Oral Radiological Indices Calculation Based on SAM

Yupeng Du, Qiang Qu, Cong Zhang, Xu Wang, Guowen Kuang, Mengna Wen, Man Liu, Jinfeng Yang, Fengyu Liang, and Panpan Yuan

***Abstract*—Oral radiological indices are important non-invasive markers for assessing disease risk. However, current related work is relatively scarce and most methods rely on manual calculations, lacking precision and efficiency. This study aims to improve the performance of oral radiological index calculations using the Segment Anything Model (SAM). A method based on SAM was developed with a modified use of a point-prompt approach to improve the robustness and efficiency of the segmentation. Also, a dental panoramic radiograph dataset, $ORIC340$, was constructed and annotated for fine-tuning. Finally, the method was tested on $ORIC340$ and a publicly available dental panoramic radiograph dataset from $Kaggle$. The experimental results showed that errors of the proposed method do not exceed 8.75% on dataset $ORIC340$ and 5.88% on a dataset from $Kaggle$ for Panoramic Mandibular Index(PMI). The method's robustness was also demonstrated, while reducing calculation time per image from 90 seconds manually to 7 seconds. By integrating SAM with medical image processing, our method offers a robust and efficient solution for calculating oral radiological indices, enhancing early disease risk assessment. This method provides a reliable tool for clinical practice, enhancing the radiological index calculations with better performance, low user requirements and convenient implementation.**

***Index Terms*—Oral radiological index, SAM, Point prompt, Deep learning.**

Manuscript received March 3, 2025; revised June 24, 2025

Yupeng Du is a postgraduate student of the Institute of Applied Artificial Intelligence of the Guangdong-Hong Kong-Macao Greater Bay Area, Shenzhen Polytechnic University, PR China (e-mail: 1004962884@qq.com).

Qiang Qu is a professor of the School of Electronic and Information Engineering, University of Science and Technology Liaoning, PR China (e-mail: quqiang@ustl.edu.cn).

Cong Zhang is an attending physician of the School of Medical Technology and Nursing, Shenzhen Polytechnic University, PR China (e-mail:congpaopao1988@szpu.edu.cn).

Xu Wang is a vice researcher of the Institute of Applied Artificial Intelligence of the Guangdong-Hong Kong-Macao Greater Bay Area, Shenzhen Polytechnic University, PR China (corresponding author to provide phone: +86 18038161609; fax: 26018406; e-mail: wangxu@szpu.edu.cn).

Guowen Kuang is a vice researcher of the Institute of Applied Artificial Intelligence of the Guangdong-Hong Kong-Macao Greater Bay Area, Shenzhen Polytechnic University, PR China (e-mail:gkuang@szpu.edu.cn).

Mengna Wen is an assistant researcher of the Institute of Applied Artificial Intelligence of the Guangdong-Hong Kong-Macao Greater Bay Area, Shenzhen Polytechnic University, PR China (e-mail:WMN251718@163.com).

Man Liu is an associate chief physician of the School of Medical Technology and Nursing, Shenzhen Polytechnic University, PR China (e-mail:liumandentist@126.com).

Jinfeng Yang is a professor of the Institute of Applied Artificial Intelligence of the Guangdong-Hong Kong-Macao Greater Bay Area, Shenzhen Polytechnic University, PR China (e-mail:jfyang@szpu.edu.cn).

Fengyu Liang is a postgraduate student of the Institute of Applied Artificial Intelligence of the Guangdong-Hong Kong-Macao Greater Bay Area, Shenzhen Polytechnic University, PR China (e-mail:fengyuliang@stu.ustl.edu.cn).

Panpan Yuan is a postgraduate student of the Institute of Applied Artificial Intelligence of the Guangdong-Hong Kong-Macao Greater Bay Area, Shenzhen Polytechnic University, PR China (e-mail:2532849624@qq.com).

I. Introduction

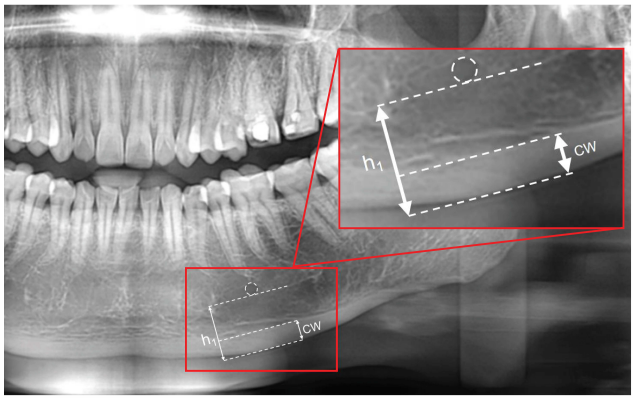

Fig. 1. The definition of oral radiological index PMI

Fig. 2. The definition of MRI

DENTAL panoramic radiograph plays an important role in contemporary dental care, providing intricate details of dental structures, including bone loss and lesions not apparent in conventional oral examinations [1, 2]. This advanced imaging technique is crucial for the early detection of cysts, tumors, and jawbone abnormalities, significantly improving the likelihood of successful treatments [3]. Also, dental panoramic radiographs have been used to predict low bone mineral density in patients [4–6]. Specifically, a number of oral radiological indices, including the Mandible Resorption Index (MRI)and Panoramic Mandibular Index (PMI), have been developed to assess and quantify the quality of mandibular bone mass and resorption for diagnose of osteopenia [7–15].

PMI is the ratio of the thickness (CW) of the mandibular cortical bone to the distance (h_1) from the mental foramen to the edge of the mandibular cortical bone, as shown in Fig. 1. MRI is the ratio of the distance (H) from the upper border of the mandible in the mental foramen area to the distance (h_2) between the center of the mental foramen and the lower edge of the mandible, as shown in Fig. 2. PMI and MRI can reflect the bone density of the human body, which is valuable for the early screening of osteoporosis [16].

Various methods have already been developed for the calculation of these indices. In 2006, Arifin et al. developed a computer-aided system to measure mandibular cortical width (MCW) on panoramic radiographs to aid diagnosing osteoporosis in postmenopausal women [17]. In 2012, a support vector machine-based computer-aided diagnosis system was developed by Kavitha et al. to detect osteoporosis using dental panoramic radiographs [18]. Also, in 2019, Muramatsu et al. developed a method for measuring mandibular cortical width on dental panoramic radiographs [19]. In 2020, Aliaga et al. developed an automated method based on fuzzy K-means classification and texture analysis to compute mandibular indices in dental panoramic radiographs [20]. To our knowledge, this is the latest work on calculating mandibular cortical indices. It employed a series of traditional image processing techniques, including fuzzy K-means clustering, texture analysis-based tangent line adjustments, multi-threshold binarization, morphological operations and feature description methods to compute relevant oral radiological indices. This method allows for the automatic identification and measurement of mandibular features to calculate the mandibular cortical width (MCW), Panoramic Mandibular Index (PMI) and mandibular ratio (M/M). However, the robustness of these methods depend on the quality of the input images with massive hyperparameter adjustment, which hinder the widespread application in routine clinical practice. Related work remains scarce, and the data and code are not publicly available.

In other side, deep learning has shown immense potential in various medical image processing. Precise segmentation would provide accurate volumetric and morphological information of the target structures, thereby supporting a wide range of subsequent clinical applications [21–34]. Furthermore, with the advent of large-scale pre-trained models [35] and its corresponding revolutionary shift (zero-shot and few-shot generalization capabilities) in the field of artificial intelligence, many foundational models have emerged and been applied to medical image analysis. In 2022, a study developed a model to identify human activities and assess hazard levels in hospital environments using deep learning models like YOLOv2, VGG16, and MobileNetv2 SSD. The VGG16 model achieved the highest accuracy (93.33%) in classifying hazardous activities. The research highlighted the need for more data to improve model performance in real-world settings [36]. This approach aligns with our work in leveraging deep learning for medical applications, where accurately classifying and assessing risks from radiographic images can similarly benefit from robust, data-driven models. Wong et al. conducted experiments using the ScribblePrompt interactive segmentation tool on a variety of biomedical imaging datasets, providing rapid and flexible annotations. However, it is more suitable for annotation instead of application in clinical or research settings [37]. Ma et al. conducted experiments using the MedSAM model on diverse medical imaging datasets, demonstrating its high accuracy and robustness across various tasks. However, MedSAM's performance in oral imaging did not demonstrate the high performance found in other medical datasets, almost failing to segment complete teeth in dental panoramic radiographs [38].

Among these pre-trained models, SAM [39] is a versatile segmentation model with remarkable zero-shot generalization ability, capable of generating masks of various granularities through point prompts, mask prompts, box prompts, or text prompts. It consists of three core components: the image encoder, the prompt encoder, and the mask decoder [39]. The image encoder adopts a Vision Transformer (ViT) [40] pre-trained with Masked Autoencoder (MAE) [41], emphasizing scalability and robustness. The prompt encoder processes two types of prompts: sparse (such as points, boxes, text) and dense (mask). In this work, the point prompt is adopted to improve segmentation accuracy, which is easy and convenient to use in practice. The mask decoder utilizes Transformer [42–45] decoder modules to generate mask, by leveraging prompt self-attention and cross-attention mechanisms. However, numerous studies indicate a significant performance gap between SAM and current state-of-the-art models within medical field. Roy et al. conducted experiments on the AMOS22 abdominal CT organ segmentation dataset [46], and the results show that the performance of SAM was inferior to SOTA models. Furthermore, according to literature [47], the experiments indicate that SAM could achieve better performance with prompts, while there remains a substantial performance gap compared to the classic U-Net architecture [48]. One prime reason for this gap is the lack of proper fine-tuning of SAM model or the absence of specialized medical training dataset [49]. No pre-trained model (e.g. SAM) based method for calculating oral radiological indices has been found yet.

To address these challenges, a SAM-based method is developed. Also, a dental panoramic radiograph dataset $ORIC340$ was built and annotated with PMI and MRI value, including corresponding masks, i.e. the masks of the mental foramen, the second premolar, the alveolar bone adjacent to the second premolar, and the mandibular cortical bone for fine-tuning and verification. Besides, the way of using prompt point is modified, with the decoder of SAM being fine-tuned to improve performance of medical tasks and ease of use. Finally, the proposed method is tested on datasets $ORIC340$ and $Kaggle$ [50].

Our contributions can be summarized as follows:

1) A method was developed for calculating oral radiological indices by integrating SAM with medical image processing.
2) Integrating human-in-loop with SAM through modified point prompt, i.e. only one single click in the assigned feature region, which improves the efficiency, accuracy and reliability of oral radiological indices calculation.
3) Corresponding post-processing is designed based on the segmentation output of SAM and the oral radiological index's definition.
4) The dataset $ORIC340$ of dental panoramic radiographs was built and annotated with PMI and MRI

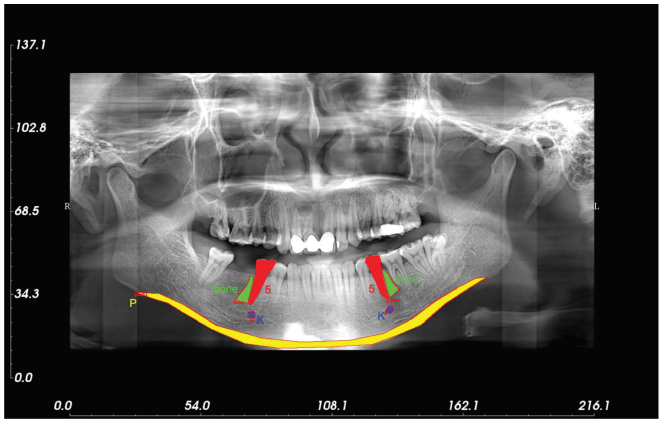

Fig. 3. Annotated dental panoramic radiograph with color-coded masks

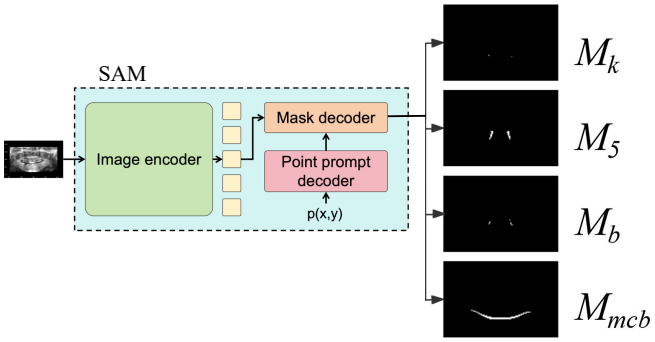

Fig. 4. The SAM-based method for dental panoramic radiograph segmentation for PMI and MRI

values and their corresponding masks, which is used to fine-tune SAM.

II. Material and Methods

A. Dataset

To verify the effectiveness of the proposed method and fine-tune SAM, a dental panoramic radiograph dataset $ORIC340$ was built. It contains 340 dental panoramic radiographs, annotated by several radiologists. For each dental panoramic image, four feature region masks are annotated, i.e. the (left and right) second premolar mask (red), the cortical bone mask (yellow), the (left and right) mental foramen mask (blue), and the area adjacent to the second premolar mask (green), as shown in Fig. 3. Also, the corresponding groundtruth of PMI and MRI are calculated by hand. The dataset is divided into training set with 323 dental panoramic radiograph images and test set with 17 images.

Furthermore, the public dental panoramic radiographs dataset on $Kaggle$ was downloaded and annotated by hand to verify the effectiveness of proposed method. The annotation is the same as $ORIC340$, as shown in Fig. 3. After removing the samples with unclear mental foramen or no mental foramen, the dataset contains 104 samples, which is also divided into training set with 86 samples and test set with 18 samples. However, due to the Kaggle dataset containing many samples with missing teeth which makes it impossible to locate the second premolar, only PMI is calculated.

B. Framework

The framework of the proposed method is shown in Fig. 4, which is comprised of three main components: the image encoder, the point prompt encoder and the mask decoder.

The image encoder utilizes the structure of SAM's encoder, including corresponding weights, which are fixed during training. This is for two reasons. Firstly, SAM has been pre-trained on large-scale datasets, its powerful image encoder is capable of extracting feature from dental panoramic radiographs. Secondly, this encoder can be directly applied to calculating other oral radiological indices.

Note that the point prompt is different from the original version of SAM. In SAM, one point prompt only corresponds to one region, i.e., one mask. In contrast, the point prompt in this method corresponds to all regions that required by oral radiological index, which means that only one point prompt is needed to calculate oral radiological index. Take PMI for example, the prompt point is selected from the mental foramen region. The point corresponds not only to the mental foramen region, but also to the mandibular cortical bone region. Although the prompt point is outside mandibular cortical bone regions, their relative positions remain fixed, which makes it feasible to predict the masks for these regions only by one outside prompt point. Furthermore, the mental foramen region is not always well caught by photography equipment, and these low-quality images without clear mental foramen region could be filtered out during this prompt procedure by users. In Fig. 4, $p(x, y)$ represents the point coordinates corresponding to the user's click on the mental foramen.

As mentioned above, the mask decoder of the proposed method is trained to generate masks for the feature regions required by oral radiological index with only a single point prompt. The output channel dimension N of the mask decoder is equal to the categorie number of the index feature regions, e.g. $N = 2$ for PMI index. To calculate other indices, only the output channel dimension N of the mask decoder need to be modified, with its weight being trained or fine-tuned. In this work, N is set to 4 for calculating both PMI and MRI, where the output masks are the second premolar mask M_5, the cortical bone mask M_{mcb}, the mask of the feature area adjacent to the second premolar M_b, and the mental foramen mask M_k, respectively, as shown on the right side of Fig. 4. Note that only the mask decoder is trained during the fine-tuning to accurately segment key regions from medical images, while the encoder is kept fixed.

The detailed procedure to calculate the oral radiological index is as follows:

First, the output dimension N is determined by the categories of the feature regions according to the definition of the oral radiological index.

Second, choose a critical feature region and set it as the prompt region from which the point prompt is picked. Here, a point prompt is for all the regions related with the oral radiological index, in comparison with SAM where one point prompt only corresponds to one region.

Then, SAM segments the input image and generates the masks of regions based on the point prompt.

Finally, the oral radiological index is calculated by the region masks according to its definition in the post-processing.

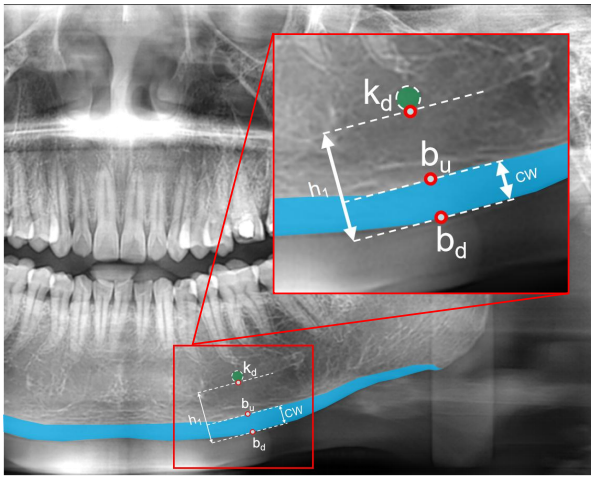

Fig. 5. The calculation of PMI with mask segmentation.

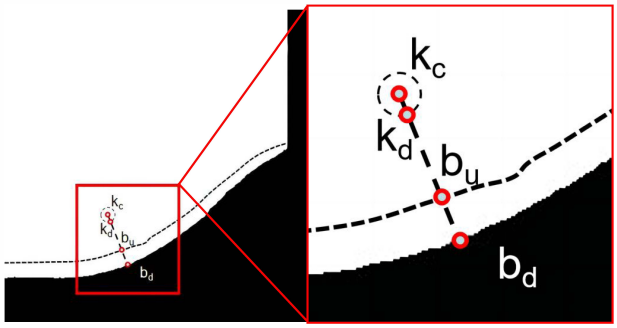

Fig. 6. Identify the inferior edge of the mandibular cortical bone

C. Post-processing

With the output masks of the feature regions, the index could be calculated according to its definition in the post-processing. Considering that the definitions are different from each other, the post-processing is specific for each oral radiological index. For instance, to calculate PMI, the post-processing is as follows:

1) Employ erosion and dilation operation to polish the masks M_k and M_{mcb}(the green area and blue area in Fig. 5). To filter out noise, the kernels are 5×5 and 3×10 for M_k and M_{mcb} respectively. Then, the largest connected region is selected as the feature region.
2) Obtain the center point k_c on the mask M_k of the mental foramen region and turn each pixel point above the mask $mask_{mcb}$ into white. Find out the point b_d in the black area A_{black} in Fig. 6, satisfying

$$b_d = argmax_{p \in A_{black}} ||p - k_c||_2 \tag{1}$$

where $||p - k_c||_2$ is the Euclidean distance between points p and k_c. Thus, the obtained b_d is on the edge of lower bound of $mask_{mcb}$ with minimal distance to k_c, as shown in Fig. 6.
3) Line from k_c to b_d intersecting with the boundary of $mask_k$ and $mask_{mcb}$ to obtain the two boundary points k_d and b_u, as shown in Fig. 6.
4) Calculate h_1, i.e. the distance between the lower edge of the mandible and the lower edge of the mental foramen, by $h_1 = ||k_d - b_d||_2$. Calculate CW, i.e. the thickness of the cortical bone by $CW = ||b_u - b_d||_2$.
5) Finally, obtain the value of PMI by $PMI = CW/h_1$.

D. Loss Function

The Dice-CE loss [51–53] is employed for the training of the mask decoder. The Dice-CE loss contains Dice loss (Equ. 2) and CE loss (Equ. 3),

$$loss_{Dice} = 1 - \frac{2 \times \ TP}{2 \times \ TP + \ FP + \ FN} \tag{2}$$

where TP, FP and FN represents true positives, false positives and false negatives respectively, according to the mask ground truths.

$$loss_{CE} = -\left(y \log(p) + (1 - y) \log(1 - p)\right) \tag{3}$$

where $y \in \{0, 1\}$ represents the ground truth label for oral image segmentation. The label 1 indicates that a pixel belongs to a target area such as the mandibular cortex, mental foramen, etc., while 0 signifies that a pixel is not part of the target area. $p \in [0, 1]$ denotes the probability predicted by the model that a pixel belongs to the target area.

Dice-CE loss(Equ. 4) is as follows:

$$loss_{Dice-CE} = \alpha \times loss_{Dice} + (1 - \alpha) \times loss_{CE} \tag{4}$$

where $\alpha \in [0, 1]$ is used to balance the contributions of Dice loss and CE loss in training.

Thus, the total loss $loss_{total}$ is as follows:

$$loss_{total} = \sum_{i=1}^{N} \alpha_i \cdot loss^i_{Dice-CE} \tag{5}$$

where N is the number of output channel, i.e. the number of the feature mask categories, and $N = 4$ for calculating PMI and MRI. In this work, all α_i is set to 1 during training.

III. Results

To test and verify the effectiveness of the proposed method, three experiments were conducted. Initially, the proposed method was tested on the dental panoramic radiograph dataset $ORIC340$. Then, the robustness for point prompt was tested(shown in Section IV). Finally, an open-source dataset from Kaggle was relabeled and used to verify the performance of the proposed method. The results were only compared with the human labels. The results of the model segmentation are visualized in Fig. 7. The example panoramic dental image shows that the mandibular cortical bone and mental foramens to calculate MPI are accurately segmented. A comparison of the results before and after fine-tuning is shown in Fig. 8. It is evident that the model output after fine-tuning using the proposed method has a significant visual improvement over the original model.

The experiments were executed on a PC with an Nvidia RTX 3080 GPU(10G). Excluding the 2 seconds required to upload the image, our method only takes 7 seconds per image to compute the indices. In comparison, manual calculation takes as long as 90 seconds per image on average.

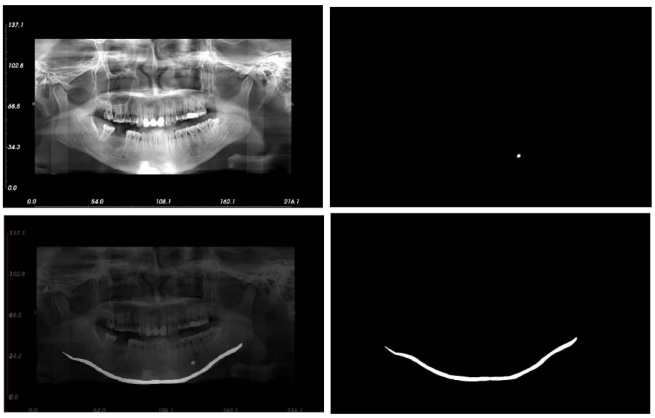

Fig. 7. Visualization of feature region segmentation results for calculating PMI

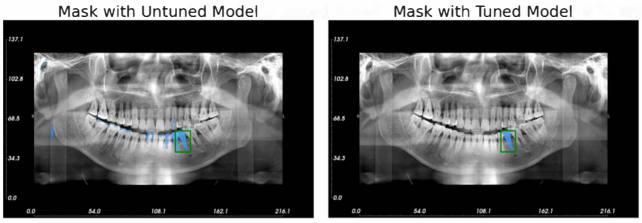

Fig. 8. Comparison and visualization of segmentation effects before and after fine-tuning

A. Evaluation metrics

In this work, the relative error is used to evaluate the performance of the proposed method, as shown in Equ. 6.

$$\text{error} = \frac{y_{label} - y_{output}}{y_{label}} \times 100\% \tag{6}$$

here, y_{label} represents the human label, and y_{output} is the output of the method.

B. Experiment results

The experiment result on Dataset $ORIC340$ is shown in Tab. I. Besides the value of PMI and MRI, Tab. I also gives out the value of the distance h_1 from the lower edge of the mental foramen to the lower edge of the cortical bone, the thickness CW of the cortical bone, the distance h_2 from the center of the mental foramen to the lower edge of the cortical bone, and the distance H from the apex of the alveolar bone of the second premolar to the lower edge of the cortical bone, as shown in the definitions of PMI and MRI(Fig. 1 and 2). These are helpful for the analysis of the performance of the proposed method. Furthermore, all experiments in this work use the pixel as unit.

Also, the proposed method is tested on the dataset $Kaggle$ as mentioned above, and the result for calculating PMI is shown in Tab. II. It is obvious that with point prompt and fine-tune, the performance of the proposed method is satisfactory from one dataset to another dataset. In summary, the proposed method could provide feasible estimation of the oral radiological indices under practical conditions.

Additionally, we tested the robustness of our system by rotating the test samples from the ORIC dataset by 5 degrees, as shown in Table VI. It can be observed that, despite the rotation of the samples, the errors in the results calculated by the system remain within a relatively small range. This demonstrates that the system maintains its accuracy even when the input images undergo slight transformations

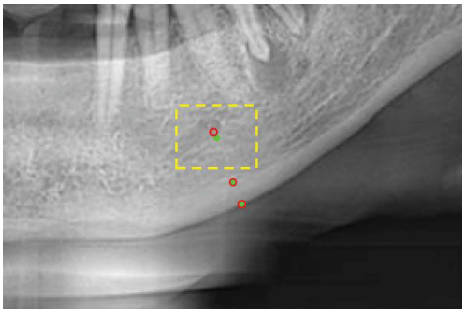

Fig. 9. The comparison between the feature points and their labels of the test sample S15 for calculating PMI

IV. Discussion

In this study, by integrating the pre-trained model SAM, a method was developed to calculate oral radiological indices through a simple click on the mental foramen. Compared with traditional manual methods, the proposed method significantly improves efficiency and accuracy, verified by the experiments. However, there are certain limitations in the proposed method. For instance, the performance of the model may be affected by the quality of the input data, such as the clarity of the mental foramen in the training set, which could impact the actual reference value of the PMI [54]. Also, the method has to be fine-tuned when being applied to a new dataset, and the post-process depends on the definition of the oral radiological indices.

In Tab. I, the proposed method could provide valid estimation of PMI and MRI. The maximum error is 8.75% in calculating PMI for the sample S15. In sample S15, the value of CW is 16.155, very close to label value 16.54, while h_1 value is 48.104, smaller than the label value 53.56, as shown in Fig. 9. In Fig. 9, the red round circles are hand labels, and the green points are generated by our proposed method. The error comes mainly from the difference between the point k_d and its label, as shown in the yellow box in Fig. 9, which is due to the unclear boundary of the structure of the mental foramen. Such error also appears in calculating MRI for sample S22, as shown in Fig. 10. In practice, this kind of error seems inevitable even by hand, where the value differs from person to person.

Additionally, some test samples were marked as "None" in the Tab. I due to the absence of the key feature teeth, making it impossible to give out value for MRI. In practical usage, these dental panoramic radiographs would be considered as fault and filtered out by hand in the prompt procedure. Then, the error of the $Kaggle$ dataset is shown in Fig. 11. It is easy to find that the blurred mandibular edge leads to calculation errors. Finally, to evaluate the impact of the coordinates of the prompt point on the accuracy of the feature index, an experiment is conducted on an image with a clearly visible mental foramen from $ORIC340$. 10 points, shown in Fig. 12, were randomly picked from the the foramen region as the point prompts to calculate PMI, and their results are shown in Tab. III.

TABLE I
THE EXPERIMENT RESULT ON DATASET $ORIC340$ FOR CALCULATING ORAL RADIOLOGICAL INDICES PMI AND MRI.

number	Ground Truth						Inference						Error	
	CW	h_1	h_2	H	PMI	MRI	CW	h_1	h_2	H	PMI	MRI	$error_{PMI}$	$error_{MRI}$
Z2	21.63	58.09	64.6	162.88	0.372	2.521	21.471	59.481	67.535	159.665	0.361	2.364	-3.06%	-6.23%
Z24	23.01	66.47	73.1	168.18	0.346	2.301	20.248	59.237	70.937	166.571	0.342	2.348	-1.26%	2.06%
Z27	18.96	82.76	92.05	165.06	0.229	1.793	18.788	78.791	88.640	166.027	0.238	1.873	4.09%	4.46%
Z32	20.35	76.82	84.17	175.33	0.265	2.083	21.024	75.584	82.292	176.664	0.278	2.147	5.00%	3.06%
Z43	16.76	56.53	62.37	151.33	0.296	2.426	17.464	54.571	63.789	None	0.320	None	7.94%	None
Z53	23.65	63.11	68.27	168.33	0.375	2.466	22.361	56.365	67.543	169.523	0.397	2.510	5.86%	1.79%
Z56	21.19	64.03	70.79	161.58	0.331	2.283	19.235	61.814	77.201	166.760	0.311	2.160	-5.97%	-5.36%
Z58	23.84	70.91	76.90	None	0.336	None	23.259	65.307	75.153	178.941	0.356	2.381	5.94%	None
Z60	27.39	61.09	66.92	164.19	0.448	2.454	25.000	59.481	65.803	165.130	0.420	2.509	-6.26%	2.28%
Z61	17.96	52.19	57.92	147.25	0.344	2.542	15.652	46.957	58.138	159.656	0.333	2.746	-3.14%	8.02%
S15	16.54	53.56	59.97	124.98	0.309	2.084	16.155	48.104	58.873	127.224	0.336	2.161	8.75%	3.69%
S16	20.45	62.94	70.16	None	0.325	None	18.788	59.933	68.877	None	0.313	None	-3.52%	None
S22	22.79	50.62	59.03	128.78	0.450	2.182	23.409	49.041	59.808	144.312	0.477	2.413	6.03%	10.60%
S29	17.62	67.88	74.37	157.85	0.249	2.122	18.028	66.731	75.273	161.307	0.270	2.143	4.08%	0.96%
S35	14.07	82.53	91.24	196.31	0.170	2.152	13.000	73.539	85.615	198.540	0.177	2.319	3.69%	7.78%
S40	21.23	64.07	72.79	158.83	0.331	2.182	19.925	63.246	70.520	160.963	0.315	2.283	-4.92%	4.61%
S44	18.78	69.14	75.06	144.73	0.272	1.928	19.235	67.082	None	None	0.287	None	5.57%	None

TABLE II
THE RESULT OF PMI ESTIMATION ON DATASET $Kaggle$

name	p(x,y)	Ground Truth			Inference			Loss error
		CW	h1	PMI	CW	h1	PMI	PMI
2	[1982, 972]	47.449	182.483	0.260	47.424	182.082	0.260	0.17%
8	[2023, 1054]	52.947	176.930	0.299	51.971	183. 371	0.283	-5.29%
10	[1922, 970]	49.978	171.817	0.291	51.923	171.292	0.303	4.21%
14	[1038, 948]	35.518	113.592	0.313	37.483	113.600	0.330	5.53%
18	[1021, 1055]	51.72	127.350	0.406	50.804	119.050	0.427	5.08%
24	[1873, 1013]	52.342	157.027	0.333	53.151	156.461	0.340	1.91%
25	[1862, 975]	38.982	117.000	0.333	40.000	118.000	0.339	1.74%
26	[1884, 912]	48.659	124.753	0.390	48.662	124.615	0.390	0.12%
31	[1160, 983]	64.548	173.081	0.373	44.283	112.752	0.393	5.31%
33	[1166, 997]	69.442	166.704	0.417	66.098	168.585	0.392	-5.88%
34	[1953, 927]	45.909	167.913	0.273	45.706	166.904	0.274	0.16%
50	[1938, 852]	49.098	187.432	0.262	48.166	187.601	0.257	-1.99%
67	[1079, 851]	37.992	116.011	0.327	38.000	116.000	0.328	0.03%
70	[1764, 893]	38.102	132.890	0.287	37.108	133.154	0.279	-2.80%
93	[2111, 936]	52.452	150.319	0.349	55.579	162.779	0.341	-2.15%
107	[2111, 1059]	57.016	154.993	0.368	56.000	156.000	0.359	-2.42%
111	[973, 920]	49.263	164.669	0.299	50.804	165.560	0.307	2.57%
112	[1905, 932]	40.026	163.240	0.245	38.639	163.957	0.236	-3.89%

TABLE III
THE EXPERIMENT RESULT WITH 10 RANDOM PROMPT POINTS IN DENTAL PANORAMIC RADIOGRAPH S29 FOR CALCULATING PMI.

name	p(x,y)	Ground Truth			Inference		
		CW	h1	PMI	CW	h1	PMI
S29	[887,642]	17.62	67.88	0.260	18.028	67.067	0.269
S29	[887,643]					67.067	0.269
S29	[890,645]					67.067	0.269
S29	[896,643]					66.731	0.270
S29	[889,644]					67.067	0.269
S29	[894,641]					66.731	0.270
S29	[893,642]					66.731	0.270
S29	[890,641]					67.067	0.269
S29	[892,640]					66.731	0.270
S29	[890,644]					67.067	0.269

It is obvious that the value of PMI is nearly not affected by the exact value of the prompt point's coordinate at all, which indicates that the proposed method can be applied conveniently in practice, with only one human action, i.e.,

TABLE IV
EXPERIMENTAL RESULTS ON THE CALCULATION TIME OF INDICATORS FOR EACH ORAL PANORAMA ON THE $ORIC340$ DATASET.

name	p(x,y)	Ground Truth PMI	Inference PMI	Loss error	time	total time	average time	time variance
Z2	[550, 687]	0.372	0.361	-3.06%	4.958			
Z24	[861, 696]	0.346	0.342	-1.26%	4.031			
Z27	[873, 685]	0.229	0.238	4.09%	3.942			
Z32	[526, 685]	0.265	0.278	5.00%	3.389			
Z43	[862, 689]	0.296	0.320	7.94%	3.619			
Z53	[539, 661]	0.375	0.397	5.86%	2.938			
Z56	[543, 680]	0.331	0.311	-5.97%	4.070			
Z58	[892, 688]	0.336	0.356	5.94%	3.246			
Z60	[881, 653]	0.448	0.420	-6.26%	4.069	63.884	3.757	0.403
Z61	[870, 692]	0.344	0.333	-3.14%	2.883			
S15	[838, 595]	0.309	0.336	8.75%	3.800			
S16	[547, 601]	0.325	0.313	-3.52%	2.598			
S22	[848, 642]	0.450	0.477	6.03%	4.457			
S29	[891, 640]	0.260	0.270	4.08%	4.407			
S35	[526, 639]	0.170	0.177	3.69%	4.269			
S40	[526, 669]	0.356	0.315	-4.92%	3.219			
S44	[550, 626]	0.359	0.287	5.57%	3.989			

TABLE V
EXPERIMENTAL RESULTS ON THE CALCULATION TIME OF INDICATORS FOR EACH ORAL PANORAMA ON THE $Kaggle$ DATASET.

name	p(x,y)	Ground Truth PMI	Inference PMI	Loss error	time	total time	average time	time variance
2	(1982,972)	0.260	0.260	0.17%	4.412			
8	[2023,1054]	0.299	0.283	-5.29%	4.251			
10	[1922,970]	0.291	0.303	4.21%	4.619			
14	[1038,948]	0.313	0.330	5.53%	4.290			
18	[1021,1055]	0.406	0.427	5.08%	4.430			
24	[1873,1013]	0.333	0.340	1.91%	5.228			
25	[1862,975]	0.333	0.339	1.74%	4.269			
26	[1884,912]	0.390	0.390	0.12%	4.222			
31	[1160,983]	0.373	0.393	5.31%	3.493	76.711	4.265	0.146
33	[1166,997]	0.417	0.392	-5.88%	3.827			
34	[1953,927]	0.273	0.274	0.16%	4.310			
50	[1938,852]	0.262	0.257	-1.99%	4.020			
67	[1079,851]	0.327	0.328	0.03%	4.138			
70	[1764,893]	0.287	0.279	-2.80%	4.279			
93	[2111,936]	0.349	0.341	-2.15%	3.622			
107	[2111,1059]	0.368	0.359	-2.42%	4.430			
111	[973,920]	0.299	0.307	2.57%	4.433			
112	[1905,932]	0.245	0.236	-3.89%	4.498			

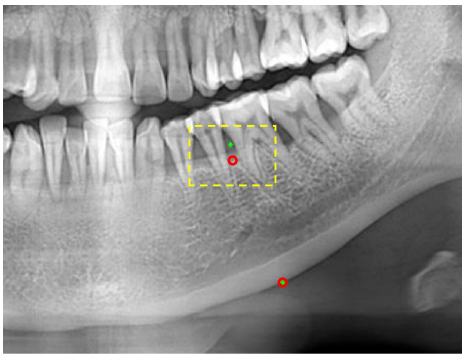

Fig. 10. The comparison between the feature points and their labels of the test sample S22 for calculating MRI

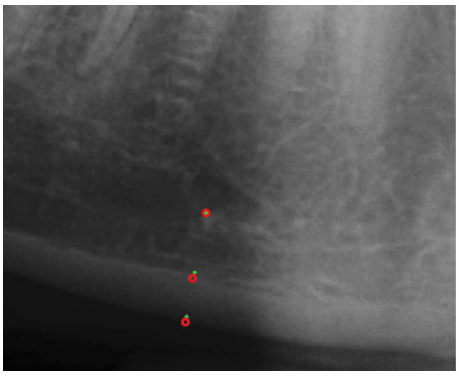

Fig. 11. The comparison between the feature points and their labels of the test sample 18 for calculating PMI

one click in the foramen region.

We also recorded the time taken to calculate the PMI for samples from the Kaggle dataset and the ORIC340 dataset, as shown in Tables V and IV. It can be observed that our system takes approximately 3 to 4 seconds to calculate the Oral Radiological Indices, while the traditional manual calculation method takes about 60 seconds per image. This clearly demonstrates that the efficiency of our proposed method significantly outperforms that of the manual approach.

TABLE VI
EXPERIMENTAL RESULTS OF CALCULATING PMI FOR $ORIC$ DATASET AFTER ROTATION

name	p(x,y)	Ground Truth			Inference			Loss error
		CW	h1	PMI	CW	h1	PMI	PMI
Z2	[857,678]	21.63	58.09	0.372	21.954	60.902	0.360	-3.19%
Z24	[534,688]	23.01	66.47	0.346	20.248	63.285	0.320	-7.57%
Z27	[867,682]	18.96	82.76	0.229	18.028	79.712	0.226	-1.28%
Z32	[528,690]	20.35	76.82	0.265	20.125	75.133	0.268	1.11%
Z43	[538,680]	16.76	56.53	0.296	19.698	68.447	0.288	-2.93%
Z53	[875,668]	23.65	63.11	0.375	21.190	56.303	0.376	0.43%
Z56	[550,672]	21.19	64.03	0.331	19.235	61.984	0.310	-6.23%
Z58	[555,686]	23.84	70.91	0.336	21.954	64.140	0.342	1.81%
Z60	[883,656]	27.39	61.09	0.448	25.000	59.203	0.422	-5.82%
Z61	[557,709]	17.96	52.19	0.344	17.720	51.923	0.341	-0.83%
S15	[588,595]	16.54	53.56	0.309	14.142	50.359	0.281	-9.06%
S16	[551,602]	20.45	62.94	0.325	17.889	59.481	0.301	-7.44%
S22	[551,663]	22.79	50.62	0.450	20.396	46.872	0.435	-3.35%
S29	[891,640]	17.62	67.88	0.260	16.763	62.642	0.268	3.09%
S35	[873,646]	14.07	82.53	0.170	13.000	77.078	0.169	-1.07%
S40	[526,664]	21.23	64.07	0.331	20.881	63.600	0.328	-0.92%
S44	[544,631]	18.78	69.14	0.272	19.235	67.082	0.287	5.57%

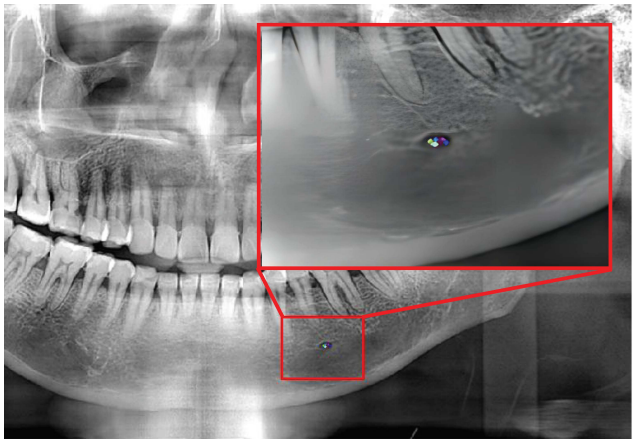

Fig. 12. 10 random points picked from the right mental foramen of the dental panoramic radiograph S29

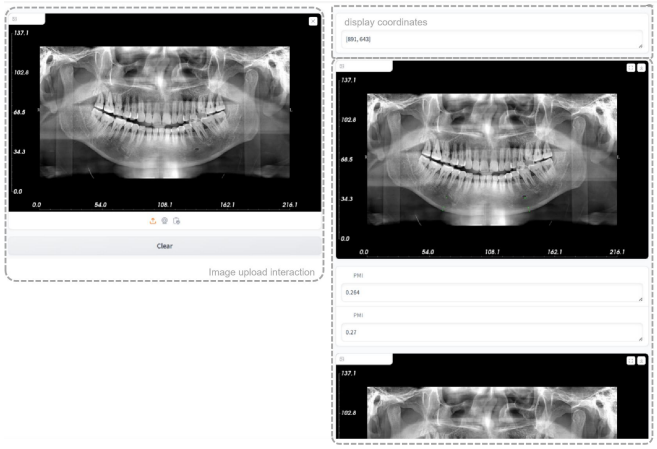

Fig. 13. Front-End Interface

V. ENGINEERING IMPLEMENTATION

The system adopts a client-server separation architecture, with integrated front-end and back-end development implemented through the Gradio framework. The front-end module constructs the interactive interface using Gradio's UI components, including an image upload window , an interactive image annotation area, and a result display panel, as shown in Fig. 13. The image upload window (left box in Fig. 13) arranges the upload button supporting both drag-and-drop upload. The right-up area in Fig. 13) is the corresponding interactive annotation area for click annotation by doctors or other users, and the results,including segmentation masks and the oral radiological indices are given in a result display panel (right-down box in Fig. 13). Further, the system support both PNG and DICOM files, where pydicom library is applied to extract metadata. Then, the DICOM data is normalized and scaled to 1024×1024 by OpenCV, eliminating device differences that may affect model inference.

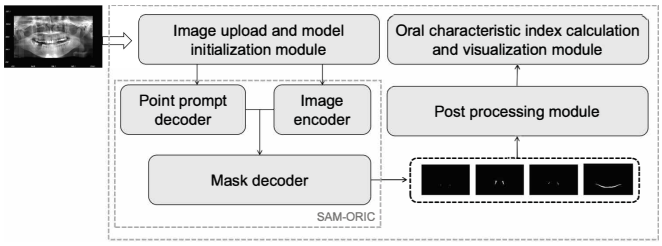

Fig. 14. Oral index intelligent calculation system framework

The system architecture is shown in Fig. 14, where the back-end module consists the image encoder, the prompt encoder, the mask decoder and post-processing. All these back-end functions runs in Python, and the @cache decorator is used to optimize model singleton loading, with the local storage module utilizing the tempfile library to manage temporary data. The image encoder is the pretrained SAM model with frozen weight, while the prompt encoder converts

user-clicked positions into point prompt vectors in the format [[x, y, 1]] through coordinate mapping. Then, the fine-tuned mask decoder outputs four kinds of masks (for the premolar, mandibular cortical bone, second premolar, and its adjacent regions) in the result display panel. Finally, after the oral radiological indices being calculated and shown in the result display panel, all the input data is automatically destroyed to ensure zero retention of patient privacy data. The system avoids the I/O bottleneck of databases and compresses the processing time for a single image to under 7 seconds, about 2 to 4 times faster than cloud services.

VI. Conclusion

In this work, a SAM-based method is developed to assist the calculation of feature indices for dental panoramic radiographs. The results on datasets $ORIC340$ and $Kaggle$ show that this approach significantly improves the efficiency of calculating oral indices, while being robust against users, offering an efficient and reliable solution for the early screening and assessment of oral diseases. The future work will extend this method to handle 3D data, such as CBCT, which could relief the limitations, such as the clarity of the mental foramen depending on the perspective of the camera. All source codes and models will be publicly available after the acceptance of this paper.

References

[1] V. Tsapaki, "Radiation protection in dental radiology – recent advances and future directions," *Physica Medica*, vol. 44, pp. 222–226, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S112017971730234X

[2] V. Geethasree, C. Sai Swapna Sri, V. Sravani, K. Bhaskari, and P. Manne, "A survey on image processing techniques for detection of cavities in dental x-ray images," *Proceedings of Fourth International Conference on Communication, Computing and Electronics Systems*, V. Bindhu, J. M. R. S. Tavares, and C. Vuppalapati, Eds. Singapore: Springer Nature Singapore, 2023, pp. 805–814.

[3] L. T. Kawamoto, W. O. Kawamoto, A. Formigoni, E. F. Rodrigues, I. P. de Arruda Campos, and S. C. M. Rodrigues, "Quality comparison of analog and digital x-ray equipment and materials in a dental clinic," *Key Engineering Materials*, vol. 660, pp. 330–334, 2015. [Online]. Available: https://doi.org/10.4028/www.scientific.net/KEM.660.330

[4] K. Seki, M. Nagasaki, T. Yoshino, M. Yano, A. Kawamoto, and O. Shimizu, "Radiographical diagnostic evaluation of mandibular cortical index classification and mandibular cortical width in female patients prescribed antiosteoporosis medication: A retrospective cohort study," *Diagnostics*, vol. 14, no. 10, 2024. [Online]. Available: https://www.mdpi.com/2075-4418/14/10/1009

[5] E. Hastar, H. H. Yilmaz, and H. Orhan, "Evaluation of mental index, mandibular cortical index and panoramic mandibular index on dental panoramic radiographs in the elderly," *European Journal of Dentistry*, vol. 5, no. 01, pp. 060–067, 2011.

[6] L. Munhoz, L. Morita, A. Y. Nagai, J. Moreira, and E. S. Arita, "Mandibular cortical index in the screening of postmenopausal at low mineral density risk: a systematic review," *Dentomaxillofacial Radiology*, vol. 50, no. 4, p. 20200514, 02 2021. [Online]. Available: https://doi.org/10.1259/dmfr.20200514

[7] A. Taguchi, M. Tsuda, M. Ohtsuka, and et al., "Use of dental panoramic radiographs in identifying younger postmenopausal women with osteoporosis," *Osteoporosis International*, vol. 17, no. 3, pp. 387–394, 2006. [Online]. Available: https://doi.org/10.1007/s00198-005-2029-7

[8] Y. Ariji, A. Katsumata, R. Kubo, and et al., "Factors affecting observer agreement in morphological evaluation of mandibular cortical bone on panoramic radiographs," *Oral Radiology*, vol. 33, no. 2, pp. 117–123, 2017. [Online]. Available: https://doi.org/10.1007/s11282-016-0253-9

[9] D. Preiss, P. Welsh, S. Murphy, J. H. Waters, D. Demicco, C. Cannon, M. Sabatine, S. Seshasai, and P. Barter, "Osteoporosis prevention, diagnosis, and therapy," *Journal of the American Medical Association*, vol. 285, no. 6, pp. 785–795, Feb 2001. [Online]. Available: https://doi.org/10.1001/jama.285.6.785

[10] L. G. Chandak, V. K. Lohe, R. R. Bhowate, K. P. Gandhi, and N. V. Vyas, "Correlation of periodontitis with mandibular radiomorphometric indices, serum calcium and serum estradiol in postmenopausal women: A case–control study," *Indian Journal of Dental Research*, vol. 28, no. 4, pp. 388–394, Jul–Aug 2017. [Online]. Available: doi.org/10.4103/ijdr.IJDR_532_16

[11] M. Tofangchiha, M. Khorasani, M. Shokrimozhdehi, and A. Javadi, "Diagnosis of osteoporosis using cortex mandibular indices based on cortex thickness and morphology in comparison with visual assessment of the cortex," *Journal of Craniomaxillofacial Research*, vol. 4, no. 2, pp. 345–351, Jul. 2017. [Online]. Available: https://jcr.tums.ac.ir/index.php/jcr/article/view/154

[12] E. Calciolari, N. Donos, J. Park, A. Petrie, and N. Mardas, "Panoramic measures for oral bone mass in detecting osteoporosis: A systematic review and meta-analysis," *Journal of Dental Research*, vol. 94, no. 3_suppl, pp. 17S–27S, 2015, pMID: 25365969. [Online]. Available: https://doi.org/10.1177/0022034514554949

[13] A. Taguchi, M. Sanada, E. Krall, T. Nakamoto, M. Ohtsuka, Y. Suei, K. Tanimoto, I. Kodama, M. Tsuda, and K. Ohama, "Relationship between dental panoramic radiographic findings and biochemical markers of bone turnover," *Journal of Bone and Mineral Research*, vol. 18, no. 9, pp. 1689–1694, 12 2009. [Online]. Available: https://doi.org/10.1359/jbmr.2003.18.9.1689

[14] A. Taguchi, Y. Suei, M. Sanada, M. Ohtsuka, T. Nakamoto, H. Sumida, K. Ohama, and K. Tanimoto, "Validation of dental panoramic radiography measures for identifying postmenopausal women with spinal osteoporosis," *American Journal of Roentgenology*, vol. 183, no. 6, pp. 1755–1760, 2004, pMID: 15547223. [Online]. Available: https://doi.org/10.2214/ajr.183.6.01831755

[15] A. Taguchi, K. Tanimoto, Y. Suei, K. Ohama, and T. Wada, "Relationship between the mandibular and lumbar vertebral bone mineral density at different postmenopausal stages," *Dentomaxillofacial Radiology*, vol. 25, no. 3, pp. 130–135, 01 2014. [Online]. Available: https://doi.org/10.1259/dmfr.25.3.9084261

[16] B. Cakur, S. Dagistan, A. Sahin, A. Harorli, and A. Yilmaz, "Reliability of mandibular cortical index and mandibular bone mineral density in the detection of osteoporotic women," *Dentomaxillofacial Radiology*, vol. 38, no. 5, pp. 255–261, 01 2014. [Online]. Available: https://doi.org/10.1259/dmfr/22559806

[17] A. Arifin, A. Asano, A. Taguchi, and et al., "Computer-aided system for measuring the mandibular cortical width on dental panoramic radiographs in identifying postmenopausal women with low bone mineral density," *Osteoporosis International*, vol. 17, pp. 753–759, 2006. [Online]. Available: https://doi.org/10.1007/s00198-005-0045-2

[18] M. Kavitha, A. Asano, A. Taguchi, and et al., "Diagnosis of osteoporosis from dental panoramic radiographs using the support vector machine method in a computer-aided system," *BMC Medical Imaging*, vol. 12, p. 1, 2012. [Online]. Available: https://doi.org/10.1186/1471-2342-12-1

[19] C. Muramatsu, T. Matsumoto, T. Hayashi, and et al., "Automated measurement of mandibular cortical width on dental panoramic radiographs," *International Journal of Computer Assisted Radiology and Surgery*, vol. 8, pp. 877–885, 2013. [Online]. Available: https://doi.org/10.1007/s11548-012-0800-8

[20] J. Wu, Z. Wang, M. Hong, W. Ji, H. Fu, Y. Xu, M. Xu, and Y. Jin, "Medical sam adapter: Adapting segment anything model for medical image segmentation," *Medical Image Analysis*, vol. 102, p. 103547, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841525000945

[21] J. Ma, Y. Zhang, S. Gu, C. Zhu, C. Ge, Y. Zhang, X. An, C. Wang, Q. Wang, X. Liu, S. Cao, Q. Zhang, S. Liu, Y. Wang, Y. Li, J. He, and X. Yang, "Abdomenct-1k: Is abdominal organ segmentation a solved problem?" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 10, pp. 6695–6714, 2022.

[22] M. Antonelli, A. Reinke, S. Bakas, and et al., "The medical segmentation decathlon," *Nature Communications*, vol. 13, p. 4128, 2022. [Online]. Available: https://doi.org/10.1038/s41467-022-30695-9

[23] J. Wasserthal, H.-C. Breit, M. T. Meyer, M. Pradella, D. Hinck, A. W. Sauter, T. Heye, D. T. Boll, J. Cyriac, S. Yang, M. Bach, and M. Segeroth, "Totalsegmentator: Robust segmentation of 104 anatomic structures in ct images," *Radiology: Artificial Intelligence*, vol. 5, no. 5, p. e230024, 2023. [Online]. Available: https://doi.org/10.1148/ryai.230024

[24] D. Zhang, C. Wang, T. Chen, W. Chen, and Y. Shen, "Scalable swin transformer network for brain tumor segmentation from incomplete mri modalities," *Artificial Intelligence in Medicine*, vol. 149, p. 102788, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0933365724000307

[25] X. Shu, J. Wang, A. Zhang, J. Shi, and X.-J. Wu, "Csca u-net: A channel and space compound attention cnn for medical image segmentation," *Artificial Intelligence in Medicine*, vol. 150, p.

102800, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0933365724000423
[26] Y. Wang, Z. Yang, X. Liu, Z. Li, C. Wu, Y. Wang, K. Jin, D. Chen, G. Jia, X. Chen, J. Ye, and X. Huang, "Pgkd-net: Prior-guided and knowledge diffusive network for choroid segmentation," *Artificial Intelligence in Medicine*, vol. 150, p. 102837, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0933365724000794
[27] N. Athulya, S. Ramachandran, N. George, N. Ambily, and L. Shine, "Enhancing outdoor mobility and environment perception for visually impaired individuals through a customized cnn-based system," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 9, 2023.
[28] R. Sreelatha and L. R. Roopa, "Deep learning-based detection system for heavy-construction vehicles and urban traffic monitoring," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 3, 2022.
[29] S. N. A. Mohd Robi, N. Ahmad, M. A. M. Izhar, H. M. Kaidi, and N. M. Noor, "Utilizing uav data for neural network-based classification of melon leaf diseases in smart agriculture." *International Journal of Advanced Computer Science & Applications*, vol. 15, no. 1, 2024.
[30] E. A. Elatrash, N. E. Ghannam, M. S. Zaki, and R. S. El-Sayed, "Enhancing deep learning-based models for early detection of colon cancer." *IAENG International Journal of Computer Science*, vol. 52, no. 4, pp. 1148–1158, 2025.
[31] S. Kong, J.-S. Wang, and S.-Y. Jin, "Cell nucleus image segmentation method and performance comparison." *IAENG International Journal of Computer Science*, vol. 52, no. 2, pp. 298–306, 2025.
[32] S. Takezaki and K. Kishida, "Construction of cnns for abnormal heart sound detection using data augmentation," *Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists*, 2021, pp. 20–22.
[33] I. Baokar and L. He, "Memristive neural networks application in predicting of health disorders," *Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2023*, 5-7 July, 2023, Hong Kong, pp. 94–99.
[34] N. M. Cong, H.-C. Liu, V. R. Mekala, E. Zaenudin, E. B. Wijaya, and K.-L. Ng, "Identify gene-gene regulatory modules for patients with renal clear cell tumor metastasis," *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering*, vol. 2245, 2023, pp. 31–35.
[35] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever, "Learning transferable visual models from natural language supervision," *Proceedings of the 38th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, M. Meila and T. Zhang, Eds., vol. 139. PMLR, 18–24 Jul 2021, pp. 8748–8763. [Online]. Available: https://proceedings.mlr.press/v139/radford21a.html
[36] K. Shahrim, A. H. Abd Rahman, and S. Goudarzi, "Hazardous human activity recognition in hospital environment using deep learning," *IAENG International Journal of Applied Mathematics*, vol. 52, no. 3, pp. 748–753, 2022.
[37] H. E. Wong, M. Rakic, J. Guttag, and A. V. Dalca, "Scribbleprompt: Fast and flexible interactive segmentation for any biomedical image," Cham, pp. 207–229, 2025.
[38] J. Ma, Y. He, F. Li, and et al., "Segment anything in medical images," *Nature Communications*, vol. 15, p. 654, 2024. [Online]. Available: https://doi.org/10.1038/s41467-024-44824-z
[39] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollar, and R. Girshick, "Segment anything," *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2023, pp. 4015–4026.
[40] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," 2020, https://doi.org/10.48550/arXiv.2010.11929.
[41] K. He, X. Chen, S. Xie, Y. Li, P. Dollár, and R. Girshick, "Masked autoencoders are scalable vision learners," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022, pp. 16 000–16 009.
[42] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc., 2017. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf
[43] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, 2019, pp. 4171–4186. [Online]. Available: https://aclanthology.org/N19-1423/
[44] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever *et al.*, "Improving language understanding by generative pre-training," 2018.
[45] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever *et al.*, "Language models are unsupervised multitask learners," *OpenAI Blog*, vol. 1, no. 8, p. 9, 2019, https://insightcivic.s3.us-east-1.amazonaws.com/language-models.pdf.
[46] T. Wald, S. Roy, G. Koehler, N. Disch, M. R. Rokuss, J. Holzschuh, D. Zimmerer, and K. Maier-Hein, "SAM.MD: Zero-shot medical image segmentation capabilities of the segment anything model," *Medical Imaging with Deep Learning, short paper track*, 2023. [Online]. Available: https://openreview.net/forum?id=iilLHaINUW
[47] C. Hu and X. Li, "When sam meets medical images: An investigation of segment anything model (sam) on multi-phase liver tumor segmentation," 2023, https://doi.org/10.48550/arXiv.2304.08506.
[48] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer International Publishing, 2015, pp. 234–241. [Online]. Available: http://lmb.informatik.uni-freiburg.de/people/ronneber/u-net
[49] J. Wu, Z. Wang, M. Hong, W. Ji, H. Fu, Y. Xu, M. Xu, and Y. Jin, "Medical sam adapter: Adapting segment anything model for medical image segmentation," p. 103547, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841525000945
[50] A. H. Abdi, S. Kasaei, and M. Mehdizadeh, "Automatic segmentation of mandible in panoramic x-ray," *Journal of Medical Imaging*, vol. 2, no. 4, pp. 044 003–044 003, 2015.
[51] X. Li, X. Sun, Y. Meng, J. Liang, F. Wu, and J. Li, "Dice loss for data-imbalanced nlp tasks," 2019, https://doi.org/10.48550/arXiv.1911.02855.
[52] B. Liu, J. Dolz, A. Galdran, R. Kobbi, and I. Ben Ayed, "Do we really need dice? the hidden region-size biases of segmentation losses," *Medical Image Analysis*, vol. 91, p. 103015, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S136184152300275X
[53] A. Galdran, G. Carneiro, and M. A. G. Ballester, "On the optimal combination of cross-entropy and soft dice losses for lesion segmentation with out-of-distribution robustness," *Diabetic Foot Ulcers Grand Challenge*. Cham: Springer International Publishing, 2023, pp. 40–51.
[54] A.-Y. Kwon, K.-H. Huh, W.-J. Yi, S.-S. Lee, S.-C. Choi, and M.-S. Heo, "Is the panoramic mandibular index useful for bone quality evaluation?" *Imaging Science in Dentistry*, vol. 47, no. 2, p. 87, 2017, https://doi.org/10.5624/isd.2017.47.2.87.