5 were employed in the experiment), the alveolar bone on the right in Fig. 4c was not correctly recognized as the alveolar bone.



a b c

Fig. 4. Examples of missclassification results.

The CBCT scan shows that greater density alveolar bones seem whiter, whereas lower density alveolar bones appear darker. Most type 1 alveolar bone is white, indicating entirely homogeneous cortical bone. Meanwhile, quality type 4 is white with a thin border of alveolar bone surrounding the darker bone in the center. Types 1 and 4 bones can be distinguished more easily, however types 2 and 3 bones may not be as easily distinguished. The majority of false positives in the studies occurred in alveolar bone, which was classified as type 2 and type 3. Variations in cortical and trabecular bone can contribute to differences in subjective classification [9]. By taking into consideration all possible combinations of cortical and trabecular bone and being able to reduce false positives of bone quality classification types 2 and 3, it is necessary to develop a bone quality classification system based on the revised version of bone quality classification from Lekholm and Zarb.

5 Conclusion

The YOLO4-tiny detection and classification method successfully detected and classified bone quality based on the bone quality classification proposed by Lekholm and Zarb with outstanding performance. The performance of alveolar bone classification into four types of bone quality achieves the best mean average precision of 99.91%. These excellent outcomes have the potential to enable automated bone quality evaluation in dental implant planning. To improve the determination of bone quality types 2 and 3, a bone quality classification system can be developed using the newer version of the YOLO method and the revised version of the bone quality classification system by Lekholm and Zarb.

References

1. Jacobs R, Salmon B, Codari M, et al (2018) Cone beam computed tomography in implant dentistry: Recommendations for clinical use. BMC Oral Health 18:1–17. https://doi.org/10.1186/s12903-018-0523-5
2. Juodzbalys G, Kubilius M (2013) Clinical and Radiological Classification of the Jawbone Anatomy in Endosseous Dental Implant Treatment. J Oral Maxillofac Res 4:. https://doi.org/10.5037/jomr.2013.4202
3. Hao Y, Zhao W, Wang Y, et al (2014) Assessments of jaw bone density at implant sites using 3D cone-beam computed tomography. Eur Rev Med Pharmacol Sci 18:1398–1403
4. Sorkhabi MM, Saadat Khajeh M (2019) Classification of alveolar bone density using 3-D deep convolutional neural network in the cone-beam CT images: A 6-month clinical study. Measurement (Lond) 148:106945. https://doi.org/10.1016/j.measurement.2019.106945
5. Widiasri M, Arifin AZ, Suciati N, et al (2021) Alveolar Bone Detection from Dental Cone Beam Computed Tomography using YOLOv3-tiny. AIMS 2021 - International Conference on Artificial Intelligence and Mechatronics Systems. https://doi.org/10.1109/AIMS52415.2021.9466037
6. Widiasri M, Arifin AZ, Suciati N, et al (2022) Dental-YOLO: Alveolar Bone and Mandibular Canal Detection on Cone Beam Computed Tomography Images for Dental Implant Planning. IEEE Access 10:101483–101494. https://doi.org/10.1109/ACCESS.2022.3208350
7. Wang S, Zhao J, Ta N, et al (2021) A real-time deep learning forest fire monitoring algorithm based on an improved Pruned + KD model. J Real Time Image Process. https://doi.org/10.1007/s11554-021-01124-9
8. Bernaerts A, Vanhoenacker FM, Chapelle K, et al (2006) The role of dental CT imaging in dental implantology. J Belge Radiol 89:32–42
9. Al-Ekrish A, Widmann G, Alfadda S (2018) Revised, Computed Tomography–Based Lekholm and Zarb Jawbone Quality Classification. Int J Prosthodont 31:342–345. https://doi.org/10.11607/ijp.5714
10. Sarma A (2022) Introduction to the YOLO Family. https://pyimagesearch.com/2022/04/04/introduction-to-the-yolo-family/. Accessed 5 Jun 2023
11. Jiang Z, Zhao L, Shuaiyang LI, Yanfei JIA (2020) Real-time object detection method for embedded devices. ArXiv 3:1–11
12. Chen J, Ma A, Huang L, et al (2023) GA-YOLO: A Lightweight YOLO Model for Dense and Occluded Grape Target Detection. Horticulturae 9:. https://doi.org/10.3390/horticulturae9040443
13. Pedregosa F, Varoquaux G, Gramfort A, et al (2011) Scikit-Learn: Machine Learning in Python. J Mach Learn Res 12:2825–2830