# **Cephalogram Landmark Detection: A Survey Paper**

Rakesh Kumar, Jayasurya, and Ayyappan

Vellore Institute of Technology, Chennai, India

{rakeshkumar.ks, jayasurya.s, ayyappan.t}2020@vitstudent.ac.in

**Abstract**

Cephalometric analysis provides important information about craniofacial structures, which is useful in the fields of orthodontics, maxillofacial surgery, facial reconstruction and to diagnose abnormal growth in the face. An essential part of this analysis is cephalometric landmark detection, which is the process of locating and identifying particular anatomical spots on the radiographs. The merging of computer vision, machine learning, and image processing techniques has led to notable breakthroughs in this process in recent years. Techniques involving template matching, feature extraction machine learning and deep learning are being used for precise landmark detection. This paper offers an overview of the rapidly developing progress on cephalometric landmark detection, report the research gaps and the challenges present

**Keywords:** Landmark Detection, Computer Vision, Machine Learning

1. **Introduction**

In the domains of craniofacial analysis and orthodontics, accurate identification of cephalometric landmarks is a prerequisite for diagnosis, treatment planning, and result evaluation. A fundamental component of orthodontics, cephalometric analysis is measuring and interpreting different skeletal and dental connections using certain anatomical landmarks seen on radiographic images. Digital Image Processing is revolutionizing multiple domains in the medical field, of which cephalometric analysis is one. As a result, automated approaches for landmark recognition have emerged that are faster, more reliable, and less prone to human interpretation errors.

Cephalometric landmark detection, a subdomain within medical image analysis, focuses on developing algorithms and techniques to automatically locate key anatomical points in cephalometric radiographs. Cephalometric Analysis is the process of analyzing cephalogram radiographs to identify the skeletal structure and relative positions of facial structures of the human skull, especially the maxilla, mandible regions of the skull. This analysis is often used as a step for planning maxillofacial surgery, orthodontics, craniofacial reconstructive surgery. Cephalograms are analyzed by medical professionals for identification of landmarks. Landmarks in a cephalogram are points of reference that are used for measurements and further analysis. The landmarks can also be joined together to form a line. These can be seen as features which are used for distance, angle, coordinate analysis and in some cases arcial analysis. All of these measurements are used to establish the relationship in skeletal structure of the human skull.

These landmarks assist doctors examine craniofacial anatomy to develop accurate treatment plans by acting as reference points for quantitative assessments. The area has seen a remarkable development in recent times, fueled by advances in better computing hardware, image processing, machine learning, and computer vision.

Cephalometric landmarks were traditionally identified by hand, which required exacting attention to detail and advanced anatomical understanding by an expert. Nevertheless, this approach required a lot of work, took a long time, and was subject to observer variability. As the digital era emerged, researchers tried to use technology to improve landmark recognition's effectiveness and precision.

A paradigm change in orthodontics was brought about by the switch from manual to computerized cephalometric analysis. The widespread use of digital cephalometric radiographs made patient data exchange, retrieval, and storage simpler. In parallel, researchers investigated computer-aided techniques for landmark recognition in an effort to mechanize a procedure that has historically depended on human knowledge. Early efforts used simple image processing methods and rule-based systems. Eventually, the use of machine learning signaled a revolutionary shift in the recognition of cephalometric landmarks. Neural networks and other supervised learning algorithms have shown promising findings in automatically learning the intricate patterns connected to cephalometric landmarks. Because these algorithms were trained on annotated datasets, they can correctly recognize landmarks on fresh radiographs and generalize their expertise.

The development of convolutional neural networks (CNNs), in particular and deep learning greatly advanced the area. CNNs performed well when it came to recognizing hierarchical characteristics in pictures, which made them a good fit for the complex cephalometric radiographs. With minimal labeled cephalometric data, transfer learning approaches enabled the creation of strong landmark identification models by utilizing pre-trained models on large datasets. Recently, attention based models utilizing Squeeze-and-Excitation blocks, Vision Transformers, and so on have been gaining popularity as well.

Cephalometric landmark detection is still confronted with a number of difficulties despite significant advancements. Algorithmic accuracy is hindered by variations in patient anatomy, picture quality, and the existence of anatomical abnormalities. Robustness to such fluctuations continues to be an important area of study. There are still issues, such as the requirement for sizable annotated datasets for deep learning model training and the moral dilemmas associated with patient data privacy.

These difficulties, nevertheless, also present opportunities for creativity. Scholars are investigating innovative methods such as combining 3D imaging, multi-modal data fusion, and attention processes to improve the precision of landmark identification and overcome the drawbacks of conventional 2D cephalometry. Interprofessional collaborations among radiologists, computer scientists, and orthodontist specialists are promoting multidisciplinary methods to address these issues.

This comprehensive survey aims to present an overview of the most recent developments in cephalometric landmark detection. It investigates the development of methods, the influence of machine learning, and the current trends reshaping the subject by a thorough examination of the literature. The survey will also draw attention to the ongoing difficulties and new directions in research. This survey attempts to be a useful tool for researchers, clinicians, and technologists working at the interface of orthodontics and medical image processing by consolidating current information.

1. **Datasets Used**

The most common dataset used by researchers in this topic is the ISBI 2015 Grand Challenge Dataset. Many papers also use private datasets that they obtained from a medical institution and hence those are not publicly accessible. These datasets obtained from medical institution by the researchers were usually either X-ray or CT images.

*2.1 ISBI 2015 Grand Challenge Dataset*

The IEEE International Symposium on Biomedical Imaging (ISBI) held on 2015 released a dataset containing 400 cephalograms obtained from a Hospital in Taipei, Taiwan. These cephalograms were captured by Soredex CRANEX® Excel Ceph machine and processed by Soredex SorCom software (3.1.5 version 2.0) [1]. The resolution of these images is 1935 x 2400 pixels. The cephalograms are obtained from a wide range of age group ranging from 6 to 60. These cephalogram images are saved as a TIFF format image that stores raster images. This is the most widely used dataset among researchers in automated cephalometric analysis. The dataset has 19 landmarks marked as (x, y) coordinates of pixels by two experts with an experience of 15 and 6 years respectively. Each expert marked the landmarks twice and the average of these 4 markings were taken as the ground truth to minimize inter- and intra-observer variation. It has a pixel sizing of 0.1 mm x 0.1 mm

The landmarks marked in this dataset are: the sella, the nasion, the orbitale, the porion, the subspinale, the supramentale, the pogonion, the menton, the gnathion, the gonion, the lower incisal incision, the upper incisal incision, the upper lip, the lower lip, the subnasal, the soft tissue pogonion, the posterior nasal spine, the anterior nasal spine, the anterior nasal spine and the articulate [1]

*2.2 PDDCA*

Public Domain Database for Computational Anatomy (PDDCA) dataset contains CT scans of 48 patients and has manual organ segmentation as well as five bony landmarks. Since, it is a CT scan, the data of the skull is used for 3D automated cephalometric analysis and have been used by some researchers. The CT scans are available for download at The Cancer Imaging Archive (TCIA). It is originally from the Radiation Therapy Oncology Group (RTOG) 0522 study.

**Table 1** - Overview of the datasets used

|  |  |  |  |
| --- | --- | --- | --- |
| **Dataset** | **Modality** | **Dimension** | **# of images** |
| ISBI 2015 Grand Challenge | X-ray | 2D | 400 |
| PDDCA | CT | 3D | 48 |

1. **Imaging Modalities**

X-rays, CBCT and CT imaging is commonly used for cephalometric analysis. The most common open-source dataset for cephalometric analysis uses X-rays. CBCT and CT are the modern imaging modality for performing cephalometric analysis as their measurements tend to be more precise than X-rays. CBCT is preferred over CT due to its lesser degree of radiation exposure. MRI has no radiation exposure but it is expensive compared to other modalities and consequently used less.

*3.1 X-rays*

X-ray is the most widely utilized imaging modality in cephalometric analysis because of its affordability and simplicity. This also applies to recent researches done in this area because the most often utilized dataset is the well-known ISBI 2015 Grand Challenge dataset. Usually, X-ray imaging modality is limited to 2D images. 3D X-ray can be taken as well by rotating an array of X-ray scanners, but the prominent one is 2D. X-ray is obtained by exposing the patient to radiation, similar to other modalities mentioned. Cephalometric analysis is used to study the relationship between bony and soft tissue landmarks and X-ray is good at imaging soft tissues. The ISBI Grand Challenge contains 400 X-ray images that were then marked and reviewed by medical professionals. Some studies reviewed in this survey also obtain X-ray images from hospital(s).

*3.2 CT*

Computed Tomograph (CT) scans are sets of cross-sectional X-ray scans. An X-ray source beam and detector is placed facing each other in a rotating setup which passes through the subject and creates multiple slices of cross-sectional scans. It can be visualized as a triangular beam passing through the body from the source to the detector. The subject is moved while the scanner rotates to capture these slices of scans. These scans can then be viewed in 2D or be stacked together and processed by a software to create a 3D representation of it. This is its main advantage over X-ray as X-ray can only produce 2D output. CT images are more detailed than conventional X-ray images. It reveals bones and soft tissues. Exposure of radiation higher than other modalities mentioned here is a big disadvantage for CT as long or higher dose of exposure has a small chance of causing cancer.

*3.3 CBCT*

Cone Beam Computed Tomography (CBCT) is an extensively used modality for cranial and skull imaging. It is a variant of CT and is preferred due to its less radiating nature. CBCT achieves this by passing a cone shaped beam instead of a triangular beam, capturing more volumetric data in a single pass. CBCT can produce 2D as well as 3D cephalograms that can then be used for analysis. 3D cephalograms are recommended for maxillofacial surgeries as they can provide more information and hence enable surgeons to precisely decide on the medical procedure. It’s ability to model cranial and dental regions in 3D is one of its main benefit. Due to its radiating nature, it is not to be used frequently on the same person, especially on children [2].

*3.4 MRI*

Magnetic Resonance Imaging (MRI) is a technique that exploits magnetic properties to produce a detailed image of the body. It does not produce any ionizing radiation and hence does not pose any threat of cancer to its subject or operator. This is its biggest advantage. However, it does cost more to take an MRI and hence its use and well as MRI data for cephalometry is quite limited. MRI captures soft tissues really well and it is frequently used to diagnose and plan the treatment of implants, jaw lesions, etc., to achieve better prognosis [3]. MRI is usually viewed as 2D output slices but it can be processed and converted to 3D using specialized software, similar to CT.

**Table 2** - Comparison of Medical Imaging Modalities

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Imaging Modality** | **Output Dimension** | **Radiation Exposure** | **Cost** | **Time Taken** |
| X-ray | 2D | Low | $ | \*\* |
| CT | 2D or 3D | High | $$ | \*\*\* |
| CBCT | 2D or 3D | Medium | $$ | \* |
| MRI | 2D or 3D | None | $$$ | \*\*\*\*\*\* |

Here, $ and \* are used to show the relative cost and time taken

1. **Metrics Used**

Different techniques and architectures use different training metrics but most used the same evaluation metrics that we discuss below. These were the most common ones used by researchers

* 1. *Training Metrics*

*4.1.1 Mean Squared Error Loss*

In case of training a regression model for the coordinates, the mean of or commonly known as MSE is usually used as the loss function

*4.1.2 Cross Entropy Loss*

Some methods predict a landmark heatmap instead of predicting a coordinate using regression. They do this by convert the ground truth coordinate into a probability distribution heatmap and making the model predict the same. For such probability distributions, Cross Entropy (CE) loss is used.

Where, is the ground truth probability and is the predicted probability

*4.2 Evaluation Metrics*

*4.2.1 Mean Radial Error (MRE)*

Mean Radial Error is the mean of squared error between predicted and ground truth coordinates. While evaluating, often MRE is used after converting the pixel values to millimeters using a conversion factor *S* which depends on the scanner used to take the radiogram.

Where,

The landmarks predicted by automated systems should ideally have an error of <= 2 mm relative to the ones marked by medical professionals [1]

*4.2.2 Mean Absolute Error (MAE)*

It is derived by taking the mean of the absolute error between the predicted and the ground truth values

*4.2.3 Standard Deviation (SD)*

Standard Deviation is a statistical measure of how much a data varies around its mean. This is usually combined with the MRE to give an numerical representation of how well the landmark detection system works.

It is written with distance as follows:

*4.2.3 Successful Detection Rate (SDR)*

In landmark detection tasks, a threshold is used to measure the performance of the detection system alongside a distance metric. SDR is used to give an idea about how many landmarks were identified within a maximum range of some threshold *T*. Mathematically, it can be written as follows

Where,

Common thresholds that are used by most papers in this survey are 2.0 mm, 2.5 mm, 3.0 mm, 4.0 mm

*4.2.4 Successful Classification Rate (SCR)*

Some methods predict each landmark as a class and so, SCR is a metric that describes the accuracy of this classification step.

Where,

1. **Techniques Used**

*Template Matching Based*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Year and Author** | **Method Used** | **Number of Landmarks** | **Dataset Size** | **Imaging Modality** | **Results** |
| Kaur et al. 2015 [4] | Zernike Moments and Template Matching | 19 | 135 | X-ray | 89.5 % < 2 mm |
| Koga et al. 2023 [5] | Robotic Process Automation | 49 | 300 | CT | -- |

*Knowledge Based*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Year and Author** | **Method Used** | **Number of Landmarks** | **Dataset Size** | **Imaging Modality** | **Results** |
| Yang et al. 2001 [9] | Computerized Cephalometric analysis system based on digital image processing and computer graphics techniques. | 19 | 500 | X-ray | Mean error = 0.9 mm and 1.2 deg |
| Gupta et al. 2016 [8] | An Automatic knowledge-based landmark detection algorithm | 21 | 30 | CBCT | Mean error = 1.67 mm  S.D = 0.83 mm. |
| Vezzetti et al. 2018 [7] | Deterministic model with thresholding technique | 68 | 3362 | CT | Mean error = 4.75 mm |
| Alessandri et al. 2023 [6] | Artificial Intelligence-Assisted model | 25 | 13 | X-ray | Mean error = 0.807mm and 1.854mm for Posterior Facial Height and Facial Axis Angle respectively |

*Machine Learning Based*

Machine Learning is a set algorithms with a statistical approach of training and predicting data. Algorithms like Random Forest tend to perform really well when the features for it are selected right. So, features are extracted from these cephalograms and are used to identify the landmarks

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Year and Author** | **Method Used** | **Number of Landmarks** | **Dataset Size** | **Imaging Modality** | **Results** |
| Chakrabartty et al. 2003 [10] | Support Vector Machines | 19 | 130 | X-ray | Mean error = 1.84 |
| Vucinic et al. 2010 [11] | Combination of active appearance models (AAMs) and, statistical shape and appearance model. | 50 | 60 | X-ray | 61% < 2mm |
| Keustermans et al. 2010 [12] | Shape and local appearance models based predictor using statistical features. | 19 | 400 | CBCT | Mean error = 1.64 mm |
| Mirzaalian et al. 2014 [13] | Random Forest Classifier | 20 | 200 | X-ray | 87.68% < 4 mm |
| Chen et al. 2015 [14] | An active contour model for tooth segmentation | 80 | 1000 | CBCT | DSC = 0.92 |
| Wang et al. 2018 [15] | Multiresolution decision tree regression voting model | 19 | 465 | X-ray | 73.37% and 72.08% < 2mm for test1data1 and database2 |
| Montúfar et al. 2018 [16] | Active Shape  Models in Related Projections | 19 | 120 | CBCT | Mean error = 3.64 |
| Suhail et al. 2022 [17] | Ensemble of Regression Trees | 26 | 362 | X-ray | 58.58 % < 2 mm |

*Deep Learning Based*

Deep Learning is a subfield of Machine Learning that focuses on deep neural networks. Neural networks are a high level replica of biological neurons. A single unit of Artificial Neural Network (ANN) is called as a Perceptron. Perceptron is a simple unit that takes in a numerical value as input and performs a linear operation on it. A coefficient is multiplied to scale the input and a value is used to offset it. This coefficient and offset is known as the weight and the bias. More than one perceptron can be stacked together horizontally to form a layer of perceptrons. Such a layer can take in a vector of input wherein each input value can be mapped to one or more perceptron. If we stack more than one layer together, it is known as a Mutli-Layer Perceptron (MLP). These MLPs, when stacked for more layers (also called as the depth of the network) becomes deep and hence they are called as deep neural networks. The weights and biases of these networks can be trained to obtain the input-output mapping that we desire by using an optimization technique called backpropagation. Researchers have found numerous ways to improve these ANNs by using activation functions to introduce non-linearity, etc. Relating to image processing, Convolutional Neural Networks (CNNs) are the most predominant type of neural networks that perform convolution operation on images to extract the features needed for the task. We’ve curated, studied and listed some papers that perform automatic cephalogram analysis and have listed it

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Year and Author** | **Method Used** | **Number of Landmarks** | **Dataset Size** | **Imaging Modality** | **Results** |
| Montúfar et al. 2018 [18] | A hybrid approach combining Deep Learning and active shape models. | 19 | 1,000 | CBCT | Mean error = 2.51 |
| Dai et. al, 2019 [19] | CNN based Encoder-Decoder and Regression Voting | 19 | 400 | X-ray | SDR:  38% < 2.0 mm  80% < 2.0 mm |
| Kim. et. al, 2020 [20] | CNN based, Stacked Hourglass | 23 | 2075 | X-ray | SDR:  84.53% < 2.0 mm  96.79% < 4.0 mm  Overall Error: 1.37 ± 1.79 mm  SCR: 88.43% |
| Lee et. al, 2020 [21] | Bayesian based CNN | 19 | 400 | X-ray | SDR:  82.11% < 2.0 mm  88.63% < 2.5 mm  92.28% < 3.0 mm  95.96% < 4.0 mm |
| Noothout et. al, 2020 [22] | Fully Convolutional Network | 19 | 400 | X-ray | MSE:  0.95 mm on test set 1  1.07 mm on test set 2 |
| Takeda et. al, 2020 [23] | CNN with fully connected layers | 4 | 400 | X-ray | MAE (SGD Optimizer): 1.67 ± 1.77 mm  SDR (SGD Optimizer)  29.5% < 2.0 mm |
| Zeng et. al, 2021[24] | Cascaded CNN | 19 | 400 | X-ray | SDR:  76.82% < 2 mm  95.58% < 4mm |
| Kim et. al, 2021[25] | Multi-Stage CNN | 15 | 430 Lateral + 430 MIP Lateral | CBCT | MRE: 1.03 ± 1.29 mm  SDR:  87.13% < 2.0 mm  96.59% < 4.0 mm |
| Wang et. al, 2021[26] | Iterative Deep CNN | 19 | 400 | X-ray | SDR:  87.51% < 2.0 mm |
| Šavc et. al., 2022 [27] | CNN – Spatial Configuration Net | Public: 19  Private: 72 | Public: 400  Private: 4695 | X-ray | Public:  MRE: 1.13 ± 1.11 mm on test set 1 of ISBI  85.61% < 2.0 mm  Private:  MRE: 11.26 ± 17.51 px |
| Rashmi et al. 2022 [28] | Combination of U-net based CNN and heatmap regression | 19 | 400 | X-ray | 85.36% < 2mm |
| Nishimoto et. al, 2023 [29] | Multi-phased CNN (Transfer Learning for feature extraction) | 16 | 120 | CT | MSE:  11.60 px in phase-1  4.66 px in phase-2 |
| Neeraja et. al, 2023 [30] | Attention based CNN | 19 | 400 | X-ray | SCR: 97.72%  SDR:  *Test 1*  88.06% < 2.0 mm  98.68% < 4.0 mm  *Test 2*  91.26% < 2.0 mm  95.15% < 4.0 mm |

1. **Challenges in Cephalometric Landmark Detection**

The process of cephalometric analysis starts with the identification of the landmarks which are used for further measurements. Identifying these landmarks precisely requires expert level knowledge and it is still time consuming. It takes about 16 minutes, of which about 4 minutes if for the identification of the landmarks [x]. By developing automated systems that can identify the landmarks, we can save this valuable time and let the experts attend other important things. Through this survey, we have identified a few challenges that is hindering the overall reliability of such systems.

*6.1 Lack of open-source dataset*

From our survey, most researchers use the ISBI 2015 Grand Challenge dataset which only contains 400 images of which only 150 is used for training. This does not provide the models with a lot of patterns to learn from. Many researchers also use private data that they have obtained from institutes but they are inaccessible by other fellow researchers that can potentially improve the metrics such as SDR more. Also, since most researchers have used the ISBI dataset for comparison, it raises a question about the generalization of the proposed systems as the test set size is not large either (150 + 100). Thus, there is a very compelling requirement for a large, public dataset.

* 1. *Requirement of high computational power*

Deep convolutional networks require high amount of VRAM to train and infer. Also, resizing the input scans to very low resolution results in a loss of information that ultimately contributes to errors. This ultimately increases the cost of such an automatic system as such GPUs are very expensive. This could also potentially bottleneck the research on this area as researchers from developing countries may not have access to such high-end computational power.

* 1. *Lack of practicality*

One of the techniques [23] identified only 4 landmarks. Considering that it only takes about 4 minutes to do 23 measurements in a cephalogram [31] such proposed systems lose their practicality as these automated systems carry a % of error alongside them. An expert can manually mark 4 landmarks in a few seconds and be confident about 0 % error.

* 1. *Lack of applicability*

Some of the studies [19, 23] have less than 50% SDR in the preferred <= 2.0 mm error range, which makes them unattractive. Many of the studies have less than 85% SDR in the <= 2.0 mm which in cases where precision matter the most, such as maxillofacial surgery which can affect the way a person looks, may be unsuitable.

* 1. *Complex landmarks*

Landmarks like pogonion, menton, gnathion, soft tissue pogonion are close to each other and hence share similar visual features in the image. This is also true for lower incisal incision and upper incisal incision. Since they share similar features, models often have a hard time. Landmarks like upper lip and lower lip have gradient edges between the human face and the background and hence are more prone to errors. All these factor in the overall error that an automated system can induce

* 1. *Lack of Diversity in the data*

Most of the methods have been trained and/or evaluated on the ISBI 2015 Grand Challenge dataset. The cephalograms in this dataset was collected from a single hospital in Taiwan. Also, most of the private dataset collected by the researchers are from either a single source or from the same location. Thus, there is a lack of race or ethnic diversity in their data. This hinders the applicability on such models in a diverse scenario. The features of the human skull in one race is different from another and there is no sure-shot way to ensure that a model performing excellent in the ISBI dataset would perform just as good when used in an hospital with patients from multicultural backgrounds or of other races and ethnicities. Also, abnormalities, disabilities, accidents, and so on may distort the shape of the skull and so the systems that are not trained to handle such variations will fail on these data. Thus, there is a need to publish a standardized dataset that contains people of various races, ethnicities, and also includes data from people with distorted skull.

1. **Conclusion**

This survey summarizes this progressive journey by highlighting the development of computer-assisted landmark recognition and the critical function of digital cephalometric radiographs in orthodontics. The advent of machine learning, particularly deep learning and convolutional neural networks (CNNs), has substantially enhanced landmark detection accuracy, expediting analyses while bolstering precision. However, enduring challenges impede further advancements in this domain. Persistent issues such as limited diverse, open-source datasets and high computational demands restrict the adaptability of models across varied demographics and clinical scenarios. Concerns regarding practicality surface as systems exhibit constrained landmark identification, especially impacting crucial applications like maxillofacial surgery. Complex landmark configurations, their proximity, and insufficient diversity in training data contribute to error rates in automated systems. The absence of racial, ethnic, deformed skull diversity in datasets further confines the flexibility of these models. However, considering the progress in the domain so far, if such a diverse data is to be publicized, it will fuel the creation of robust systems. The future scope of this study is to create an automated system that can be outperform the current systems and assist doctors while being explainable and lightweight.

**References**

[1] Wang, C. W., Huang, C. T., Hsieh, M. C., Li, C. H., Chang, S. W., Li, W. C., ... & Ibragimov, B. (2015). Evaluation and comparison of anatomical landmark detection methods for cephalometric x-ray images: a grand challenge. *IEEE transactions on medical imaging*, *34*(9), 1890-1900.

[2] Shukla, S., Chug, A., & Afrashtehfar, K. I. (2017). Role of cone beam computed tomography in diagnosis and treatment planning in dentistry: an update. *Journal of International Society of Preventive & Community Dentistry*, *7*(Suppl 3), S125.

[3] Niraj, L. K., Patthi, B., Singla, A., Gupta, R., Ali, I., Dhama, K., ... & Prasad, M. (2016). MRI in dentistry-A future towards radiation free imaging–systematic review. *Journal of clinical and diagnostic research: JCDR*, *10*(10), ZE14.

[4] Kaur, A., & Singh, C. (2015). Automatic cephalometric landmark detection using Zernike moments and template matching. *Signal, Image and Video Processing*, *9*(1), 117-132.

[5] Koga, H., Taki, K., & Masugi, A. (2023). Efficient Measurement Method: Development of a System Using Measurement Templates for an Orthodontic Measurement Project. *Software*, *2*(2), 276-291.1309.

[6] J. Yang, X. Ling, Y. Lu et al., “Cephalometric image analysis and measurement for orthognathic surgery,” Medical & Biological Engineering & Computing, vol. 39, no. 3, pp. 279–284, 2001.

[7] Gupta, A., Kharbanda, O. P., Sardana, V., Balachandran, R., & Sardana, H. K. (2016). Accuracy of 3D cephalometric measurements based on an automatic knowledge-based landmark detection algorithm. *International journal of computer assisted radiology and surgery*, *11*, 1297-1309.

[8] Vezzetti, E., Marcolin, F., Tornincasa, S., Ulrich, L., & Dagnes, N. (2018). 3D geometry-based automatic landmark localization in presence of facial occlusions. *Multimedia Tools and Applications*, *77*, 14177-14205.

[9] Alessandri-Bonetti, A., Sangalli, L., Salerno, M., & Gallenzi, P. (2023). Reliability of artificial Intelligence-Assisted cephalometric analysis. A Pilot Study. *BioMedInformatics*, *3*(1), 44-53.

[10] S. Chakrabartty, M. Yagi, T. Shibata et al., “Robust cephalometric landmark identification using support vector machines,” in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. 825–828, Vancouver, Canada, May 2003.

[11] P. Vucinic, Z. Trpovski, and I. Scepan, “Automatic landmarking of cephalograms using active appearance models,” *European Journal of Orthodontics*, vol. 32, no. 3, pp. 233–241, 2010.

[12] J. Keustermans, W. Mollemans, D. Vandermeulen, and P. Suetens, “Automated cephalometric landmark identification using shape and local appearance models,” in *Proceedings of 20th International Conference on Pattern Recognition*, pp. 2464–2467, Istanbul, Turkey, August 2010.

[13] Mirzaalian, H., & Hamarneh, G. (2014, April). Automatic globally-optimal pictorial structures with random decision forest based likelihoods for cephalometric x-ray landmark detection. In *IEEE ISBI*.

[14] Chen, S. (2015, March). Tooth segmentation system with intelligent editing for cephalometric analysis. In *Medical Imaging 2015: Biomedical Applications in Molecular, Structural, and Functional Imaging* (Vol. 9417, pp. 653-658). SPIE.

[15] Wang, S., Li, H., Li, J., Zhang, Y., & Zou, B. (2018). Automatic analysis of lateral cephalograms based on multiresolution decision tree regression voting. *Journal of healthcare engineering*, *2018*.

[16] Montúfar, J., Romero, M., & Scougall-Vilchis, R. J. (2018). Automatic 3-dimensional cephalometric landmarking based on active shape models in related projections. *American Journal of Orthodontics and Dentofacial Orthopedics*, *153*(3), 449-458.

[17] Suhail, S., Harris, K., Sinha, G., Schmidt, M., Durgekar, S., Mehta, S., & Upadhyay, M. (2022). Learning Cephalometric Landmarks for Diagnostic Features Using Regression Trees. *Bioengineering*, *9*(11), 617.

[18] Montúfar, J., Romero, M., & Scougall-Vilchis, R. J. (2018). Hybrid approach for automatic cephalometric landmark annotation on cone-beam computed tomography volumes. *American Journal of Orthodontics and Dentofacial Orthopedics*, *154*(1), 140-150.

[19] Dai, X., Zhao, H., Liu, T., Cao, D., & Xie, L. (2019). Locating anatomical landmarks on 2D lateral cephalograms through adversarial encoder-decoder networks. IEEE Access, 7, 132738-132747.

[20] Kim, H., Shim, E., Park, J., Kim, Y. J., Lee, U., & Kim, Y. (2020). Web-based fully automated cephalometric analysis by deep learning. Computer methods and programs in biomedicine, 194, 105513.

[21] Lee, J. H., Yu, H. J., Kim, M. J., Kim, J. W., & Choi, J. (2020). Automated cephalometric landmark detection with confidence regions using Bayesian convolutional neural networks. BMC oral health, 20, 1-10.

[22] Noothout, J. M., De Vos, B. D., Wolterink, J. M., Postma, E. M., Smeets, P. A., Takx, R. A., ... & Išgum, I. (2020). Deep learning-based regression and classification for automatic landmark localization in medical images. IEEE transactions on medical imaging, 39(12), 4011-4022.

[23] Takeda, S., Mine, Y., Yoshimi, Y., Ito, S., Tanimoto, K., & Murayama, T. (2021). Landmark annotation and mandibular lateral deviation analysis of posteroanterior cephalograms using a convolutional neural network. Journal of Dental Sciences, 16(3), 957-963.

[24] Zeng, M., Yan, Z., Liu, S., Zhou, Y., & Qiu, L. (2021). Cascaded convolutional networks for automatic cephalometric landmark detection. Medical Image Analysis, 68, 101904.

[25] Kim, M. J., Liu, Y., Oh, S. H., Ahn, H. W., Kim, S. H., & Nelson, G. (2021). Automatic cephalometric landmark identification system based on the multi-stage convolutional neural networks with CBCT combination images. Sensors, 21(2)

[26] Wang, L., Ma, L., Li, Y., Niu, K., & He, Z. (2021). A DCNN system based on an iterative method for automatic landmark detection in cephalometric X-ray images. Biomedical Signal Processing and Control, 68, 102757.

[27] Šavc, M., Sedej, G., & Potočnik, B. (2022). Cephalometric Landmark Detection in Lateral Skull X-ray Images by Using Improved SpatialConfiguration-Net. Applied Sciences, 12(9), 4644.

[28] Rashmi, S., Murthy, P., Ashok, V., & Srinath, S. (2022). Cephalometric Skeletal Structure Classification Using Convolutional Neural Networks and Heatmap Regression. SN Computer Science, 3(5), 336.

[29] Nishimoto, S., Saito, T., Ishise, H., Fujiwara, T., Kawai, K., & Kakibuchi, M. (2023). Three-Dimensional Craniofacial Landmark Detection in Series of CT Slices Using Multi-Phased Regression Networks. Diagnostics, 13(11), 1930.

[30] Neeraja, R., & Anbarasi, L. J. (2023). CephXNet: A Deep Convolutional Squeeze-and-Excitation model for Landmark Prediction on Lateral Cephalograms. IEEE Access.

[31] Tsorovas, G., & Linder-Aronson Karsten, A. (2010). A comparison of hand-tracing and cephalometric analysis computer programs with and without advanced features—accuracy and time demands. *The European Journal of Orthodontics*, *32*(6), 721-728.