Cephalometric Landmarks Identification Through an Object Detection-based Deep Learning Model

Idriss Tafala[1], Fatima-Ezzahraa Ben-Bouazza[2], Aymane Edder[3], Oumaima Manchadi[4],
Mehdi Et-Taoussi[5], Bassma Jioudi[6]
Clinical and Medical Sciences and Biomedical Engineering Laboratory, Mohammed VI University
of Sciences and Health, Casablanca, Morocco - UM6SS.[1,2,3,4,5,6]
Faculty of Science and Technology, Hassan 1^{er} University[2]
LaMSN, La Maison des Sciences Numériques, France[2]

***Abstract*—In the field of orthodontics, the accurate identification of cephalometric landmarks in dental radiography plays a crucial role in ensuring precise diagnoses and efficient treatment planning. Previous studies have demonstrated the impressive capabilities of advanced deep learning models in this particular domain. However, due to the ever-changing technological landscape, it is imperative to consistently investigate and explore emerging algorithms to further improve efficiency in this field. The present study centers around the assessment of the effectiveness of YOLOv8, the most recent version of the 'You Only Look Once (YOLO)' algorithm series, with a particular emphasis on its autonomous capability to accurately identify cephalometric landmarks. In this study, a thorough examination was conducted to evaluate the YOLOv8 algorithm efficiency in detecting cephalometric landmarks. The assessments encompassed various aspects such as precision, adaptability in challenging conditions, and a comparative analysis with alternative algorithms. The predefined proximities of 2mm, 2.5mm, and 3mm were utilized for the comparisons. By focusing on its potential as a noteworthy breakthrough, the investigation seeks to ascertain whether the recent enhancements indeed bring about a significant stride in the precise identification of cephalometric landmarks.**

***Keywords*—Cephalometry; YOLOv8; landmark detection; orthodontics**

I. Introduction

Since its introduction by Broadbent in 1931, the cephalometer has garnered extensive recognition as a universally adopted diagnostic instrument within the realms of orthodontic practice and research [1]. Cephalometry, a field of study focused on the measurement of cranial dimensions, involves the meticulous quantification of these dimensions using either direct measurements or radiographic techniques. This process involves identifying specific anatomical landmarks as reference points for accurate and standardized measurements. The incorporation of suitable standards plays a pivotal role in the assessment of facial growth and development, rendering them an indispensable component in the diagnostic and treatment planning stages of orthodontics. In the realm of cephalometrics, the arduous task of manually identifying and annotating cephalometric landmarks on radiographic images has long been entrusted to proficient experts in the field. Hand annotation, while a valuable technique, is known for its labor-intensive nature and vulnerability to human error-induced discrepancies [2] [3].

The advent of deep learning in the field of medical imaging has brought about a profound transformation in various diagnostic techniques, such as cephalometric analysis. In this emerging era of digital diagnostics, machine learning algorithms have taken the lead, paving the way for significant advancements in orthodontic diagnostic practices and treatment strategies. Prior research has clarified the impressive potential of deep learning models in automating this critical procedure, resulting in significant improvements in both precision and speed.

In the domain of diagnostic automation, deep learning has witnessed remarkable progress. However, the quest for enhanced speed, better precision, and unwavering reliability remains unyielding. Our research is conducted within the context of an ongoing pursuit of continuous improvement.

In this study, we aim to evaluate the capabilities of YOLOv8, the most recent version of the renowned "You Only Look Once (YOLO)" models, in the realm of automated cephalometric landmark detection. This inquiry transcends pure theory and holds significant practical implications for the field of orthodontics. Previous studies have highlighted its potential, but a comprehensive exploration of its performance, especially in diverse clinical scenarios, remains lacking [4].

In our investigation, we delve into the intricate details of YOLOv8, meticulously examining its precision, speed, and robustness metrics. Simultaneously, we remain vigilant in assessing its practicality and suitability for real-world orthodontic clinical scenarios. Our main goal in this study is to find out how well YOLOv8 can bridge the gap between extremely advanced technology and the practical needs of orthodontic practice, setting new standards in the field that have never been seen before.

In this article, we will start by elucidating the underlying impetus behind our research, thereby furnishing a lucid comprehension of the motivation that propelled this work forward. In the subsequent sections, we shall embark on an in-depth examination of pertinent literature, thereby furnishing crucial background information for our research endeavor.

In the sections that follow, we will talk about all of our materials and methods in detail, including the organized way we gathered data, how carefully we prepared it, and how strictly we followed our training procedures. In this section, we will delve into the intricate methods utilized, providing a thorough understanding of the meticulous approach undertaken during the research endeavor.

In the upcoming parts, we will give the results obtained from our thorough inquiry, offering a concise summary of the findings acquired from our analysis of data and experimentation. In the last part of this essay, we will conduct a thorough examination, combining the results, extracting meaningful observations, and possibly suggesting avenues for further research.

II. Motivation

In the field of medical diagnostics, the importance of cephalometry cannot be emphasized enough. The field of craniofacial imaging has played a vital role in the medical industry for many years. Its significance lies in its ability to diagnose craniofacial malformations, facilitate detailed surgical planning and evaluation, and contribute to essential growth studies. Cephalometry, a field of study focused on craniofacial analysis, centers around the meticulous identification of craniofacial landmarks. This crucial process involves the precise detection of cephalometric landmarks on the cephalogram, serving as the fundamental initial stage in conducting any cephalometric analysis [5] [6] [7]

In the past few years, the field of deep learning has witnessed the emergence of highly sophisticated models that have demonstrated exceptional capabilities in this particular domain [8] [9] [10]. The advent of these models has brought about a paradigm shift in the field, presenting ingenious approaches to tackle the complex challenge of landmark identification. In light of the ever-changing technological terrain, it is crucial to consistently delve into and scrutinize nascent algorithms. The need to maximize efficiency and precision fuels the relentless pursuit of advancement in cephalometric analyses.

In our relentless quest for perfection, we embarked on a comprehensive exploration of the most recent breakthroughs in the realm of object detection. In our study, we sought to enhance the accuracy and efficiency of cephalometric landmark identification by harnessing the advanced capabilities of the YOLOv8 algorithm. This algorithm has gained widespread recognition for its exceptional performance in detecting objects. The YOLOv8 model distinguishes itself in the field of computer vision with its remarkable precision and efficiency. The model's exceptional performance is a result of meticulous training on an extensive and varied dataset, ensuring its ability to handle diverse visual scenarios.

The utilization of cephalometry in the medical field is of utmost importance, and with the constant advancements in technology, our investigation of YOLOv8 marks a significant leap forward. Through the utilization of this cutting-edge algorithm, our objective is to not only augment the accuracy of cephalometric landmark detection but also make substantial contributions to the continuous progress in medical diagnostics.

The motivation behind this research was clearly defined in this part. In the following section, we will conduct a thorough examination of the pertinent literature in the topic, providing useful insights into the existing body of knowledge.

III. Related Works

In the field of orthodontics, the convergence of cutting-edge machine learning and computer vision, specifically through the utilization of deep learning techniques, presents a remarkable opportunity for an upheaval. The precise diagnosis of cephalometric landmarks can be significantly improved through the implementation of automated detection techniques [11].

In the realm of orthodontics, the lack of medical imaging data presents an immense barrier. However, the imperative for collaboration between orthodontic professionals and skilled data scientists remains essential. The integration of specialized datasets and advanced techniques holds great potential in bridging the gap between deep learning algorithms and the intricacies of orthodontic imaging. The integration of advanced technologies in dental healthcare not only enhances the efficiency of diagnosis and treatment planning but also serves as a catalyst for innovation in the field [12].

Acknowledging the importance of automatic landmark detection, The ISBI, which stands for the International Symposium on Biomedical Imaging, has led the organization of a number of challenges related to the matter, namely in 2014 [13], in 2015 [14], and is set to continue its impact in the year of 2023 [15].

Over the past few decades, extensive research has been conducted on a multitude of automated techniques aimed at detecting landmarks. In first studies, Wang et al. [13] [14] spearheaded pioneering initiatives that involved the organization of public challenges. These challenges served as platforms to exhibit innovative algorithms that are at the forefront of scientific advancement. Researchers have successfully utilized random forests, a machine learning technique, to classify intensity appearance patterns with remarkable accuracy. Moreover, they have employed statistical shape analysis to gain insights into the complex spatial relationships among landmarks. This innovative approach has yielded impressive results, showcasing the potential of these cutting-edge techniques in various scientific domains.

In another study, Ibragimov and colleagues [16] have presented impressive findings by harnessing the power of Random Forest and Game Theoretic techniques. Researchers such as Chu et al. have successfully utilized tree-based methods in their studies. These methods include hierarchical random forest regression and binary pixel classification with randomized trees [17].

In same vein, Lee et al. and Arik et al. have made significant strides in the field of pixel classification by leveraging convolutional neural network (CNN) concepts. Their work has paved the way for the development of cutting-edge algorithms in this domain [18], [19]. Subsequent studies delved into the realm of deep learning, utilizing U-shaped deep convolutional neural network (CNN) structures to achieve accurate landmark estimation [20] [21].

Mehmet Ugurlu and Alshamrani Khalaf, two researchers, have achieved significant progress in the field of automated cephalometric landmark detection implementing an altered architecture known as the Feature Aggregation and Refinement Network (FAR Net) (Ugurlu, 2022) [22], and an Inception-based neural network layers [23].

In the upcoming chapter, we will conduct a thorough examination of the essential components of our research. This examination will comprehensively investigate the methodological framework that underpins our research. We will primarily concentrate on the specifics of our data gathering procedure, the methodology we have utilized, and the meticulous preparations undertaken for our training data.

IV. Materials

This chapter provides a comprehensive analysis of the complexities involved in our study technique. In the first stage of our study, we will provide a thorough examination of our data collection process, with a detailed explanation of the methods used and the sources of our data. The dataset is thoroughly documented, offering a strong foundation for our upcoming investigations.

After the comprehensive discussion on the process of data acquisition, we now delve into the intricate preparations that were meticulously carried out to enhance the quality of our dataset for the purpose of training. The methodology encompasses various essential components, such as preprocessing techniques, data cleansing methodologies, and necessary transformations implemented to guarantee the dataset's appropriateness for the intended research.

In the culmination of this chapter, we embark upon the fundamental essence of our investigation—the proposed methodology. In this article, we will explore the cutting-edge methodologies and techniques that underpin our research efforts. By the conclusion of this chapter, readers will have acquired a thorough comprehension of the meticulous procedures and approaches that form the foundation of our research methodology.

A. Data Acquisition

In a meticulous and all-encompassing investigation, a collection of lateral cephalograms was procured from a heterogeneous group consisting of 400 individuals. The subjects covered a wide age range, from 7 to 76 years, with an average age of 27.0 years. The sample comprised 235 females and 165 males, ensuring a balanced representation of both genders. The cutting-edge Soredex CRANEXr Excel Ceph machine, situated in Tuusula, Finland, was employed to capture all images in the TIFF format. These images were subsequently processed using the advanced Soredex SorCom software, specifically versions 3.1.5 and 2.0. The captured images exhibited an impressive resolution of 1935×2400 pixels, accompanied by a pixel spacing of 0.1mm, thereby guaranteeing an exceptional level of precision and meticulousness in capturing even the finest details [24]. The dataset utilized in this study, sourced from the ISBI 2015 challenge, consisted of distinct collections of data obtained from a total of 400 subjects. Each subject's data encompassed a lateral cephalogram, as well as two sets of landmark points that were meticulously plotted by skilled orthodontic specialists. Notably, both a junior and a senior specialist contributed to the manual plotting of these landmark points. Significantly, the average intra-observer variability for these specific landmarks was found to be 1.73 mm and 0.90 mm, respectively, indicating a high level of precision [19]. In the captured images, every individual pixel corresponded to a precise 0.1 mm square region. These pixels were characterized by grayscale values spanning from 0 to 255. In order to ensure a good training phase, a total of 150 images were subjected to augmentation techniques, the test1 set was merged to the training set, resulting in an increase to over 900 images. These augmented images were then carefully allocated for rigorous training purposes. Additionally, the remaining 100 images underwent meticulous testing. This comprehensive approach provided a robust assessment of the overall performance of the system. See the annotated image below in Fig. 1.

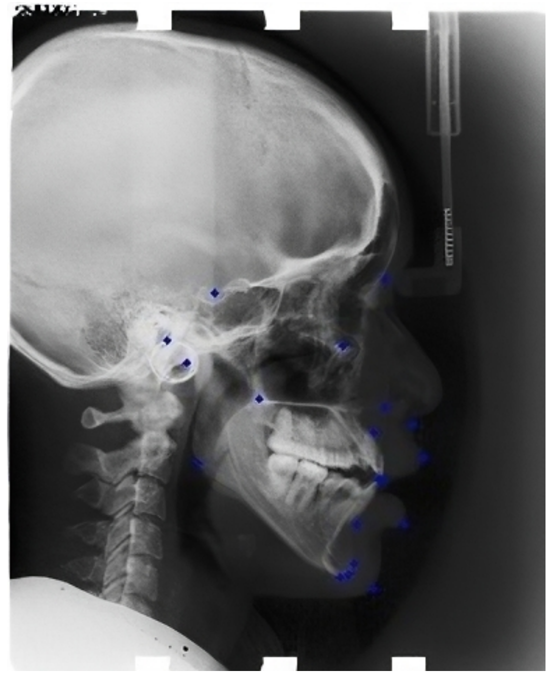

Fig. 1. Annotated image.

B. Data Preparation

In the pursuit of scientific inquiry, we diligently undertook a comprehensive analysis of the data preparation phase, breaking it down into various discrete stages. This approach diverged significantly from the methodologies employed in previous studies. In our study, we departed from traditional methodologies and instead, we embraced a simultaneous detecting approach across all data points.

The dataset being examined, as previously discussed, was burdened with limitations due to its size. In order to overcome these limitations, the incorporation of data augmentation techniques became crucial (Fig. 2). In our study, we have discovered that while mosaic augmentation possesses inherent advantages within the framework of YOLOv8, it unfortunately falls short in meeting our specific requirements. In pursuit of enhancing the dataset, our research endeavors prompted us to extract supplementary images from our preexisting repository of unprocessed visual data.

In the quest for enhancing data quality, we employed three pivotal techniques for augmentation. Color jittering has emerged as a powerful tool that is widely employed in the fields of computer vision and image processing. Through the intentional introduction of controlled randomness into the color attributes, our augmented dataset has successfully attained a heightened level of diversity, in term of saturation,contrast, brightness, and hue. The process of diversification has played a crucial role in enabling machine learning models to adapt and generalize effectively across a wide range of real-world situations. This adaptability is particularly evident in the models' ability to handle changes in lighting conditions and variations in color.

Furthermore, our study carefully utilized Gaussian noise, a fundamental augmentation technique deeply rooted in the fields of machine learning and computer vision. Through the careful manipulation of noise intensity and the introduction of stochastic perturbations into the input data, our augmentation methodology has successfully attained a remarkable level of precision. This includes the ability to accurately account for subtle differences in lighting conditions and address inaccuracies in sensor measurements.

Moreover, we used another technique on the experimental procedure, Random contrast, which involved the manipulation of image contrast through the expansion or compression of the range of pixel intensity values. The process of diversification played a crucial role in effectively addressing the challenge of analyzing images taken under different lighting conditions. By incorporating this approach, our model was able to successfully adapt to and manage the diverse levels of contrast encountered in various real-world situations. The implementation of random contrast adjustment has resulted in a notable improvement in the robustness of our model.

Another crucial aspect of our data preparation process was the transformation of labels. The recorded labels, originally in the form of coordinate dataframes, underwent a transformation to conform to the YOLO annotation format. The utilization of this particular format offers an additional advantage in the form of its inherent normalization feature. The successful outcome of this endeavor led to the smooth incorporation of our image data into the system, eliminating the necessity for additional adjustments such as resizing, scaling, or normalization procedures. The incorporation of this particular aspect has demonstrated itself to be a significant time-saving element within our data preparation procedure.

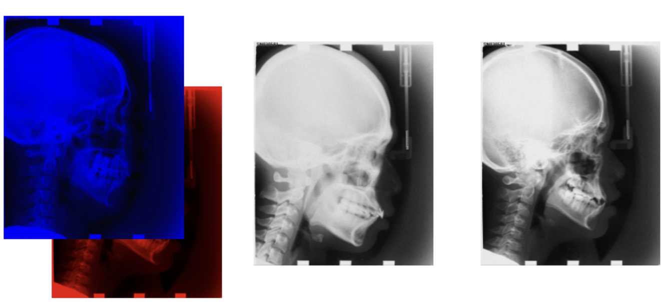

Fig. 2. Data augmentation.

V. Experimental Results

After establishing the underlying motivation behind our research and presenting a thorough summary of the existing literature in the field, we now proceed to explore the core aspects of our study. In this section, we introduce our proposed methodology that centers around leveraging the advanced functionalities of the most recent release of the state-of-the-art object detection algorithm, YOLOv8. Present the experiment's results, and conclude with a comparison with relevant findings from the previous studies.

A. The Proposed Approach

In our study, we adopted YOLOv8 as the base model, employing a meticulous approach to effectively tackle the complexities inherent in our research goals. The decision to utilize this particular choice was driven by its well-documented track record of high efficiency, exceptional accuracy, and remarkable adaptability. These qualities render it an optimal platform for conducting our scientific investigations. In the subsequent sections, we expound upon the intricate intricacies of our methodology, shedding light on the alterations, advancements, and refinements we incorporated to customize YOLOv8 for the precise challenges presented by our research inquiries.

Prior methods frequently employed sliding windows coupled with a classifier, necessitating hundreds or thousands of iterations per image, or employed more refined techniques that divided the task into two steps. The initial phase would identify prospective regions containing objects (referred to as "regions of interest"), and the subsequent phase would evaluate the presence of objects in these proposed regions using a classifier. Our model, only requires a single pass of the network to perform the detection task.

The structure of our model consists of multiple essential components, as illustrated in Fig. 3. The Conv block, also known as the beginning block, is composed of a conv2d layer, batch normalization, and a SiLu activation function. The parameters for this function include the input channels (c1), output channels (c2), kernel size (k), and stride (s). The following block, referred to as the c2f block, comprises Conv blocks in which the generated feature maps are distributed between the bottleneck block and the concat block. The parameters for this block consist of c1, c2, the quantity of shortcuts (n), and a boolean value indicating the utilization of shortcuts. The third block, known as the Spatial Pyramid Pooling Fast (SPPF) block, combines a Convolutional (Conv) block with three Max pooling layers. Significantly, every resultant feature map is merged together prior to the completion of the Spatial Pyramid Pooling Function (SPPF). The parameters accepted by this block include c1, c2, n, and a shortcut indicator. The last component, known as the Detect block, consists of many Conv blocks that have two separate tracks—one for bounding box data and another for class data. The combination of these blocks is detailed in Table I and visually depicted in Fig. 3.

In the field of deep learning research, much attention is typically given to the design of model architecture. However, it is imperative to recognize the significant impact that training procedures. The neural network model used in our study had 295 layers, with a total of 25,867,321 parameters and 25,867,305 gradients. We meticulously adjusted a set of hyperparameters in order to get optimal performance. Mosaic augmentation was activated, which introduced a dynamic element to the training data by merging numerous pictures. To optimize the process, we utilized the AdamW optimizer with a designated learning

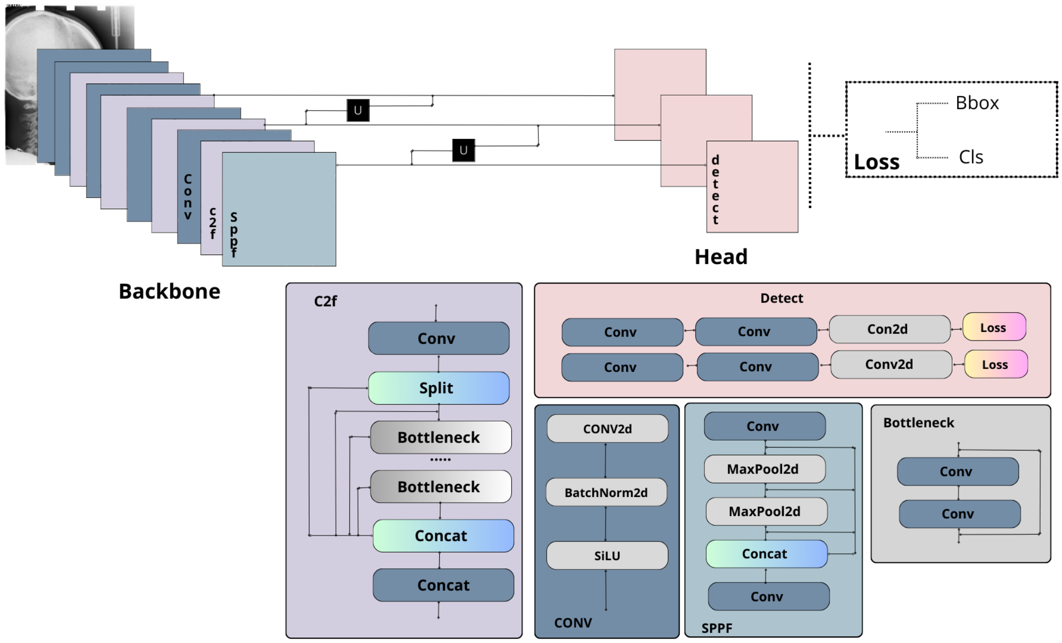

Fig. 3. YOLOv8 Architecture.

TABLE I. MODEL SUMMARY

Params per layer	Module Block	Arguments
1392	ultralytics.nn.modules.conv.Conv	[c1=3,c2= 48, k=3, s=2]
41664	ultralytics.nn.modules.conv.Conv	[c1=48, c2=96, k=3, s=2]
111360	ultralytics.nn.modules.block.C2f	[c1=96, c2=96, n=2, shortcut=True]
166272	ltralytics.nn.modules.conv.Conv	[c1=96, c2=192, k=3, s=2]
813312	ltralytics.nn.modules.block.C2f	[c1=192, c2=192, n=4, shortcut=True]
664320	ultralytics.nn.modules.conv.Conv	[c1=192, c2=384, k=3, s=2]
3248640	ultralytics.nn.modules.block.C2f	[c1=384, c2=384, n=4, shortcut=True]
1991808	ultralytics.nn.modules.conv.Conv	[c1=384, c2=576,k= 3, s=2]
3985920	ultralytics.nn.modules.block.C2f	[c1=576, c2=576, n=2, shortcut=True]
831168	ultralytics.nn.modules.block.SPPF	[c1=576, c2=576, k=5]
0	torch.nn.modules.upsampling.Upsample	[size=None, scale_factor=2, mode='nearest']
0	ultralytics.nn.modules.conv.Concat	[dimension=1]
1993728	ultralytics.nn.modules.block.C2f	[c1=960, c2=384, n=2]
0	torch.nn.modules.upsampling.Upsample	[size=None, scale_factor=2, mode='nearest']
0	ultralytics.nn.modules.conv.Concat	[dimension=1]
517632	ultralytics.nn.modules.block.C2f	[c1=576,c2= 192, n=2]
332160	ultralytics.nn.modules.conv.Conv	[c1=192, c2=192,k= 3, s=2]
0	ultralytics.nn.modules.conv.Concat	[dimension=1]
1846272	ultralytics.nn.modules.block.C2f	[c1=576, c2=384, n=2]
1327872	ultralytics.nn.modules.conv.Conv	[c1=384,c2= 384, k=3, s=2]
0	ultralytics.nn.modules.conv.Concat	[dimension=1]
4207104	ultralytics.nn.modules.block.C2f	[c1=960,c2= 576, n=2]
3786697	ultralytics.nn.modules.head.Detect	[nc=19, [192, 384, 576]]
Summary : 295 layers, 25867321 parameters, 25867305 gradients.		

rate of 0.000435, along with a momentum of 0.9. The training phase consisted of 60 epochs, with batches of size 20, which enhanced the overall resilience of the model.

After presenting a thorough clarification of the fundamental components of our methodology, our attention now shifts towards the results yielded by our endeavors. In the following sections, we shall now proceed to unveil the outcomes derived from the implementation of this particular methodology. The empirical evidence presented not only serves to confirm the strength and reliability of our methodology, but also provides valuable insights into the practical implications of our research in real-world scenarios. By conducting a meticulous examination and subsequent interpretation, we aim to elucidate the profound implications of these discoveries, effectively establishing a connection between theoretical knowledge and its real-world implementation.

B. Results

A distinctive feature of our research is the smooth integration of both unprocessed and meticulously enhanced data throughout the training stage of our YOLOv8 model. The combination of traditional imaging methods with augmented data has transformed the sector, presenting thrilling opportunities for enhanced precision and efficiency, ultimately resulting in notable progress in landmark detection. Our algorithm correctly recognized 19 anatomical landmarks on lateral cephalometric radiographs, proving its efficiency.

In order to assess the system's performance, we employed the Successful Detection Rate (SDR) score, a vital metric for quantifying its level of success. The SDR metric quantifies the accuracy of predicted landmarks by measuring the percentage of these landmarks that fall within a predetermined threshold distance from the ground truth.

Our system demonstrated exceptional performance in terms of average Successful Detection Rate (SDR) scores, operating within the specified range of 2 mm, 2.5 mm, and 3 mm. In the test set, the recorded scores were 86.31, 87.69, and 90.84, respectively. The findings show that the system's performance is consistently maintained throughout various settings, demonstrating its reliability and efficacy.

Through a rigorous examination of the data, a comprehensive analysis has unveiled captivating patterns that warrant further investigation. The 'S' point, scientifically referred to as Sella, the 'Po' point referring to the Pogonion, 'Gn' point and othres, has garnered significant attention due to the outstanding performance the model in prediction them in the threshold of 2mm.For other points, such as the lower lip, the precision of the model experienced a remarkable increase between the different thresholds, reaching a flawless 100.00 Successful Detection Rate within the specified thresholds of 2.5 mm and 3 mm, respectively. The Porion, A-point, and

Gonion point on the hand, have presented significant obstacles, according to the findings. The detection of these landmarks in lateral cephalometric radiographs has proven to be remarkably elusive, highlighting the complicated process of their identification (see Fig. 4 to 8).

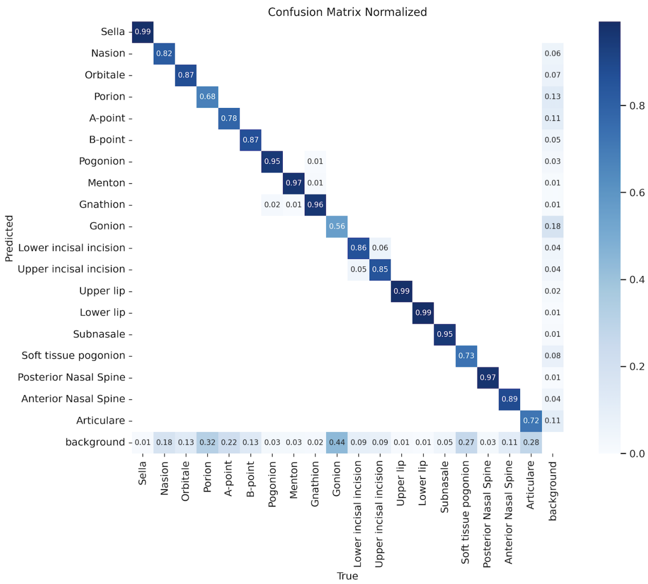

Fig. 4. Normalized confusion matrix for 2mm SDR.

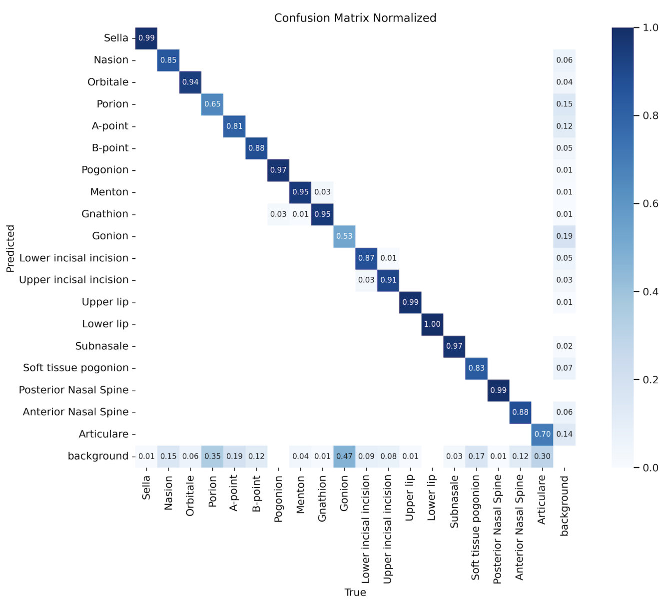

Fig. 5. Normalized confusion matrix for 2.5mm SDR.

To gain a comprehensive grasp of our findings and enhance our expertise, we did a meticulous study by meticulously comparing our results with those of previous studies that used the same dataset and followed the same data splitting techniques. The inclusion of a comparative analysis in our research has proven to be highly informative, as it has provided us with valuable context that allows us to assess the significance of our findings in relation to the existing body of knowledge (Table II, and Fig. 9).

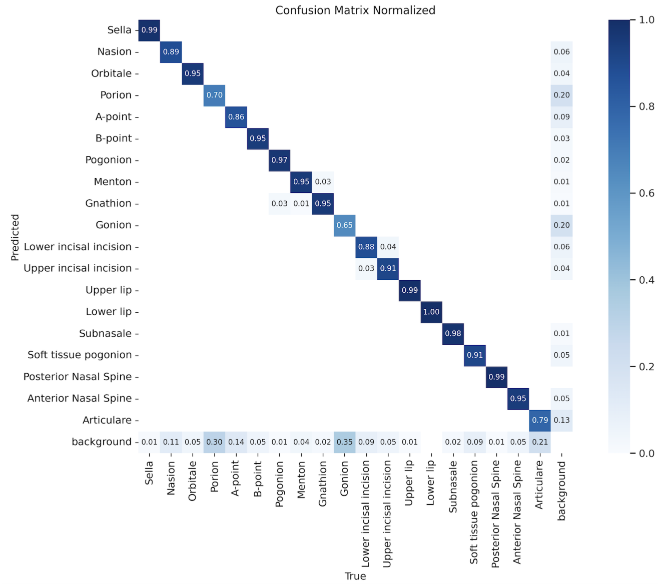

Fig. 6. Normalized confusion matrix for 3mm SDR.

TABLE II. Comparison of Results

	2mm	2.5mm	3mm
Ibragimov et al.[25]	62,74	70,47	76,53
Lindner et al. [26]	66,11	72,00	77,63
Arik et al.[19]	67,68	74,16	79,11
Qian et al. [27]	72,40	76,15	79,65
Oh et al. [28]	75,90	83,40	89,30
CephaX [29]	74,58	83,40	89,30
Our Model	86,31	87,69	90,84

After a thorough presentation of our findings and a meticulous examination of existing literature, our attention now turns towards a deeper exploration and analysis. In the forthcoming section, we embark on a more profound exploration of the complexities inherent in our discoveries. The primary objective of this discussion chapter is to provide a comprehensive context for our findings within the wider realm of established knowledge, elucidating the intricacies and ramifications of our research outcomes.

C. Discussion

The progressive assimilation of deep learning-driven artificial intelligence (AI) algorithms in the realm of medical image analysis has established a paradigm-shifting domain, particularly within the discipline of orthodontics. At the core of this metamorphosis lie cephalometric images, which serve as a crucial diagnostic instrument in assessing the complex interconnections among the mandible, maxilla, dentoalveolar structures, and in identifying dental and skeletal irregularities. The indisputable significance of cephalometric analysis in the field of orthodontics cannot be overstated. Nevertheless, it is imperative to acknowledge that this particular procedure is a complex and laborious undertaking, as its results are prone to fluctuations due to inherent variances in individual anatomical

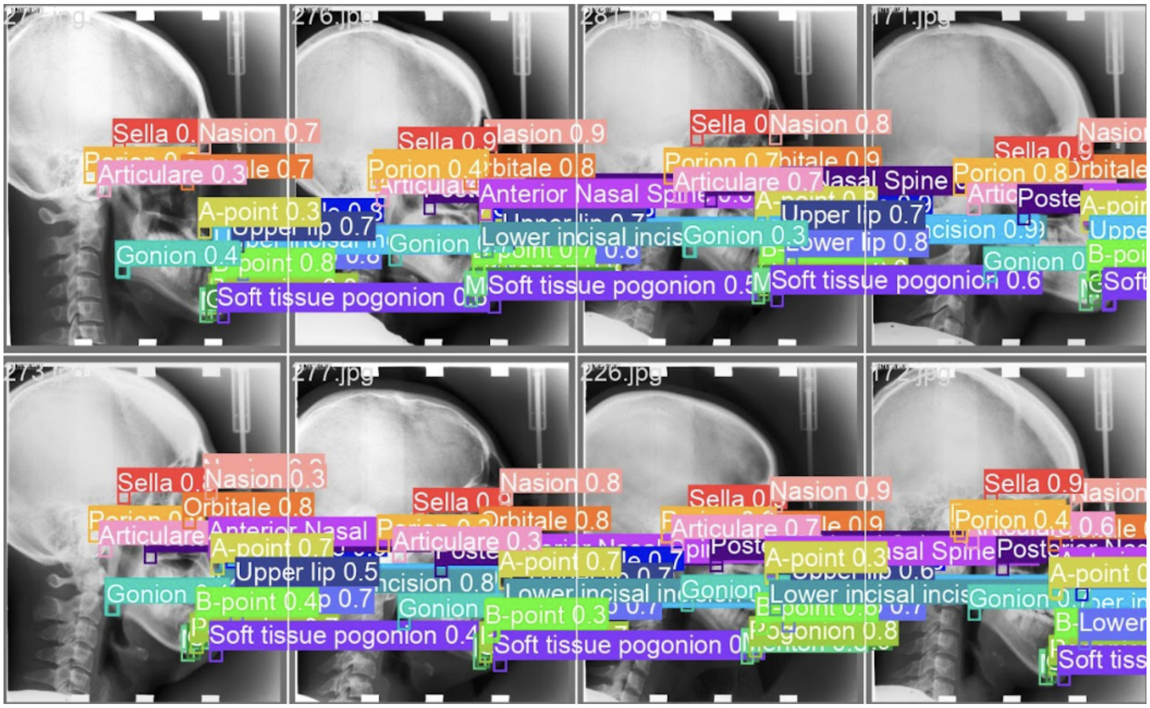

Fig. 7. Predicted batch sample.

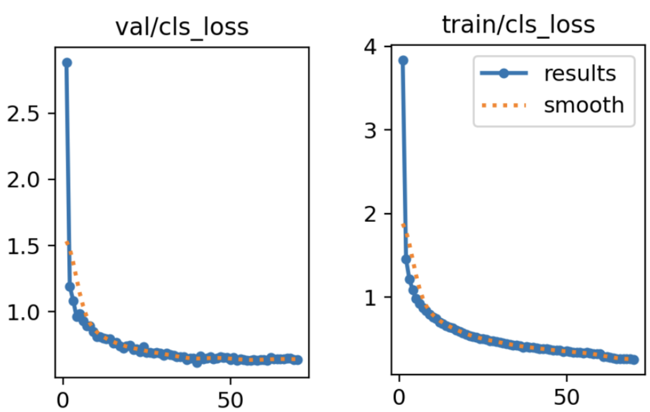

Fig. 8. Loss function.

structures.

The potential impact of AI algorithms in this particular context should not be underestimated. Numerous studies present in current scholarly literature provide substantial evidence supporting the effectiveness of diverse artificial intelligence (AI) techniques in facilitating and optimizing cephalometric analysis. The present study serves to augment the expanding reservoir of knowledge in this field. The YOLOv8 model exhibited varying levels of performance within the specified range of dimensions, specifically 2 mm, 2.5 mm, and 3 mm. Notably, the average SDR scores achieved were 86.31, 87.69, and 90.84, respectively. These results underscore the model's promising capabilities and aptitude in the given context.

Yet, a dive into literature and our observations indicate challenges in the automated detection of certain cephalometric landmarks. Specifically, points like A point, Go, Pg', and Or have traditionally posed challenges, often recording higher error rates or comparatively lower SDR scores. The Soft tissue pogonion point , as alluded to in [28], further exemplifies this challenge.

Notably, the A-point position being susceptible to variations based on head positioning, often complicates its precise tracing. This vulnerability of the A-point, corroborated by prior studies, underscores it as a landmark frequently marred by identification errors [5].

Notwithstanding the valuable insights provided by our research, it is imperative to duly recognize and address the limitations associated with our study. The images utilized in this study were obtained exclusively from a single source, ensuring consistency in terms of exposure parameters.

Furthermore, the labeling of these images was performed by an orthodontist, thereby ensuring accuracy and expertise in the categorization process. Moreover, it should be noted that the lack of external dataset validation and the restricted range of cephalometric landmarks examined could potentially impact the applicability of our results.

The story does not culminate at this juncture, as there exists an additional salient aspect necessitating a comprehensive examination. In the realm of orthodontics, the soft tissue paradigm [30] has ushered in a new era of comprehensive analysis, where the influence of facial soft tissue is taken into account in various jaw and tooth movements. Cephalometric studies encompass a range of soft tissue parameters, including but not limited to facial convexity, nasolabial angle, the positioning of the upper and lower lips, the mentolabial sulcus, as well as the positioning of the soft tissue chin and lower anterior face height [2].

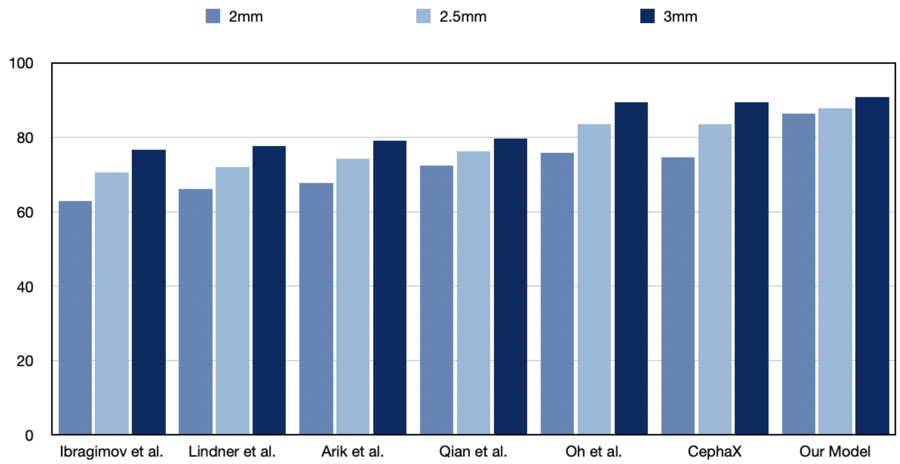

Fig. 9. Comparison of results.

The aforementioned characteristics are crucial in the field of orthodontics for making well-informed decisions on whether to pursue extraction or non-extraction treatment [31]. They play a crucial role in determining the degree of anterior teeth retraction, assessing growth changes, and evaluating surgical movements of the maxilla and mandible. In a recent investigation, it has come to light that the currently accessible datasets for soft tissue cephalometric analyses suffer from a notable limitation. These datasets only provide a small number of four soft tissue landmarks, making them insufficient for undertaking thorough analysis of soft tissue structures.

Soft tissue cephalometric analyses play a crucial role in various fields, including orthodontics, plastic surgery, and facial reconstruction. These analyses involve the examination and measurement of soft tissue landmarks to assess facial proportions, symmetry, and other relevant parameters. However, the limited number of soft tissue landmarks available in publicly accessible datasets severely hampers the accuracy and comprehensiveness of such analyses. The scarcity of soft tissue landmarks in these datasets poses a significant challenge for Furthermore, it is noteworthy that the existing datasets lack crucial occlusal landmarks, which play a pivotal role in the establishment of the occlusal plane. This plane holds significant implications in orthodontic diagnosis and treatment planning, as it has the potential to undergo alterations throughout the course of treatment. Consequently, a pressing demand arises for a novel dataset focused on cephalometric landmark detection. This dataset would efficiently address the current constraints and assist academics in developing sophisticated algorithms that might greatly improve cephalometric decision-making processes.

It is also worth noting that, Cephalometric analysis stands as a cornerstone in both orthodontics and orthopedics, relying on accurately identifying cephalometric landmarks. These landmarks are crucial reference points for a variety of measures and evaluations essential for diagnostic processes [32]. Steiner analysis and Tweed analysis are approaches that use several cephalometric landmarks to assess facial proportions, jaw connections, and dental inclinations. Dr. Tweed's research focused on the prognostic significance of landmark configurations [33]. Cephalometric study is useful not only for assessing dental parameters but also for evaluating upper airway dimensions and potential blockages. Precise recognition of landmarks is essential for assessing parameters like nasopharyngeal airway (NPA) depth and width, soft tissue thickness at various airway levels, and the relationship between the hyoid bone and the mandible, all of which impact airway openness and diagnostic precision.

VI. Conclusion

The integration of deep learning and artificial intelligence (AI) algorithms in medical image analysis, particularly in orthodontics, offers a promising approach to improve diagnostic precision and operational efficiency. The study examining the application of YOLOv8 for cephalometric landmark identification has strengthened the potential of these technical developments. The algorithm exhibited notable levels of accuracy within certain thresholds. However, it still faces persistent challenges, particularly in consistently detecting specific landmarks. Moreover, the variations in experimental procedures and inherent constraints of the study emphasize the need for wider and more varied testing environments.As the interaction between artificial intelligence (AI) and orthodontics becomes increasingly prominent, it is crucial for the technology and clinical sectors to work together to improve the effectiveness and usefulness of these technologies. Research such as ours is crucial in shaping the future of orthodontic diagnostics, which holds the potential for a harmonious integration of human expertise and technical proficiency.

References

[1] B. Trpkova, P. Major, N. Prasad, B. Nebbe *et al.*, "Cephalometric landmarks identification and reproducibility: a meta analysis," *American journal of orthodontics and dentofacial orthopedics*, vol. 112, no. 2, pp. 165–170, 1997.

[2] W. K. Darkwah, A. Kadri, B. B. Adormaa, and G. Aidoo, "Cephalometric study of the relationship between facial morphology and ethnicity," *Translational Research in Anatomy*, vol. 12, pp. 20–24, 2018.

[3] G. de Queiroz Tavares Borges Mesquita, W. A. Vieira, M. T. C. Vidigal, B. A. N. Travençolo, T. L. Beaini, R. Spin-Neto, L. R. Paranhos, and R. B. de Brito Júnior, "Artificial intelligence for detecting cephalometric landmarks: A systematic review and meta-analysis," *Journal of Digital Imaging*, vol. 36, no. 3, pp. 1158–1179, 2023.

[4] G. Wang, Y. Chen, P. An, H. Hong, J. Hu, and T. Huang, "Uav-yolov8: A small-object-detection model based on improved yolov8 for uav aerial photography scenarios," *Sensors*, vol. 23, no. 16, p. 7190, 2023.

[5] A. R. Durão, P. Pittayapat, M. I. B. Rockenbach, R. Olszewski, S. Ng, A. P. Ferreira, and R. Jacobs, "Validity of 2d lateral cephalometry in orthodontics: a systematic review," *Progress in orthodontics*, vol. 14, no. 1, pp. 1–11, 2013.

[6] S. Corbella, S. Srinivas, and F. Cabitza, "Applications of deep learning in dentistry," *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, vol. 132, no. 2, pp. 225–238, 2021.

[7] J. Yang, Y. Xie, L. Liu, B. Xia, Z. Cao, and C. Guo, "Automated dental image analysis by deep learning on small dataset," in *2018 IEEE 42nd annual computer software and applications conference (COMPSAC)*, vol. 1. IEEE, 2018, pp. 492–497.

[8] F.-E. Ben-Bouazza, Y. Bennani, G. Cabanes, and A. Touzani, "Unsupervised collaborative learning based on optimal transport theory," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 698–719, 2021.

[9] O. Manchadi, F.-e. Ben-Bouazza, and B. Jioudi, "Predictive maintenance in healthcare system: A survey," *IEEE Access*, 2023.

[10] S. Azeroual, F.-e. Ben-Bouazza, A. Naqi, and R. Sebihi, "Triple negative breast cancer and non-triple negative breast cancer recurrence prediction using boosting models," in *International Conference on Advanced Intelligent Systems for Sustainable Development*. Springer, 2022, pp. 440–450.

[11] I. Tafala, F. Bourzgui, M. B. Othmani, and M. Azmi, "Automatic classification of malocclusion," *Procedia Computer Science*, vol. 210, pp. 301–304, 2022.

[12] J. Londono, S. Ghasmi, A. H. Shah, A. Fahimipour, N. Ghadimi, S. Hashemi, Z. K. Sultan, and M. Dashti, "Accuracy of machine learning and convolutional neural network algorithms on detecting and prediction of anatomical landmarks on 2d lateral cephalometric images-a systematic review and meta-analysis," *The Saudi Dental Journal*, 2023.

[13] C. Wang, C. Huang, C. Li, and S. Chang, "A grand challenge for automated detection of critical landmarks for cephalometric x-ray image analysis," in *IEEE International Symposium on Biomedical Imaging*, 2014.

[14] C.-W. Wang, C.-T. Huang, M.-C. Hsieh, C.-H. Li, S.-W. Chang, W.-C. Li, R. Vandaele, R. Marée, S. Jodogne, P. Geurts *et al.*, "Evaluation and comparison of anatomical landmark detection methods for cephalometric x-ray images: a grand challenge," *IEEE transactions on medical imaging*, vol. 34, no. 9, pp. 1890–1900, 2015.

[15] M. A. Khalid, K. Zulfiqar, U. Bashir, A. Shaheen, R. Iqbal, Z. Rizwan, G. Rizwan, and M. M. Fraz, "Cepha29: Automatic cephalometric landmark detection challenge 2023," *arXiv preprint arXiv:2212.04808*, 2022.

[16] B. Ibragimov, B. Likar, F. Pernus, and T. Vrtovec, "Automatic cephalometric x-ray landmark detection by applying game theory and random forests," in *Proc. ISBI Int. Symp. on Biomedical Imaging*. © Springer-Verlag Berlin Heidelberg 2014, 2014, pp. 1–8.

[17] C. Chu, C. Chen, L. Nolte, and G. Zheng, "Fully automatic cephalometric x-ray landmark detection using random forest regression and sparse shape composition," *submitted to Automatic Cephalometric X-ray Landmark Detection Challenge*, 2014.

[18] H. Lee, M. Park, and J. Kim, "Cephalometric landmark detection in dental x-ray images using convolutional neural networks," in *Medical imaging 2017: Computer-aided diagnosis*, vol. 10134. SPIE, 2017, pp. 494–499.

[19] S. Ö. Arık, B. Ibragimov, and L. Xing, "Fully automated quantitative cephalometry using convolutional neural networks," *Journal of Medical Imaging*, vol. 4, no. 1, pp. 014 501–014 501, 2017.

[20] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.

[21] E. Goutham, S. Vasamsetti, P. Kishore, and H. K. Sardana, "Automatic localization of landmarks in cephalometric images via modified u-net," in *2019 10th international conference on computing, Communication and Networking Technologies (ICCCNT)*. IEEE, 2019, pp. 1–6.

[22] M. Uğurlu, "Performance of a convolutional neural network-based artificial intelligence algorithm for automatic cephalometric landmark detection," *Turkish Journal of Orthodontics*, vol. 35, no. 2, p. 94, 2022.

[23] K. Alshamrani, H. Alshamrani, F. Alqahtani, and A. H. Alshehri, "Automation of cephalometrics using machine learning methods," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[24] C.-W. Wang, C.-T. Huang, J.-H. Lee, C.-H. Li, S.-W. Chang, M.-J. Siao, T.-M. Lai, B. Ibragimov, T. Vrtovec, O. Ronneberger *et al.*, "A benchmark for comparison of dental radiography analysis algorithms," *Medical image analysis*, vol. 31, pp. 63–76, 2016.

[25] B. Ibragimov, B. Likar, F. Pernus, and T. Vrtovec, "Computerized cephalometry by game theory with shape-and appearance-based landmark refinement," in *Proceedings of International Symposium on Biomedical imaging (ISBI)*, 2015.

[26] C. Tim, F. Cootes *et al.*, "Fully automatic cephalometric evaluation using random forest regression-voting," in *proceedings of the IEEE international symposium on biomedical imaging (ISBI)*, 2015, pp. 16–19.

[27] J. Qian, M. Cheng, Y. Tao, J. Lin, and H. Lin, "Cephanet: An improved faster r-cnn for cephalometric landmark detection," in *2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019)*. IEEE, 2019, pp. 868–871.

[28] K. Oh, I.-S. Oh, D.-W. Lee *et al.*, "Deep anatomical context feature learning for cephalometric landmark detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 3, pp. 806–817, 2020.

[29] C.-H. King, Y.-L. Wang, W.-Y. Lin, and C.-L. Tsai, "Automatic cephalometric landmark detection on x-ray images using object detection," in *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2022, pp. 1–4.

[30] Y. Almansob, M. Jubari, A. Li, L. Jun, S. Tang, A. Ali *et al.*, "Patient's facial soft tissue changes following the orthodontics treatment," *IOSR J Dent Med Sci*, vol. 18, pp. 69–78, 2019.

[31] S. Moon, A. M. A. Mohamed, Y. He, W. Dong, C. Yaosen, and Y. Yang, "Extraction vs. nonextraction on soft-tissue profile change in patients with malocclusion: a systematic review and meta-analysis," *BioMed Research International*, vol. 2021, 2021.

[32] G. Brugnara, M. Baumgartner, E. D. Scholze, K. Deike-Hofmann, K. Kades, J. Scherer, S. Denner, H. Meredig, A. Rastogi, M. A. Mahmutoglu *et al.*, "Deep-learning based detection of vessel occlusions on ct-angiography in patients with suspected acute ischemic stroke," *Nature Communications*, vol. 14, no. 1, p. 4938, 2023.

[33] Y. A. Nesterenko, "Methods of cephalometric analysis according to burstone cj, tweed ch and kim yh: opportunities, prospects and problems of use in ukraine," *Reports of Vinnytsia National Medical University*, vol. 25, no. 2, pp. 336–339, 2021.