

A Novel Cephalometric Tool Enhanced by AI Assistance

Riccardo Zese^{1,*}, Luca Lombardo², Matteo De Maio², Michael Tamascelli³ and Francesca Cremonini²

¹DOCPAS, University of Ferrara, Via Luigi Borsari 46, I-44121, Ferrara, Italy

³DE, University of Ferrara, Via Saragat 1 I-44122, Ferrara, Italy

²DMTR, University of Ferrara, Via Luigi Borsari 46, I-44121, Ferrara, Italy

Abstract

Lateral radiography is one of the most important records for patients' evaluation in orthodontics and cephalometric analysis is fundamental to conduct correct diagnosis and treatment plan. This analysis includes both linear and angular measurements that quantitatively describe cranial and intermaxillary relationships. In order to obtain such measurements, anatomical landmarks are used. These reference points can be found on the soft tissue profile and on hard tissues such as teeth and skeletal contour. It is important to be extremely precise in the identification of these landmarks to compute correct measurements: even the slightest discrepancy could result in wrong values leading to different and possibly erroneous treatment plan. The automatic computerized identification of such anatomical landmarks on lateral cephalograms would greatly simplify this important step in the diagnostic process. Our aim is to apply artificial intelligence techniques for the automatic detection of these landmarks, with the final objective of developing a software, THERE (auTomatic HELpeR for cEphalometry), which exploits a predictive model that analyses teleradiographs, returns the coordinates of the anatomical landmarks, and automatically calculates the measurements necessary for diagnosis. This short paper describes the system interface and the first results obtained towards the training of the model(s) for landmarks prediction.

Keywords

Machine learning, Diagnostic methods and tools, Odontology, Artificial intelligence, Computer vision

1. Introduction


Cephalometric analysis is of primary importance for the clinical evaluation of the orthodontic patient: it is essential for a correct diagnosis and for choosing the right treatment plan. It is a descriptive quantitative analysis, performed on a specific radiograph - called lateral teleradiograph of the head - for all the patients that need to undergo an orthodontic treatment. This patient's profile radiograph must be carried out in maximum intercuspation and will result in a 1:1 image of the patient's skull, analogically printed or digitally displayed. A set of anatomical landmarks, which are reference points located both on hard and soft tissues of the profile, will


HC@AIxIA 2023: 2nd AIxIA Workshop on Artificial Intelligence For Healthcare

*Corresponding author.

✉ riccardo.zese@unife.it (R. Zese); luca.lombardo@unife.it (L. Lombardo); matteo.demaio@edu.unife.it (M. De Maio); francesca.cremonini@unife.it (F. Cremonini)

ORCID 0000-0001-8352-6304 (R. Zese); 0000-0003-3221-8197 (L. Lombardo); 0000-0002-4641-2196 (F. Cremonini)

 © 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

be identified on this image. The accuracy in the identification of these landmarks will be the basis of the reliability of the measurements taken during the cephalometric analysis, i.e., the more accurate the identification of these landmarks, the more correspondence there will be with the real facial and dental pattern of the patient under analysis. The anatomical landmarks on the radiographic image allow the computation of planes, defined by angles and lines, used to describe and, if present, quantify dentoalveolar and skeletal anomalies.

There are numerous anatomical landmarks that can be identified in a cephalometry, thus, many different measures can be defined depending on which landmarks are considered. The choice of these measures depends on the type of investigation to be conducted with the cephalometry. For instance, the anatomical landmarks S (center of the sella turcica), N (most anterior point of the fronto-nasal suture), and A (most posterior point of the maxillary anterior concavity) define an angle that represents the sagittal position of the upper jaw in relation to the cranial base, describing its normal or excessively anterior/posterior positioning. Similarly, the angle defined by S, N and B (most posterior point of the mandibular anterior concavity), represents the sagittal position of the mandible with respect to the cranial base. The difference ANB between angles SNA and SNB represents the sagittal intermaxillary relationship. Normal values range in the interval 2 ± 2 degrees, whereas it represents anomalies outside this range, skeletal Class II if > 4 degrees or a skeletal Class III if < 0 degrees.

This shows how the precise, accurate and consistent identification of the initial anatomical landmarks, on which the analysis measurements are based, is of crucial importance to give the orthodontist a correct and undistorted view of the patient's cranial relationships. Historically, cephalometric analysis was performed in pencil on transparent sheets placed on the radiograph, process that requires a certain degree of effort on the part of the medics for calibrating the image and tracing the necessary anatomical landmarks and structures. In recent years this analysis is performed digitally with specific software that asks the physician only to position the anatomical landmarks on the radiograph removing the burden of calculating angles and lines. The remaining challenge in this area lies in the ability of cephalometric analysis software to automatically detect and position the anatomical landmarks, avoiding the prior need to manually set them, in order to make the work of the orthodontist more efficient.

Thus, cephalometry is suitable to be performed with the use of automatic systems since it is performed by analysing precise and well-located landmarks on the skull. The coordinates returned by the system can be drawn on the teleradiography, allowing the clinician to check the results and ensuring maximum transparency of the model results. This transparency ensures a high explainability (or interpretability) of the model. However, in the literature there is a lack of effective procedures for self-detection of cephalometric anatomical landmarks.

Bulatova et al. in 2021 [1] studied the accuracy and reliability of the ability of cephalometric landmarks detection performed by a well-known commercial system from DDH Inc. This study considered 16 landmarks taken from 110 teleradiographs of the skull. No significant differences were found for 12 out of 16 landmarks between the position automatically identified by the system and those identified by trained human operators, empirically demonstrating that AI is a promising tool to facilitate the performance of cephalometric analysis in routine clinical practice and can speed up the analysis of large databases for research purposes. Similarly, Kim et al. [2] analysed a database of 2075 images to automatically identify the position of important landmarks in cephalometric analysis, obtaining an anatomical diagnosis of subjects based on

the correct landmark placement in 88.43% of cases.

In this short paper we present preliminary work to develop a system able to return the coordinates of the landmarks and to calculate all the needed measurements for the correct classification of the three classes of malocclusion: protrusion, retrusion or absence of malocclusion. This allows the medics to have a powerful diagnostic tool on his side that speeds up the diagnosis. The main objective of the project will therefore be the training of a model capable of correctly identifying malocclusions and the development of a system, distributed as a Web application, using this model, called THERE. The final system will allow the orthodontist to upload the radiographs and obtain the cephalometry analysis, accompanied by the landmarks obtained, the classification and the measurements. This will allow even less experienced orthodontists to minimise errors, even in cases where landmarks identification appears more complex, e.g., due to poor image quality. THERE will also allow the user to check the landmarks obtained and, if necessary, correct them if wrongly placed. The implementation of THERE as a web application allows for easy distribution of the system, which can be used by the user in any situation, easy maintenance, as each update will be made directly available without any action on the part of the user, ensuring maximum anonymisation of patients, as it will not require any type of data other than the image of the cephalometry, and will allow expert users to correct wrong predictions on the one hand, and the application to continuously collect new data in order to constantly improve its accuracy on the other hand.

2. THERE Web Application

The application has been developed using the Flask Framework, written in Python. It is designed to allow to focus on application-level business logic, without unnecessary ties to specific deployment environments. The application is developed as a package and exploit Flask Blueprints to be easily extensible in future. They allows the encapsulation of functionality, such as views and templates, helping the adoption of a Model-View-Controller pattern.

The workflow of the application is linear and has been kept simple with the aim of developing an application which is straightforward to use for a user, who could be not proficient in computer use. Moreover, since the application has to work with teleradiographs from patients, the application must ensure the highest privacy. For this reason, no sensible data is collected by the application. However, if the user decides to significantly change the coordinates of the anatomical landmarks found, the application will save internally the image, which does not contain sensible data of the patient, and the corrected coordinates in order to improve the performance of the underlying neural network.

Basically, the Web application is composed of three main views: **Home**, which is the landing page of the URL of the application, from which the user must start using the application; **Calibration**, the page where the user has to set the application to work on the uploaded image; **Dashboard**, the operating page, where the uploaded image is analysed, the anatomical landmarks are shown and the user can both modify the coordinates of the landmarks and read the measurements necessary for the diagnosis.

The complete workflow is depicted in Figure 1. The user visits the **Home** page and uploads an image. The user will be directed to the **Calibrate** page, where the uploaded image is shown,

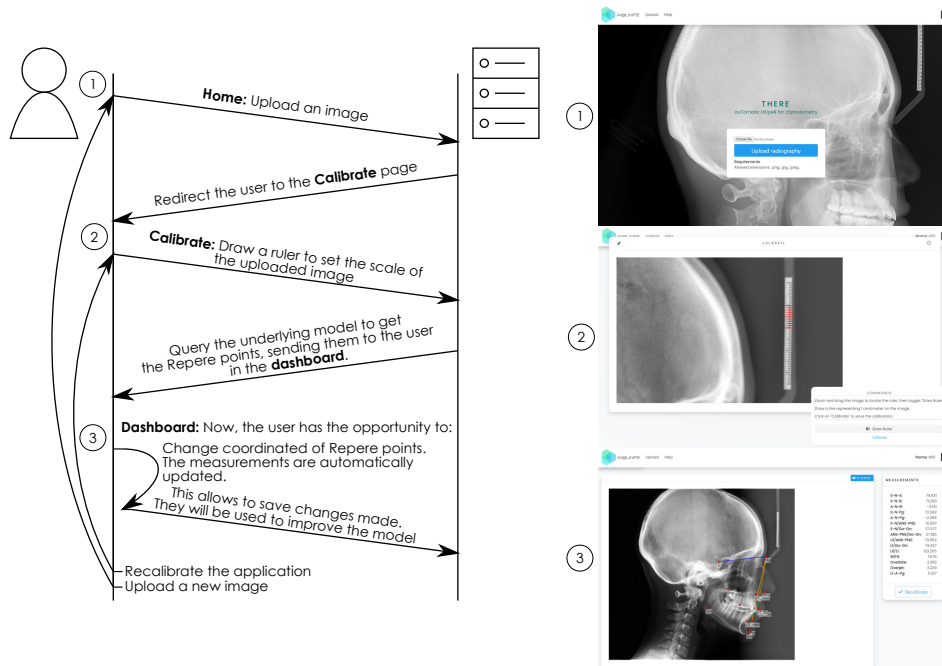


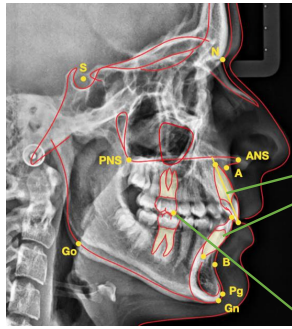
Figure 1: Workflow of the application and screenshots for the three pages.

and the user is called to indicate the scale of the image. Cephalometries include a ruler, used to calculate the correct scale of the image. So, in the **Calibrate** page, the user has to draw a ruler, as shown by ② in Figure 1, to calibrate the application in the same way they should do with standard cephalometries. Finally, the user will enter the main **Dashboard**, where they can modify the anatomical landmarks and read all the important measurements they need instantaneously.

3. Anatomical landmarks detection

The current dataset is composed of 1732 images (PNG or JPG) of anonymized lateral teleradiographs of the head of different patients. Images are taken using different devices and present different levels of saturation, colour and brightness. Each image is labelled with the coordinates in pixels of the 14 anatomical landmarks considered in this project, shown in Figure 2. Images size spans between 2685×2232 and 316×224 pixels, most of them in RGB format. All teleradiographs are taken with the patient looking to the right side of the picture, as shown in Figure 2.

One of the biggest problems to be solved in analysing images is their dimensionality. The needs of high accuracy in the coordinates detection of the anatomical landmarks is the main objective we must pursue when modelling and training the underlying model of the application. Moreover, the number of landmarks to be found and their proximity further increase the complexity of the problem. Preliminary experiments have shown how the complexity of teleradiography and



- S:** Central point of the Sella Turcica.
- N:** Deepest antero-posterior point of the nose-frontal suture.
- ANS:** Most anterior bony point of the anterior nasal spine.
- PNS:** Radiological point determined by the perpendicular drawn from the apex of the pterygopalatine fossa to the bispal plane.
- A:** Most posterior point of the anterior concavity of the maxillary alveolar process.
- B:** Most recessed point of the anterior concavity of the mandibular alveolar process.
- U1 root/U1 tip:** Root apex/incisal edge of the upper central incisor.
- L1 root/L1 tip:** Root apex/incisal edge of the lower central incisor.
- Pg:** Most anterior point of the mental symphysis contour.
- GN:** Lowest point of the mental symphysis contour.
- Go:** Geometric point constructed at the point where the tangent to the ascending branch of the mandible meets the plane of the mandible.
- Mesial:** apex of the mesiovestibular cusp of the upper first molar.

Figure 2: Anatomical landmarks considered.

the low variance of the coordinates values among the different examples tends to force the model to compute an average position for the different landmarks, achieving low loss values but always returning similar coordinates. In this regard, reducing the size of images too much could be a tricky path to follow, because it tends to facilitate the occurrence of this problem. In our first tests we followed two different approaches: (1) we divided the anatomical landmarks in four subsets, containing landmarks used together when computing the skull measurements, and trained four models specialized on a single subset, the final results will be obtained by the ensemble of the four models; and (2) we applied bigger networks to the whole set of landmarks.

Setting 1. In this setting all the images were resized to 224x224 pixels, the size of the smallest image, and transformed in RGB format. Data was augmented by flipping images and applying rotation (± 10) and scale (50%-70%) both with a probability of 0.3. The pixels values were scaled in the range $[-1,1]$. We built four networks, one for each subset of landmarks, namely (S,N,A,B), (ANS, PNS, GN, Go), (U1 root, U1 tip, L1 root, L1 tip), and (Pg, Mesial). The backbone of the network is a EfficientNetB7 pretrained on Imagenet, which is followed by two Separable Convolution layers with kernel size of 5x5 the first and 3x3 the latter, stride 1, and 8 filters. We used Nadam as optimizer with a learning rate of 0.001, which is reduced by a factor 0.2 when a plateau is reached, and early stopping. As loss we considered the mean squared error. The best model for each group achieved a test loss values between 0.00045 and 0.00068. All the models need 9MB to be saved in memory. Figure 3 shows some examples of the results obtained by the ensemble of the four networks. On average, the distance between the actual and predicted points is 6.3 ± 4.6 mm.

Setting 2. Before training the model, data in the training set were pre-processed by converting all the images to greyscale, normalizing the pixels values, and resizing them into 1000x1000 pixels. Then, data augmentation was performed by adding to the training set three new images for each original image created by: flipping the original image, computing linear contrast and adding gaussian blur with a probability of 80%, and rotating the original and the flipped images of ± 15 degrees. To try to facilitate training, the coordinates have been standardised in the range $[-1,1]$. We tried two different types of networks: (N1) a classical CNN composed of five blocks applying convolution (with 32, 64, 128, 256, and 512 filters with kernel sizes

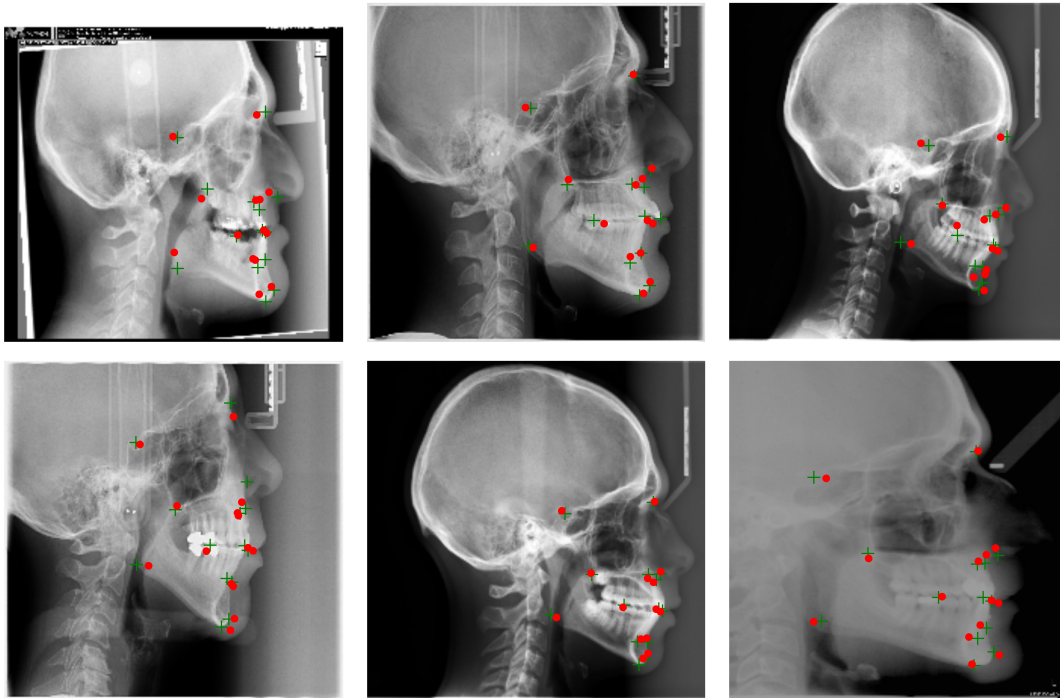


Figure 3: Examples of results for the ensemble model, label landmarks are depicted with a green cross, while predicted landmarks with a red dot.

of 3x3 except for the first block with a kernel size of 5x5) and max pooling, followed by a global average pooling and a dense layer of 256 neurons with ReLU activation function and the output layer returning the coordinates of the 14 landmarks; (N2) a network composed of 5 inception modules, shown in Figure 4, with 64, 64, 96, 96, 128 filters for the 1x1 convolution, 64, 64, 128, 256 filters for the 3x3 convolution and 32, 32, 64, 64, 128 filters for the 5x5 convolution. The network ends with a dense layer of 1024 neurons with ReLU activation function and the output layer returning the coordinates of the 14 landmarks. We used Adam as optimizer with a learning rate of 0.001 and early stopping. Moreover, the learning rate is reduced in case of plateau by a factor 0.2. As loss we considered the mean squared error. Both models achieved a mean squared error computed on the test set near 0.03. On average, the distance between the actual and predicted points computed by N1 network is 24.8 ± 7.9 mm, while that of N2 network is 11.5 ± 5.4 mm. We observed that, in general, the model based on inception (N2) was less prone to return the same coordinates for each image, as can be seen in Figures 5 (green cross for label landmarks, red dots for predicted), but it is significantly bigger than the model without inception blocks (more than $9 \cdot 10^6$ parameters - 103MB vs $5.6 \cdot 10^6$ parameters - 20MB).

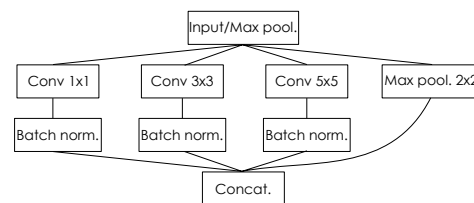


Figure 4: Inception block

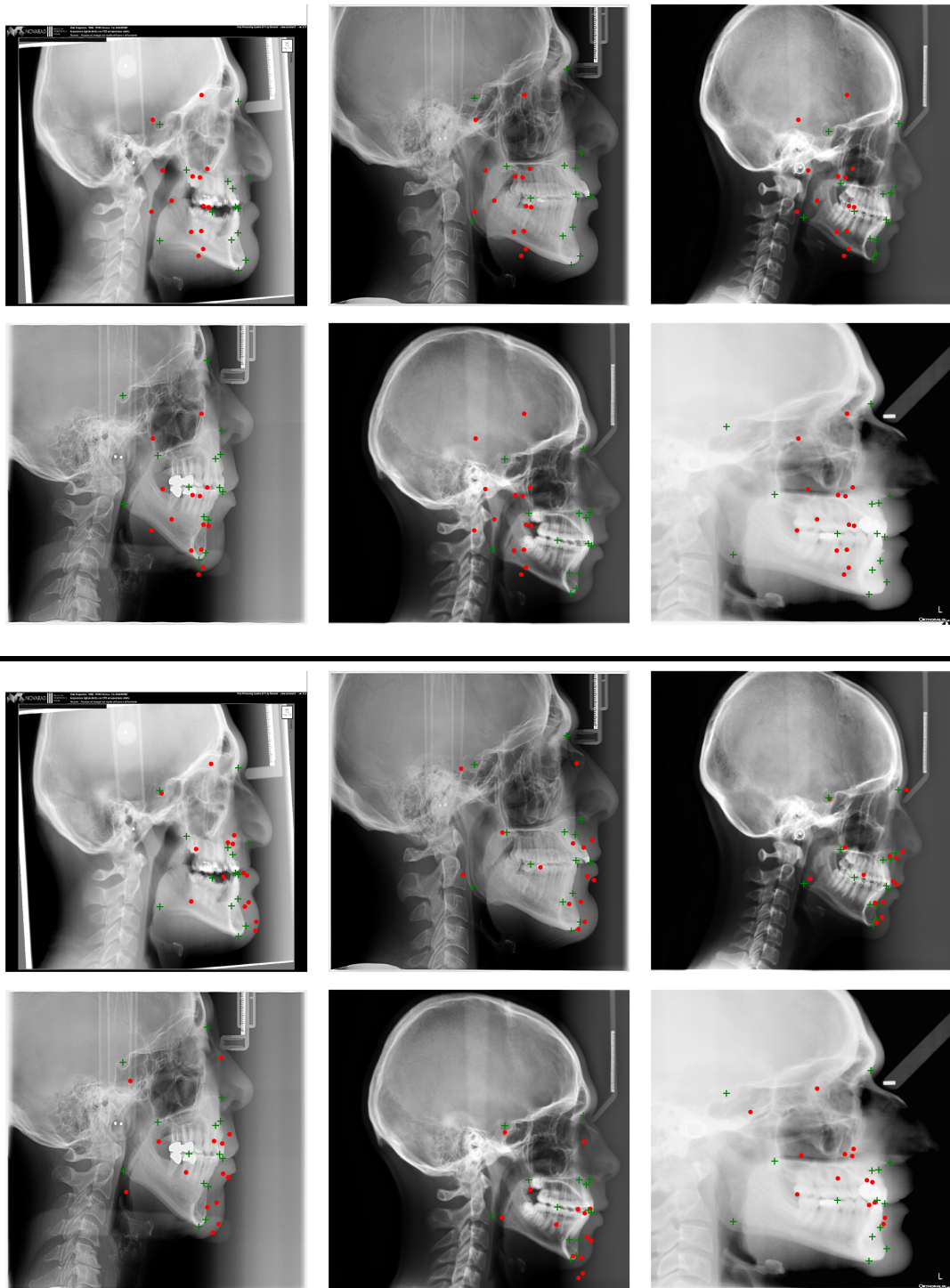


Figure 5: Examples of results for CNN model (first six) and Inception model (last six), label landmarks are depicted with a green cross, while predicted landmarks with a red dot.

4. Conclusions and Future Work

In this paper we presented a Web application called THERE, that allows users to upload a teleradiography which is analysed by an underlying neural network to locate the coordinates of the 14 anatomical landmarks necessary to perform a cephalometry. Preliminary work on the neural network shows that the localization of these landmarks is not trivial to perform automatically due to the type of images used. The considered models tend to learn the average position of each landmark instead of concentrating on the image itself, leading to the need of an accurate investigation of how the images should be fed to the model and how the model should be designed. Better results have been achieved by considering subsets of landmarks separately. We are currently working on different models, from sequential CNNs to the more complex transformer [3] and YOLO-Pose [4] architectures. We also plan to study possible masks to apply to the image or the possibility to apply patching, i.e., create sub figures from the whole teleradiography and train the network on the sub figures instead of the entire image. We also plan to perform a validation of the effectiveness and usability of the system by administering a questionnaire to the users of the application, based on Post-Study System Usability Questionnaire (PSSUQ) Version 3 template [5, 6]. Finally, we plan to add functionalities to the system, such as performing classification about other pathologies using intra-oral pictures of patients. **Acknowledgments** This work is financed by "Bando Giovani anno 2022 per progetti di ricerca finanziati con il contributo 5x1000 anno 2020".

References

- [1] G. Bulatova, B. Kusnoto, V. Grace, T. P. Tsay, D. M. Avenetti, F. J. C. Sanchez, Assessment of automatic cephalometric landmark identification using artificial intelligence, *Orthodontics & Craniofacial Research* 24 (2021) 37–42. doi:10.1111/ocr.12542.
- [2] H. Kim, E. Shim, J. Park, Y.-J. Kim, U. Lee, Y. Kim, Web-based fully automated cephalometric analysis by deep learning, *Computer Methods and Programs in Biomedicine* 194 (2020) 105513. doi:10.1016/j.cmpb.2020.105513.
- [3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, in: I. Guyon, U. von Luxburg, S. Bengio, H. M. Wallach, R. Fergus, S. V. N. Vishwanathan, R. Garnett (Eds.), *Advances in Neural Information Processing Systems* 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA, 2017, pp. 5998–6008.
- [4] D. Maji, S. Nagori, M. Mathew, D. Poddar, Yolo-pose: Enhancing YOLO for multi person pose estimation using object keypoint similarity loss, in: *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2022, New Orleans, LA, USA, June 19-20, 2022, IEEE, 2022*, pp. 2636–2645. doi:10.1109/CVPRW56347.2022.00297.
- [5] J. R. Lewis, IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use, *International Journal of Human-Computer Interaction* 7 (1995) 57–78. doi:10.1080/10447319509526110.
- [6] J. R. Lewis, J. Sauro, Revisiting the factor structure of the system usability scale, *Journal of Usability Studies* 12 (2017) 183–192.