



OPEN

## An interpretable machine learning approach to study the relationship between retrognathia and skull anatomy

Masrour Makaremi<sup>1,2✉</sup>, Alireza Vafaei Sadr<sup>3,4</sup>, Benoit Marcy<sup>5</sup>, Ikram Chraïbi Kaadoud<sup>6</sup>, Ali Mohammad-Djafari<sup>7</sup>, Salomé Sadoun<sup>1</sup>, François De Brondeau<sup>1</sup> & Bernard N'kaoua<sup>2</sup>

Mandibular retrognathia (C2Rm) is one of the most common oral pathologies. Acquiring a better understanding of the points of impact of C2Rm on the entire skull is of major interest in the diagnosis, treatment, and management of this dysmorphism, but also permits us to contribute to the debate on the changes undergone by the shape of the skull during human evolution. However, conventional methods have some limits in meeting these challenges, insofar as they require defining in advance the structures to be studied, and identifying them using landmarks. In this context, our work aims to answer these questions using AI tools and, in particular, machine learning, with the objective of relaying these treatments automatically. We propose an innovative methodology coupling convolutional neural networks (CNNs) and interpretability algorithms. Applied to a set of radiographs classified into physiological versus pathological categories, our methodology made it possible to: discuss the structures impacted by retrognathia and already identified in literature; identify new structures of potential interest in medical terms; highlight the dynamic evolution of impacted structures according to the level of gravity of C2Rm; provide for insights into the evolution of human anatomy. Results were discussed in terms of the major interest of this approach in the field of orthodontics and, more generally, in the field of automated processing of medical images.

With an incidence of 44% among 12-year-old children, Class II by mandibular retrognathia (C2Rm) is the most common craniofacial dysmorphism<sup>1,2</sup> with an unequal distribution according to sex and ethnicity<sup>3</sup> as well as a hereditary component. This pathophysiology is characterized by a too short mandible that leads to a loss of occlusion with the maxilla. Conversely, prognathism corresponds to the situation where either the maxilla or the mandible is positioned forward compared to normal positioning.

Retrognathia, and malocclusions in general, are associated with changes in craniofacial morphology. For example, a link between fusions of cervical vertebrae, deviations in the cranial base and mandibular retrognathia has been documented radiographically<sup>4</sup>. Indeed, the association between the angle of the cranial base and the morphology of the cervical spine plays a central role in establishing the diagnosis of malocclusion. These morphological changes are also important to consider when searching for etiology, especially in cases of severe skeletal malocclusions<sup>5</sup>.

Overall, the morphology of the face depends on many factors such as sex, ethnicity, climate, nutrition, genetic constitution, etc<sup>6</sup>. Regarding ethnicity, for example, cephalometric analysis of Bangladeshi and Japanese adults<sup>7</sup> showed that the mandibular antero-posterior position among the Bangladeshi was more protrusive compared with that of Japanese subjects. Proffit et al.<sup>8</sup> found that the SNA angle (the angle between the sella, nasion and anterior nasal spine) was significantly greater in African Americans than in Caucasians, resulting in a more prognathic facial profile in African Americans compared to Caucasians. The results of these studies are valuable for orthodontics, insofar as the analysis of craniofacial morphology provides information on the typology of the patient, on the shift between the bony bases, as well as on the diagram of mandibular growth. These elements

<sup>1</sup>Dentofacial Orthopedics Department (UFR de Sciences Odontologiques), University of Bordeaux, Bordeaux, France. <sup>2</sup>Bordeaux Population Health (Team ACTIVE), INSERM U1219, University of Bordeaux, Talence, France. <sup>3</sup>Department of Public Health Sciences, College of Medicine, The Pennsylvania State University, Hershey, PA 17033, USA. <sup>4</sup>Theoretical Physics Department, University of Geneva, 1211 Geneva, Switzerland. <sup>5</sup>University of Toulouse, Toulouse, France. <sup>6</sup>LUSSI Department, IMT Atlantique, 29238 Brest, France. <sup>7</sup>LSS, Centrale Supélec, Paris, France. ✉email: masrour@makaremi.fr

(in particular, the extent of malocclusion and the changes in craniofacial morphology) constitute the basis of the diagnosis and make it possible to develop a treatment strategy, or a treatment plan, adapted to each type of malocclusion, according to its importance and its orientation.

Moreover, the understanding of C2Rm represents an anthropological challenge insofar as Retrognathia is at the heart of the evolution of hominins<sup>9–12</sup>.

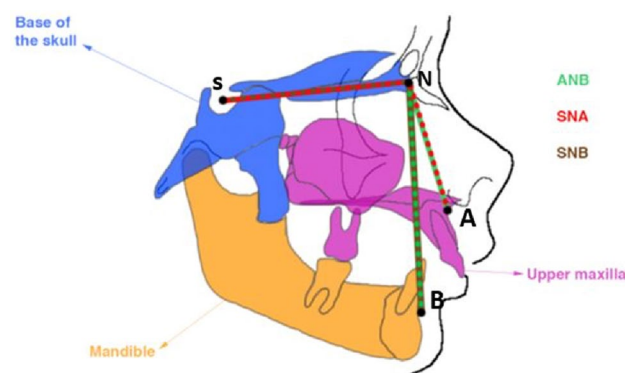
Since the differentiation of the genus *Homo*, a concomitant kinematics of transition from more prognathic forms to more retrognathic faces has been observed<sup>13</sup> but its phylogenetic and ontogenetic mechanisms are still unknown. This decrease in mandibular length was accompanied by an increase in brain volume. Even if the relationship between the length of the mandible and the craniofacial architecture remains to be clarified<sup>14</sup>, the transition from prognathism to an increasingly accentuated retrognathia has been accompanied by a series of craniofacial modifications central to human evolution<sup>11</sup>. Mandibular growth is sensitive to food consistency and changes in occlusal forces (mastication/breathing/swallowing). However, the industrial revolution led to a new change in eating habits with the emergence of processed products, mostly soft, and a more sedentary lifestyle. These changes caused an imbalance in the forces experienced by the skull during its development, which ultimately led to an increase of the number of cases of C2Rm and of their severity<sup>15</sup>. A better understanding of the craniofacial changes induced by the increase in mandibular retrognathia therefore appears to be a good model for studying the craniofacial evolution of hominins.

In orthodontic practice, the sagittal relationship between the maxilla and the mandible is an important diagnostic criterion for C2Rm. A classical parametric model with three standard cephalometric angles is widely used to characterize this relationship<sup>16</sup>: the Maxillary-Nasion-Mandibular (ANB) angle, the Sella-Nasion-Maxillary (SNA) and Sella-Nasion-Mandibular (SNB) angles. As illustrated in Fig. 1, the ANB angle expresses the relative position of the maxilla to the mandible, whereas the SNA and SNB angles indicate respectively the anteroposterior projection of the maxilla and the mandible.

Class II by mandibular retrognathia is characterized by an increased ANB angle (ANB [ $> 4$ ]) and a decreased SNB angle (SNB [ $< 76$ ]). The increased ANB angle implies that the mandible occupies a retrusive position in relation to the maxilla, taking as a reference the base of the skull. The reduction in the SNB angle induces that the mandible is in a retrusive position relatively to the base of the skull.

While being widely used in orthodontic practice, these cephalometric angles present some major limits. First, they focus on the jawbone, and therefore fail to highlight the rest of the structures potentially impacted by C2Rm physiopathology. Regarding this limit, decades of research based on cephalometric parameters investigated craniofacial alterations correlated to C2Rm. These studies aim to identify the impact of C2Rm on the shape of the skull, but also to predict the pattern of mandibular growth. Indeed, growth prediction plays a significant role in accurate diagnosis and treatment planning of orthodontics patients. For example, it has been considered that the unique pattern of pneumatization of the frontal sinus could be used as a predictor of growth<sup>17</sup>. Other statistical correlations have been highlighted, such as cervical vertebrae<sup>18</sup> or the cranial base<sup>19</sup>, but the results thus obtained vary from one study to another (depending on the sample size, the employed method, etc.). This is partly because the methods typically used rely on biologically relevant landmarks (based on predefined biological knowledge) that are manually assigned to images. This assignment is error-prone and makes it difficult to discover new anatomical features related to retrognathia that may prove relevant in understanding and managing this malocclusion<sup>20</sup>.

These limits have been partly circumvented by the development of geometric morphometrics<sup>21</sup>. Landmark-based geometric morphometric methods involve summarizing shape in terms of a landmark configuration. The direct analysis of these points would not be relevant due to the possible variations in position, orientation or scale of the specimens. Therefore, variations unrelated to shape are mathematically eliminated by superposition methods. The Procrustes approach is one of the superposition methods classically used to quantify the similarity or dissimilarity between two landmark configurations (physiological class I and pathological class II, for example) by calculating the Procrustes distance between them<sup>22</sup>.



**Figure 1.** Retrognathia and the three standard cephalometric angles widely used in evaluating anteroposterior apical base relationship: ANB in green, SNA in red, SNB in brown (ANB = SNA – SNB). The unit of measurement of angles is the degree.

But, even if the geometric morphometric techniques have found numerous applications in the biological sciences, the limits of this approach have been widely underlined<sup>23</sup>. In particular, errors may be related to the superimposition procedure. For example, Rohlf<sup>24</sup> shows that in certain circumstances, the variation observed after a Procrustes superimposition is very different from the model on which it was based. And Walker<sup>25</sup> offered a detailed description of some problems in estimating the pattern of variances and covariances of landmarks after superimposition. Applied to retrognathia, this means that sometimes the mandible position cannot be controlled during the image acquisition process. As a result, its morphological or metric characteristics cannot be compared between individuals since any observed differences would reflect more this methodological bias than a true interindividual variability.

Another important limitation of geometric morphometry applied to the analysis of medical images lies in the difficulty of understanding the dynamic of interactions between structures at different stages of evolution of a physiological or pathological process. For example, Gkantidis and Halazonetis<sup>26</sup> evaluated the morphological covariation between the face and the basicranium (midline and lateral line) at two specific developmental stages (children and adults). Lateral cephalometric radiographs were digitized and a total of 28 landmarks were placed in three areas (the medial cranial base, the lateral cranial base and the face). Geometric morphometric methods were applied and partial least squares analysis was used to assess the correlation between the three shape blocks. The results highlight the interactions between the structures of the middle cranial base and facial patterns in the group of children, as well as between the lateral elements of the cranial base and facial patterns at a later lifestage, such as during adulthood. Freudenthaler et al.<sup>27</sup> assessed the relationship between craniofacial shape and malocclusion using geometric morphometry. Cephalometric radiograph tracings of 88 untreated Caucasians (age range: 7–39 years) were assigned to four groups according to the nature of their occlusion.

The results show that the craniofacial shape was clearly associated with dental malocclusion with considerable variations between the four groups. But if these studies have allowed some advances in the field of orthodontics, Lieberman et al.<sup>28</sup> conclude that, overall, these methods require independent structure analysis and fail to isolate and define the actual interacting morphogenetic units, identify and quantify their direct and indirect interactions, and understand the processes through which they interact.

In this context, it appears relevant to develop new morphometric tools which do not involve defining areas of interest a priori, nor defining shapes using a set of benchmarks, and which allow to follow the dynamic evolution of structures impacted by the severity of mandibular retrognathia.

In recent years, the progress of research in the use of Artificial Intelligence (AI) in orthodontics has been considerable and many applications have been developed. For example, Thurzo et al.<sup>29</sup>, in a documented review, mention the use of machine learning algorithms to automate the analysis of dental images, the development of AI tools to detect oral diseases or to optimize diagnosis and treatment planning. In the field of medical imaging, Deep Neural Networks (DNNs) and particularly Convolutional Neural Networks (CNNs) as well as additional interpretable methods of saliency maps such as class activation mapping (CAM) and its variant Score-CAM<sup>30</sup> have become very popular tools. They are used in particular to locate biomarkers, i.e., observable or measurable signals that reflect the presence, severity or evolution of a disease and that can be considered as discriminating regions in abnormal images. This approach has been applied to various medical fields<sup>31</sup>, such as retinal fundus images<sup>32</sup>, cancer detection<sup>33</sup>, pneumonia detection or Alzheimer's diagnosis<sup>34</sup>. A recurrent and common result of these studies is that CNNs outperform practitioners in pathology classification<sup>35</sup>.

CAM and Score-CAM, as interpretable methods, can be applied to CNNs to provide visual identification (using a heatmap) of regions of interest (ROI), i.e. areas in medical images provided as input, on which the CNN actually relied to carry out its classification. The regions of the medical images provided as input are thus classified according to color ranging from hot zones (red color, i.e. extremely relevant for classification) to cold zones (blue color) passing through intermediate zones (orange color). Applied to the retrognathia, this means that the CNN performs a classification of the input images according to criteria provided by the practitioner (in this instance, the measurement of the angles ANB [0–5], SNB [ $< 77$ ], and SNA [79–85]) and that the interpretability methods identify the ROIs actually used by the classifier to distinguish between the input images (in class I physiologic and class II pathological).

It is therefore important to verify that the ROIs identified by the CNN correspond to biomarkers relevant from a medical point of view and already identified in literature. But this approach can make it possible to identify new regions of interest, which could allow for advances in understanding, treating or managing the pathology considered. In this perspective, interpretable AI algorithms become tools for the discovery and extraction of new knowledge, offering an innovative scientific approach to health issues<sup>36,37</sup>.

Despite the limitations that have been noted by members of the scientific community<sup>38</sup>, these methods may serve as tools in the search for information contained in sets of images. The foundations of this approach have been clarified by Wöber et al.<sup>20</sup> in a study aimed to characterize the phenotype of a fish population based on the animal's habitat. The authors start from the classic observation that assigning landmarks is tedious and error-prone and that predefined landmarks can miss information that is not obvious to the human eye. They document the new data analysis methods proposed by the machine learning (ML) community to identify subtle features in images, with an excellent predictive accuracy. The authors also insist on the contribution of interpretability methods by means of which it is possible to overcome the problem of the black box of ML methods, by making internal representations (ROI) accessible via saliency maps. The authors then conclude on how deep learning improves traditional morphometric analysis, particularly in terms of predictive performance.

Our work is inscribed in the same perspective, and aims to apply this approach to health, in particular to mandibular retrognathia, which is a central field in orthodontic research.

The main objective of our work was to propose a methodological pipeline coupling CNN and interpretability approaches to extract knowledge concerning the shape of the skull and, in particular, to evaluate the impact of a pathological process (mandibular retrognathia) on the entire skull (this is the first study ever carried out in this

area). This objective required the generation of a unique saliency map, named “global activation map”. Taking into account all pathological individuals, this map aimed to crush inter-individual variability and to highlight the biomarkers common to the totality of the study population.

A secondary and complementary objective was to use this methodology to monitor the dynamic evaluation of impacted skull structures according to the severity of retrognathia; towards this end, saliency maps were generated and compared according to different C2Rm severity levels.

## Methods: Morphological information extraction with an Interpretable CNN (MIE-ICNN) Methodology

This section presents the main steps of our methodology for extracting morphological information using interpretable convolutional neural networks (MIE-ICNN). The characteristics of the dataset are first presented, encompassing the acquisition, selection and labeling of data as well as the pre-processing step (Fig. 2A). Next, the architectural configuration and parameters of our interpretable CNN (Fig. 2B) and then the method for generating global activation maps are described (Fig. 2C).

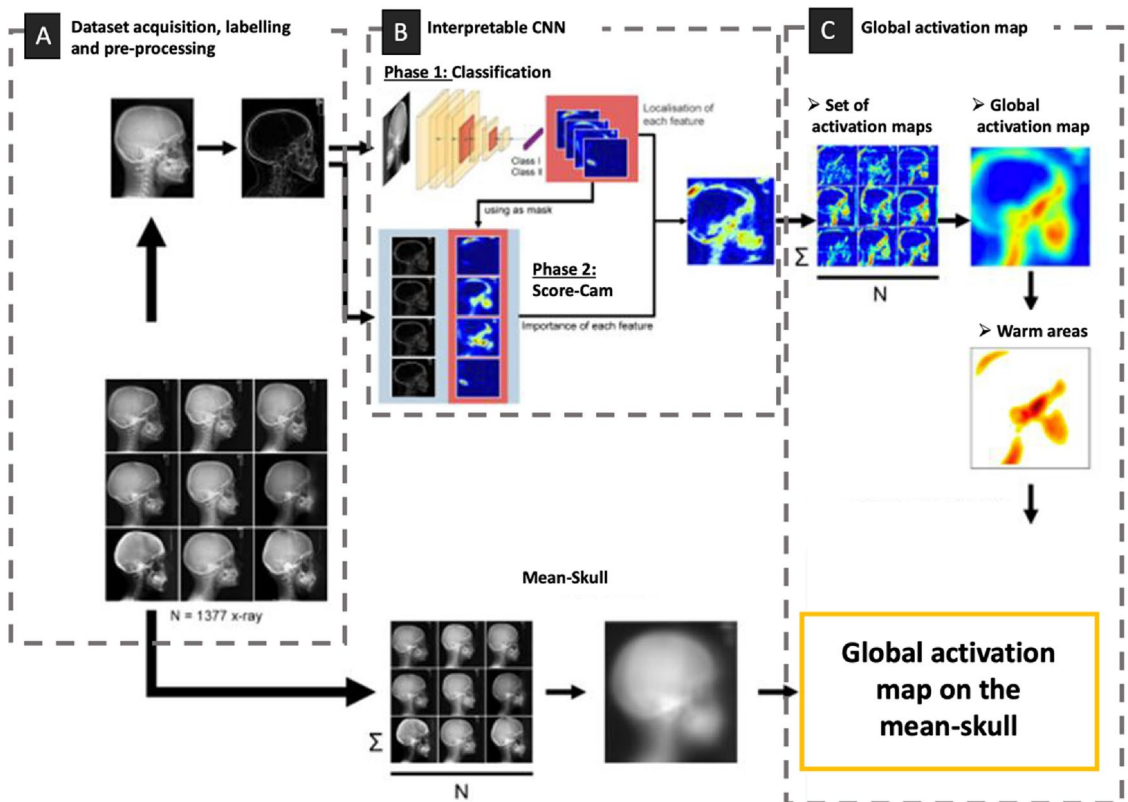
### Dataset acquisition and labeling

Our study focused on a compilation of cephalometric images, including 2694 French orthodontic patients. These were patients of both sexes and a wide age range (no selection based on age or sex was made in this study), complaining of dentofacial dysmorphism and/or malocclusion (1/3 males and 2/3 females with a mean age of  $12.63 \pm 5.06$  years).

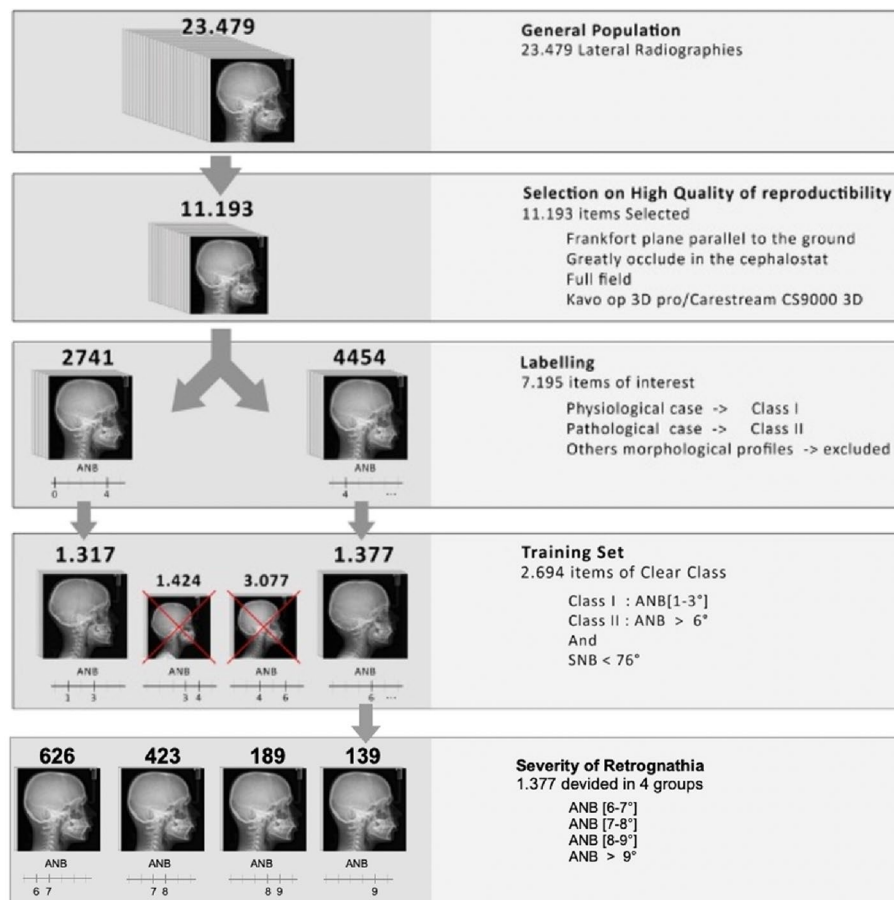
These 2694 X-rays were selected from an initial pool of 23,479 X-rays from four separate clinics. To guarantee the fidelity of the data, this initial pool included only radiographs adhering to the standardized positioning of the cephalostat, and coming from the same radiography device (CARESTREAM CS 9000-C).

A meticulous image selection and labeling process, shown in Fig. 3, was then performed. After review by dentofacial orthopedist practitioners, a subset of 11,193 images was chosen based on high quality attributes, adherence to anatomical standards (including Frankfurt plane orientation<sup>39</sup>) and a complete skull X-ray.

Then, only the images showing a Class I and a Class II with a retrognathic mandible were retained (7100 X-ray) and classified into two categories, “physiological” (2741) and “pathological” (4454), according to angle measurements (ANB [ $> 4$ ] and SNB [ $< 76$ ]). Images with other types of malocclusion were thus excluded. To reinforce the discrimination between the categories, borderline cases between physiological-pathological (ANB [ $-1$ ] –  $1$ )) were excluded. The final training set included 1317 Class I physiological images and 1377 Class II images.



**Figure 2.** Synthetic presentation of the main steps of the morphological information extraction methodology with an Interpretable CNN (MIE-ICNN) that we used. (A) Dataset acquisition, labeling and pre-processing; (B) Interpretable Deep CNN training-assessment-testing process; (C) Generation of global activation map.



**Figure 3.** Radiography images selection and labeling process for dataset elaboration.

In a second step, and in order to evaluate the modifications of the shape of the skull according to the severity of the retrognathia, the class 2 images were subsampled into 4 categories according to the amplitude of the angle ANB (from [6° to 7° [and gradually increasing to [9°<] in steps of 1). We therefore referred to Steiner's classification which defines the extent of retrognathia according to the value of the angle ANB<sup>40</sup>.

The distribution included 626 individuals exhibiting an ANB angle between [6° and 7°[, 423 exhibiting an ANB angle between [7°-8°[, 189 exhibiting an ANB angle between [8°-9°[, and 139 whose ANB angle was comprised between [9°<].

### Pre-processing step

The image preprocessing step plays a crucial role in improving the efficiency of CNN and optimizing the memory consumption of the graphics processing unit (GPU). For this, we performed the following pre-processing operations:

(i) Resizing:

Considering the importance of resizing pre-processing<sup>41</sup> and variability in our dataset, all X-ray images were standardized to 256 × 256 pixel dimensions. This resizing reduces potential distortions resulting from variability in craniofacial growth patterns (due to age, gender, ethnicity, etc.), thereby focusing attention on pathology-induced disparities.

(ii) Normalization:

To account for sigmoid or softmax functions related to the backpropagation algorithm, we normalized the pixel intensity values to a range between 0 and 1 before entering them into the neural network<sup>42</sup>.

(iii) Sobel filtering:

The use of a Sobel process<sup>43</sup> enhanced edge detection capabilities by emphasizing edge-related intensity changes.

## Interpretable deep CNN architecture

In this section, we first describe the architecture and parameters of Deep CNN, then the interpretability process that we have implemented.

### Classification

To extract bone shape features from the lateral radiographies, first, a binary classification model was built and used to classify data in class I and class II as illustrated in Fig. 2B (Phase 1).

To train the model and extract high-level features<sup>44,45</sup>, we utilized a deep-learning model. The model architecture includes seven CNN layers with 867,178 parameters that exhibited the highest accuracy regarding identification of the class, at a percentage of 97%. Each of these layers was paired with a batch normalization and a ReLU activation function<sup>46</sup>.

A SoftMax activation function was used on the last fully connected layer to make the predicted class. The model was fine-tuned using an ADAM optimizer<sup>47</sup> with a decreasing learning rate starting at  $10^{-3}$  and decreasing by 0.95 per 100 epochs. A total of 1000 epochs was reached with a batch size of 100 in a P100 Nvidia GPU, using 70 GB of memory.

In Deep Learning classification methods, when the number of parameters is large, some form of regularization is needed to ensure small generalization errors and avoid overfitting<sup>48</sup>.

In this context, statistical learning theory offers several different tools capable of controlling and avoiding overfitting. In our study, we increased and dropped data so as to control generalization errors, following the recommendations of Zhang et al.<sup>48</sup>:

- (i) Regarding data augmentation, during training, the preprocessed X-rays are augmented by the following operations: random horizontal and vertical translations; rotation with steps of 10 degrees, width and height shift, and zoom within the range of 3%. The rotation of the images was performed based on the center of the image, which is located at pixel coordinates (128, 128). This served as the reference point for the rotation process. The time taken for each rotation, was approximately 0.01 s per image.
- (ii) We used a dropout layer with a drop rate of 0.5.

To determine the model performance, we divided the dataset into training and testing subsets with an 80–20 ratio for developing and evaluating our CNN model. We used a 5-step cross-validation approach, which is common practice for datasets of this size. This technique involved splitting the dataset into five subsets, using each subset once as a test set, while the remaining four subsets were used for training in each iteration. In other words, the distribution of classes in each iteration consisted of 1317 samples (1054 for training and 263 for testing) for class I and 1377 samples (1102 for training and 275 for testing) for class II. Using cross-validation resulted in an accuracy rate of 97%.

### Interpretability through saliency maps

Once the deep CNN trained, we used it for a classification task in order to predict the presence or the absence of a C2Rm.

As illustrated in Fig. 2B (Phase 2), for each image classified as positive to C2Rm, a saliency map is generated using the Score-CAM technique. This step involves the implementation of the Score-Weighted Class Activation Mapping (Score-CAM) technique to generate informative saliency maps. These saliency maps highlight the craniofacial structures that significantly influence the CNN classification process.

The Score-CAM technique works as a multi-step process, implementing mathematical procedures to highlight the underlying elements that guide CNN predictions:

- (i) Prediction Score Extraction:

The final prediction score ( $S_c$ ) generated by the CNN represents the network's level of confidence in categorizing a given image, effectively quantifying how certain the network is about its decision.

- (ii) Backpropagation of Scores:

The identification of the relevant zones allowing the CNN to carry out its prediction goes through a process of backpropagation. This process requires: (1) quantifying the influence of the activation of each channel on the global prediction score; (2) to apply a function that accentuates regional activations that contribute significantly to the prediction<sup>49</sup>.

Mathematically, the gradients of the predicted class score ( $S_c$ ) regarding the activations ( $A_i$ ) of the last convolutional layer (i) are calculated:

$$G_i = \frac{\partial S_c}{\partial A_i}$$

These gradients quantify the influence of each neuron's activation on the overall prediction score.

- (iii) Weighted Activation Aggregation:

The weighted activations ( $M_i$ ) for the last convolutional layer were obtained by multiplying the gradients ( $G_i$ ) by the ReLU activation of the corresponding neuronal output ( $A_i$ ). This procedure aims to accentuate the activations that contribute significantly to the prediction:

$$M_i = G_i \times \text{ReLU}(A_i)$$

(iv) Saliency Map Generation:

The saliency map ( $L$ ) was created by calculating the weighted average of the activations ( $M_i$ ) from all channels using a global average pooling. This average activation value ( $\alpha$ ) highlights the significance of each channel of CNN in influencing the network's decision.

$$\alpha = \text{GlobalAvgPool}(M_i)$$

The final saliency map ( $S$ ) was obtained by linearly combining the weighted activations ( $M_i$ ) using the calculated importance ( $\alpha$ ) as the weighting element. This map captures areas of the image that contribute substantially to CNN's classification decision:

$$S = \sum (\alpha \times M_i)$$

This Score-CAM methodology was therefore applied to the 377 images labeled as C2Rm, generating a corresponding set of saliency maps. The next step consisted of averaging these individual saliency maps.

### Global activation map generation

The last step involved calculating an overarching activation map, referred to as the global activation map (depicted in Fig. 2C). To calculate the global activation map: let  $N$  be the number of images, each with its corresponding saliency map  $S_i$  for  $i = 1, 2, \dots, N$ ; these saliency maps are aligned based on the cranial positions; the global activation map  $G$  is obtained by averaging the individual saliency maps:

$$G = \frac{1}{N} \sum_{i=1}^N S_i$$

The map thus obtained constitutes a synthesis of individual saliency maps, harmonizing the cranial positions for a coherent analysis. To promote interpretability, this map is projected onto an averaged skull, corresponding to a synthesized representation made from all the input radiographs of our dataset.

### Activation map according to the severity of class II

Four saliency maps were then created according to the severity of the pathology, using the same method described above. Each map grouped individuals with ANB angles of  $[6^\circ-7^\circ]$ ,  $[7^\circ-8^\circ]$ ,  $[8^\circ-9^\circ]$  and  $[9^\circ <]$ , in order to study the evolution of the maps of salience as a class II gravity function according to Steiner's classification.

### Quantification of regions of interest

To quantify the areas of interest, we defined 2 metrics that we called "class score" and "hot surface". In order to calculate these metrics, we used the class scores (non-dimensional and normalized) obtained for each of the pixels via the score-cam method. The pixels below a threshold (0.3) have been eliminated in order to keep only the discriminating pixels. Then the various underlying bone structures were delineated and within these areas were calculated:

- *The class score*: the average of the class scores of all the pixels in the area;
- *The hot surface*: the percentage of the area whose pixels exceed the defined threshold.

These two metrics made it possible to quantify the extent and intensity of the dysmorphism in the areas of interest highlighted by the global activation maps.

### Ethical approval

The study was approved by the Research Ethics Committee of the University Hospital of Bordeaux (reference number CER-BDX-2023-25). As the study was a retrospective review and analysis of fully anonymized lateral radiographies, the Research Ethics Committee of the University Hospital of Bordeaux (reference number CER-BDX-2023-25) waived the requirement for informed consent. This study was conducted in accordance with the relevant guidelines and regulations.

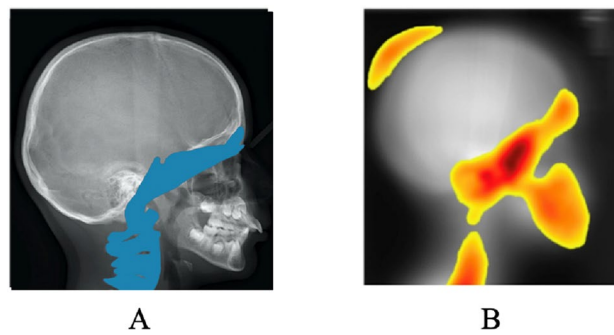
## Results

### Qualitative analysis: descriptive results analysis of saliency maps

This section presents the results obtained, first, on the whole dataset and, second, on datasets related to the different stages of C2Rm evolution, i.e., for each severity level considered in our study.

#### Global activation map

Figure 4A shows the anatomical structures (i.e. biomarkers) identified in literature as being impacted by C2Rm pathology (yellow part of the image). Figure 4B shows the global activation map obtained by the MIE-ICNN method from all class II images used in our study. This global activation map shows, through a heat map, the



**Figure 4.** Craniofacial areas: (A) Anatomical structure identified in literature as being impacted by mandibular retrognathia: the frontal sinus, the base of the skull, the sphenoid and the cervical vertebrae; (B) Anatomical structure highlighted by our global activation map according to MIE-ICNN: the structures already identified in literature, as well as the parietal bone.

areas of the images used by the network to issue its prediction. The heat map indicates, using a color code, the importance of the pixels of the image that take part in the prediction of the class. Yellow corresponds to pixels of low importance, orange to pixels of medium importance, and red to pixels of high importance. This system makes it possible to effectively and visually highlight the areas that have been considered important by the classifier in carrying out his classification.

Our results show the presence of different hot zones. Moreover, the analysis of Fig. 4 indicates that all of the highlighted areas appear to be superimposed on bony structures. The area in the upper left is superimposed on the parietal bone. The area in the lower right is centered on the jaws. The third zone is divided into two segments. The first segment covers the cervical vertebrae. It begins at C6 and continues up to the atlas. Activation follows a plane of gravity along this vertebral axis. The inflection point appears to be centered on the sphenoid bone, located in the medial part of the skull base. Finally, the second segment starts at the sphenoid and follows the base of the skull to the frontal sinus.

Our results show that there is a morphotype specific to the population extracted by the deep CNN, which is superimposed on the structures identified in literature (cervical vertebrae, sphenoid, base of the skull to the frontal sinus).

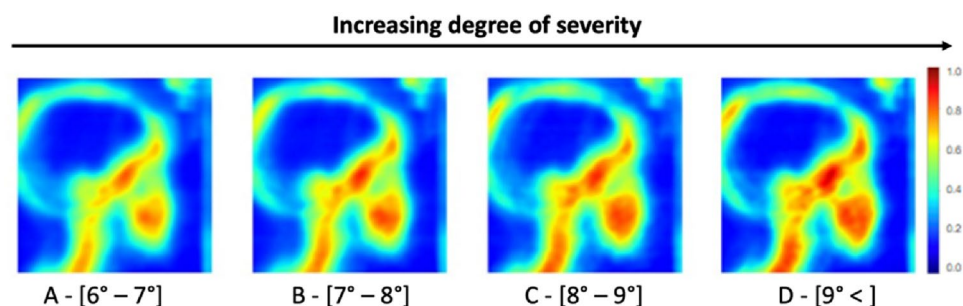
The main novelties here are: (1) the frontal sinus, whose role in C2Rm pathology is debated in literature, and which appears to be discriminating according to our global activation map, (2) the parietal bone, which appears to be discriminating whereas until now it was not cited in literature as an area involved in C2Rm.

#### *Saliency map according to the degree of severity of retrognathia*

In a second step, we identified the craniofacial structures involved according to the degree of evolution of the pathology using our MIE-ICNN methodology. Figure 5 shows the results of the saliency maps for ANB angles subsampled from  $[6^\circ-7^\circ]$  to  $[9^\circ <]$  gradually increasing by steps of 1.

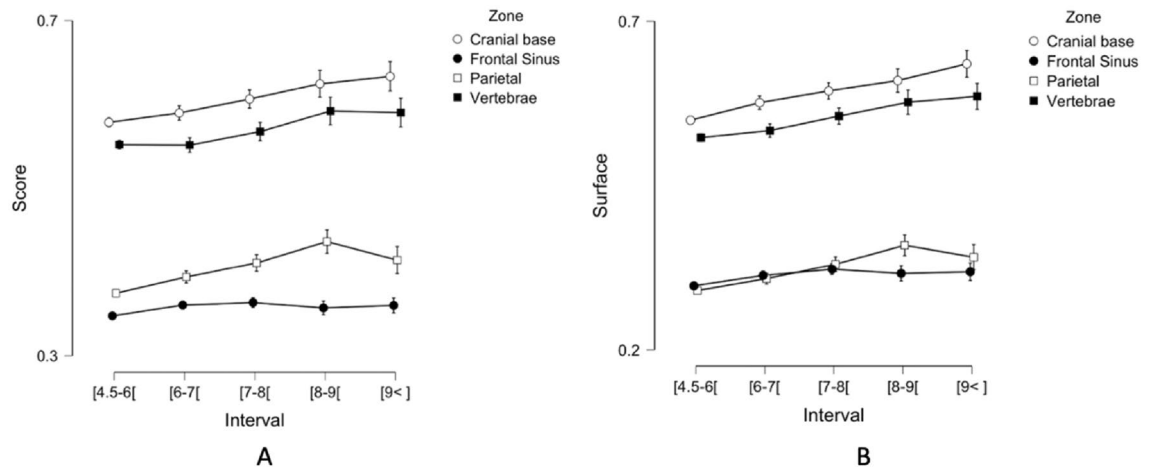
The analysis of the evolution of the heat map as a function of the severity of C2Rm offers the following results:

The first map (Fig. 6A), shows a global activation the maximum intensity of which is distributed over the bone structures in warm colors (i.e. orange). More precisely, the lower part of the cervical segment, the upper segment up to the sinus, and the mandibular halo appear in orange-red intensity. The rest of the vertebrae and the cranial base, as well as the parietal bone, are shown in yellow. The second map (Fig. 5B) shows an amplification of the intensity of the areas, which get more intense towards the red area. The cervical and basilar segments are now completely red. The diameter of the mandibular halo remains the same, but the red area is intensified. The third map (Fig. 5C) shows the thickening of the cervical-basilar segments and the activation of the inflection point



**Figure 5.** Evolution of the activation map according to the severity of class II. The color bar indicates the class score of the area.





**Figure 6.** Evolution of metrics according to the ANB angle (Interval) for each of the four skeletal structures (cranial base, frontal sinus, parietal bone, cervical vertebrae): (A) Evolution of the class score, (B) Evolution of the hot surface.

(sphenoid). The mandibular halo continues to extend into the maxilla. The parietal bone changes from yellow to orange. The fourth map presented (Fig. 5D) is characterized by the activation of the parietal bone (orange to red) and the thickening of all other red areas.

Globally, when the amplitude of the angle ANB increases, the results highlight a kinetics of increase in the intensity of the pixels and an expansion of the areas highlighted by the global activation maps.

#### Quantitative analysis: statistical results analysis

To assess the impact of increasing ANB angle (retrognathia severity) on saliency maps, we used two metrics: “class score” and “hot surface” (see point 2–6 of the methodology section). Figure 6 presents the results obtained for the four anatomical structures identified above (cranial base, frontal sinus, parietal bone, cervical vertebrae). We performed two 2-way analyses of variance (ANOVA) with ANB angle as between-subjects factor and anatomical structures as within-subjects factor. These two 2-way ANOVAs were conducted with the class score and the hot pixel area as the respective dependent variables. Post hoc comparison has been carried out by using Bonferroni test.

The analyses performed showed a significant effect of the angle factor on both class score  $F(3, 13,992) = p < 0.001$  and hot surface  $F(3, 13,992) = p < 0.001$  with:

- An increase in class score along the angle:  $\bar{x}([4.5^{\circ}-6^{\circ}[) = 0.464$ ,  $\bar{x}([6^{\circ}-7^{\circ}[) = 0.474$ ,  $\bar{x}([7^{\circ}-8^{\circ}[) = 0.487$ ,  $\bar{x}([8^{\circ}-9^{\circ}[) = 0.503$ ,  $\bar{x}(]9^{\circ}) = 0.500$ ;
- An increase in hot surface:  $\bar{x}([4.5^{\circ}-6^{\circ}[) = 0.416$ ,  $\bar{x}([6^{\circ}-7^{\circ}[) = 0.433$ ,  $\bar{x}([7^{\circ}-8^{\circ}[) = 0.451$ ,  $\bar{x}([8^{\circ}-9^{\circ}[) = 0.466$ ,  $\bar{x}(]9^{\circ}) = 0.471$ .

The analyses also showed a significant effect of the anatomical structures factor on both the class score  $F(3, 13,992) = p < 0.001$  and hot surface  $F(3, 13,992) = p < 0.001$  with:

- A higher class score for the cranial base and the cervical vertebrae ( $\bar{x}(\text{Cranial Base}) = 0.588$ ,  $\bar{x}(\text{Vertebrae}) = 0.557$ ) than for the parietal bone and the frontal sinus ( $\bar{x}(\text{Frontal Sinus}) = 0.353$ ,  $\bar{x}(\text{Parietal}) = 0.387$ );
- A larger hot surface for the cranial base and the cervical vertebrae ( $\bar{x}(\text{Cranial base}) = 0.566$ ,  $\bar{x}(\text{Vertebrae}) = 0.534$ ) than for the frontal sinus and the parietal bone ( $\bar{x}(\text{Frontal sinus}) = 0.305$ ,  $\bar{x}(\text{Parietal}) = 0.304$ ).

No interaction effect was observed for class score  $F(12, 13,992) = p = 0.379$ , nor for hot surface  $F(12, 13,992) = p = 0.530$ .

The lack of significance of the interaction is caused by the fact that the evolution as a function of the angle ANB is the same for the 4 structures.

#### Discussion

Recently, AI and, in particular, Machine Learning have come up with new tools that can help detect relevant features in images, bypassing the limitations of geometric morphometry. Moreover, interpretability methods based on the study of internal representations via saliency maps make AI accessible by explaining the elements used by the classifier to issue his prediction. In this study, we applied this innovative approach to mandibular retrognathia, a major health issue in the field of orthodontics.

In this context, we implemented an innovative interpretable AI methodology on orthodontic images by also proposing new elements such as extraction of average saliency maps (on all the individuals concerned) as well as quantification metrics of the ROIs identified by the CNN.

Our main objective was a double one. First, we intended to demonstrate that this methodology was appropriate for identifying image elements allowing a better understanding of the impact of retrognathia on the entire skull; second, we endeavoured to show that it made it possible to monitor the dynamic evolution of impacted cranial structures depending on the severity of retrognathia.

The results obtained are not only in line with these objectives—they also represent arguments in the debate on modifications to cranio-facial morphology during human evolution (linked to the progressive increase in retrognathia). Our discussion will focus on these different areas of contribution of our research.

### Interpretable AI methodology and orthodontic images

In the field of interpretable DNNs, the main contribution of our work is to propose a method of Morphological Information Extraction with Interpretable CNN (which we have called MIE-ICNN) based on a principle of global activation map. CNNs have already demonstrated their ability to extract very subtle features but often using complex and difficult to interpret models. As a result, recent publications<sup>49,50</sup> have insisted on the importance of focusing on the elements enabling CNNs to make their decisions. In this context, Wöber et al.<sup>20</sup> used two approaches [Layer-Wise Relevance Propagation<sup>51</sup>, and Grad-CAM<sup>52</sup>] to provide quantitative assessments at the pixel level about how different regions of a sample image contribute to the calculated predictions. We thus show that it was possible to go further by directly generating a global activation map of all pathological individuals to crush interindividual variability and highlight the biomarkers of a population.

The generation of a global map was made possible in particular by: (1) a preprocessing step, which consisted of aligning the images, avoiding a prediction based on noise due to rotation or offset, and (2) the superposition of the maps of individual salience, minimizing the impact of individual variability due to various parameters such as age, gender, ethnicity, etc.

Another contribution of our work made possible by the use of global activation maps is the ability to quantify the areas identified by the classifier using two complementary metrics (“classes score” and “hot surface”). This contribution is important because it offers a solution to the often qualitative approach of extracting relevant features from images. In the framework of our study, it permitted, for example, to carry out a statistical examination of the evolution of the structures involved in retrognathia, according to the severity of the malocclusion.

The methodology we propose offers perspectives on further studies, complementary to our present work. For example, as we have already indicated, numerous studies have highlighted a relationship between facial morphology and ethnicity<sup>6</sup>. Therefore, it is important to use normative cephalometric data specific to the patient's ethnicity, to ensure that the measurements are interpreted correctly and that the patient receives the most appropriate orthodontic treatment. Our study was carried out retrospectively using anonymized data which do not contain information on the ethnic origin of the patients. However, we can hypothesize that the dataset we used mainly represented a French population residing in France and receiving long-term treatment for retrognathia. In addition, our methodology which consists of superimposing individual saliency maps in order to eliminate individual variability has certainly contributed to reducing the impact of individuals who are differentiated by their ethnic origin. However, it would be of great interest to participate in the debate on the impact of ethnic origin on the shape of the skull by comparing global activation maps of individuals (suffering or not from mandibular retrognathia) from different ethnic groups. In the same way, progress in the field of three-dimensional (3D) surface imaging brings new perspectives to the field of orthodontics. Techniques such as surface laser scanning and stereophotogrammetry allow the description and comparison of 3D surfaces of the face, the study of soft tissue growth, the planning of therapy, and even the evaluation of the results of orthodontic treatment or the simulation of surgical treatments<sup>53</sup>. In recent years, three-dimensional (3D) CNNs have been applied to medical image analysis<sup>54</sup>. These methods could also constitute a promising perspective of our work insofar as they have already proved to be of interest in our scientific field. Indeed Kaźmierczak et al.<sup>55</sup> adopted this approach to obtain a good prediction of the direction of facial growth in patients with malocclusions, and consequently to provide a more personalised treatment and increased chances of success.

### Impact of retrognathia on the craniofacial morphology

In the field of orthodontic practice, our methodology has made it possible: (1) to identify the structures already known in literature as being impacted by retrognathia (frontal sinus, the base of the skull, the sphenoid and the cervical vertebrae); (2) to identify a structure that had not been identified by other methods (i.e., the parietal bone); (3) to quantify the dynamic evolution of the structures of the skulls involved according to the severity of retrognathia.

Many studies have sought to identify the impact of C2Rm on skull shape in order to predict mandibular growth pattern. Indeed, growth prediction plays an important role in the accurate diagnosis and treatment planning of orthodontic patients. As highlighted by Tehranchi et al.<sup>17</sup>, the growth pattern of the mandible, maxilla, and other craniofacial structures is critical in determining the time of onset, the duration and the prognosis of malocclusions. But in the absence of a methodology allowing a global approach, these studies were carried out using correlational methods aimed at identifying the correlations between the mandible and structures such as the cervical vertebrae<sup>18</sup>, the cranial base<sup>19</sup> or the frontal sinus<sup>17</sup>. Our study is, therefore, the first to use AI as a means to globally assess the structures impacted by retrognathia by circumventing the limitations inherent in the use of landmarks. The fact that we find the results already obtained using correlational methods argues in favor of the relevance of our approach aimed at identifying common characteristics of a population by the superposition of individual saliency maps. Similarly, the identification of a new structure (parietal bone) can be explained by

the fact that the cephalometric and Procrustes methods require a pre-definition of the structures to be measured, and that this specific structure had not been considered, a priori, to be relevant by experts.

In addition, in our study, the extraction of saliency maps according to the level of severity of the pathology made it possible to visualize the global dynamics of the evolution of the structures impacted by retrognathia. This behavior resulted in a significant “progressive inflammation” of the class score and an extension of the hot surface on a gravity-dependent axis including cervical vertebrae, the base of the skull and the frontal sinus. Moreover, the analyses showed that these kinetics were homogeneous for all the structures identified both for the class score and for the hot surface.

These results (taking into account the severity of retrognathia) must be considered with caution to the extent that the database did not allow us to identify the characteristics of the samples in terms of age distribution for each level of severity of retrognathia. However, a study conducted on school children aged 9 and 12 (in India) reached the conclusion that the severity of the malocclusions was higher in the 12-year-old age group than in the 9-year-old age group<sup>56</sup>. Despite these limitations (which have to be considered in future studies) our study confirmed an overall amplification of cranial dysmorphism (in terms of both intensity and amplitude) according to the severity of the pathology, and the involvement of frontal sinus and parietal bone from the early stages of the pathology. Further studies will be necessary to clarify the role of these two structures in mandibular retrognathia, but hypotheses can nevertheless be put forward. The frontal sinus receives the pressure of the occlusive forces through the pillars 3 (canine-canine) and constitutes a pneumatic space of adsorption. It, therefore, appears relevant that a modification of the occlusive forces as generated by the C2RM has repercussions on the sinus, an hypothesis already raised by Lieberman et al.<sup>57</sup>. The parietal bone is a point of convergence of the arcs of bio-mechanical forces of the skull. As a result, and as for the frontal sinus, it seems relevant that the alteration of the occlusive forces has an impact on its deployment and, consequently, on its morphology. Whatever the case may be, mandibular retrognathia sometimes requires, during childhood, a therapeutic intervention (for mandibular stimulation) which must necessarily take place during the peak of bone maturation. After this critical period, the only possible solution is surgery. It is, therefore, essential to be able to anticipate early (before the critical period) the need of a therapeutic intervention. In this context, the frontal sinuses and the parietal bone could constitute early predictive biomarkers of the evolution of retrognathia, and subsequent studies of the longitudinal type would be extremely relevant to test this hypothesis. It is also important to note that studying the interaction between craniofacial structures through the implementation of an integrated approach has been highlighted as a challenge by Lieberman et al.<sup>57</sup>. For these authors, “it is difficult to define or assess quantitative hypotheses of craniofacial integration because we know so little about the extent and nature of the many interactions likely to occur between cranial regions”. The MIE-ICNN approach we propose provides solutions to capture covariations involving different craniofacial structures and presents saliency maps as a new morphometric assessment tool that could be applied to many situations involving the extraction of relevant information in medical images.

### Parietal bone and modifications of craniofacial morphology during human evolution

According to our results, C2Rm, which is a pathology of the jaws, has an impact on the lower face (jaws and facial mass), but also on the cranial vault at the level of the parietal bone.

The frontal sinus (via the juxtaposed supraorbital torus) and, especially, the parietal bone are markers of paleoanthropological rusticity. The parietal bone has been identified as one of the key areas in the modification and gracilisation of the craniofacial architecture during evolution<sup>58</sup>. By demonstrating a link between malocclusion (retrognathia) and parietal bone, our study opens the way to reconsidering the masticatory theory through the hypothesis of a more extensive impact on craniofacial architecture than what is currently admitted (influence of retrognathia on changes in the shape of the skull, via the parietal bone).

Reconsidering the influence of mastication on the modification of the craniofacial phenotype during human evolution makes it possible to highlight the role of local modulating factors due to terrain and eating behaviors (such as changes in lifestyle, or in the way of preparing food, etc.). It also circumvents some of the explanatory limitations of classical models based on morphology or genetics that underestimate the role of local factors in establishing phylogenetic models.

Indeed, a progressive reduction of the mandible has taken place during human evolution, and this transition from prognathism to increasingly accentuated retrognathism has been accompanied by a series of craniofacial modifications<sup>41</sup>. For example, the passage from the Mesolithic to the Neolithic saw the emergence of a gracilisation of the craniofacial morphotype in the absence of any genetic modification. This passage was characterized by a transition from a hunter-gatherer to an agro-pastoralist lifestyle. The change in diet resulted in a sophistication of food preparation (grinding, cooking, soaking, leaching) and a much smoother texture than in the era of hunter-gatherers. A gracile morphotype appeared, mainly characterized by a flattening of the face and an increase in cranial capacities. The concomitance of gracile morphotype and dietary change had already led some authors to put forward the hypothesis of a masticatory influence, in particular on the projection of the face, but this hypothesis was not sufficient to account for changes in craniofacial morphology.

Our results shed new light on the debate on the development of cranial shape during hominin evolution. They could indeed provide more complete explanatory hypotheses (involving the parietal bone) on links between the increase in retrognathia and cranio-facial modifications during the evolution of hominins. However, these results are still to be regarded with caution and should give rise to further investigations, given that we lack information on the average age of the two groups considered (Class I and Class II) and that the identification of the parietal bone could be accentuated by an age-related effect.

## Conclusion

Even if the approach we propose seems promising for future investigations centered on medical images, our study is subject to significant limitations, which must be duly considered and discussed. The retrospective design may introduce biases, especially since half of the initial dataset was rejected to allow an optimal extraction of mean saliency maps. In addition, the data we used was anonymized and did not contain information on the ethnic origin of patients. However, as already mentioned, ethnic origin has a significant impact on the shape of the skull, and this information must be taken into account when carrying out a diagnosis of retrognathia and implementing the appropriate treatment. Similarly, our anonymized database did not contain information on sample characteristics in terms of age and sex, variables that may also affect retrognathia. Even if our methodology using an overlay technique to extract an average saliency map helped to reduce the impact of these different factors on variations in skull shape, these factors should be taken into account in subsequent studies in order to improve the predictive value of the results. Finally, our study was conducted on a relatively small sample, especially with regard to the effect of the severity of retrognathia, which means that its results may not be generalizable to a larger population. Further studies will be necessary to circumvent these problems and to define more precisely the benefits and limits as well as the fields of application of these innovative approaches based on deep learning and interpretable AI.

## Data availability

Due to the medical nature of the research, data are not publicly available, in compliance with the French Data Protection Act. The framework created during the study can be made available on reasonable request by contacting the corresponding author.

Received: 2 February 2023; Accepted: 18 October 2023

Published online: 24 October 2023

## References

- Patti, A. *Traitement des classes II. De la prévention à la chirurgie—Antonio Patti* (Quintessence international) (2010). <https://www.decitre.fr/livres/traitement-des-classes-ii-9782912550668.html>
- Proffit, W. R., Fields, H. W. & Moray, L. J. Prevalence of malocclusion and orthodontic treatment need in the United States: Estimates from the NHANES III survey. *Int. J. Adult Orthod. Orthognath. Surg.* **13**(2), 97–106 (1998).
- Darqu , J. La Classe II, division 2. *Rev Orthop. Dento-Fac.* **8**(1), 5–55 (1974).
- Sonnesen, L., Nolting, D., Engel, U. & Kjaer, I. Cervical vertebrae, cranial base, and mandibular retrognathia in human triploid fetuses. *Am. J. Med. Genet. Part A* **149A**(2), 177–187. <https://doi.org/10.1002/ajmg.a.32631> (2009).
- Anshuka, A., Vaswani, V. & Khajuria, S. Assessment and comparison of cervical column morphology and cranial base angle in three different facial types—A cephalometric study. *J. Evol. Med. Dental Sci.* **9**, 2605–2609. <https://doi.org/10.14260/jemds/2020/567> (2020).
- Darkwah, W. K., Kadri, A., Adormaa, B. B. & Aidoo, G. Cephalometric study of the relationship between facial morphology and ethnicity: Review article. *Transl. Res. Anat.* **12**, 20–24. <https://doi.org/10.1016/j.tria.2018.07.001> (2018).
- Ahsan, A., Yamaki, M., Hossain, Z. & Saito, I. Craniofacial cephalometric analysis of Bangladeshi and Japanese adults with normal occlusion and balanced faces: A comparative study. *J. Orthod. Sci.* **2**(1), 7–15. <https://doi.org/10.4103/2278-0203.110327> (2013).
- Proffit, W. R., Fields, H. W., Sarver, D. M. & Ackerman, J. L. *Cephalometric Norms for Americans aged 10–18 years: The University of Michigan cephalometric Growth Study* (University of Michigan, 1993).
- Cardini, A. & Polly, P. D. Larger mammals have longer faces because of size-related constraints on skull form. *Nat. Commun.* **4**(1), 2458. <https://doi.org/10.1038/ncomms3458> (2013).
- Ledogar, J. A. *et al.* Mechanical evidence that Australopithecus sediba was limited in its ability to eat hard foods. *Nat. Commun.* **7**(1), 10596. <https://doi.org/10.1038/ncomms10596> (2016).
- Puech, P.-F., Albertini, H. & Tyrand, H. Correlated dento-facial progression and the origin of man. *Anthropologie* (1962) **34**(1/2), 35–38 (1996).
- Spoor, F., O’Higgins, P., Dean, C. & Lieberman, D. E. Anterior sphenoid in modern humans. *Nature* **397**(6720), 572. <https://doi.org/10.1038/17505> (1999).
- Bastir, M., Rosas, A. & Sheets, H. D. The morphological integration of the hominoid skull: A partial least squares and PC analysis with implications for European middle pleistocene mandibular variation. In *Modern Morphometrics in Physical Anthropology* (ed. Slice, D. E.) 265–284 (Kluwer Academic Publishers-Plenum Publishers, 2005). [https://doi.org/10.1007/0-387-27614-9\\_12](https://doi.org/10.1007/0-387-27614-9_12).
- von Cramon-Taubadel, N. The microevolution of modern human cranial variation: Implications for hominin and primate evolution. *Ann. Hum. Biol.* **41**(4), 4. <https://doi.org/10.3109/03014460.2014.911350> (2014).
- Gazit, E. *et al.* Prevalence of mandibular dysfunction in 10–18 year old Israeli schoolchildren. *J. Oral Rehabil.* **11**(4), 4. <https://doi.org/10.1111/j.1365-2842.1984.tb00581.x> (1984).
- Katz, M. I. Angle classification revisited 1: Is current use reliable?. *Am. J. Orthod. Dentofac. Orthop.* **102**(2), 2. [https://doi.org/10.1016/0889-5406\(92\)70030-E](https://doi.org/10.1016/0889-5406(92)70030-E) (1992).
- Tehranchi, A., Motamedian, S. R., Saedi, S., Kabiri, S. & Shidfar, S. Correlation between frontal sinus dimensions and cephalometric indices: A cross-sectional study. *Eur. J. Dent.* **11**(01), 01. <https://doi.org/10.4103/1305-7456.202630> (2017).
- Gomes, L. D. C. R., Horta, K. O. C., Goncalves, J. R. & Santos-Pinto, A. D. Systematic review Craniocervical posture and craniofacial morphology. *Eur. J. Orthod.* **36**(1), 1. <https://doi.org/10.1093/ejo/cjt004> (2014).
- de Almeida, K. C. M., Raveli, T. B., Vieira, C. I. V., dos Santos-Pinto, A. & Raveli, D. B. Influence of the cranial base flexion on Class I, II and III malocclusions: A systematic review. *Dental Press J. Orthod.* **22**(5), 5. <https://doi.org/10.1590/2177-6709.22.5.056-066.oar> (2017).
- W ber, W. *et al.* Identifying geographically differentiated features of Ethiopian Nile tilapia (*Oreochromis niloticus*) morphology with machine learning. *PLoS ONE* **16**(4), 4. <https://doi.org/10.1371/journal.pone.0249593> (2021).
- Slice, D. Geometric morphometrics. *Ann. Rev. Anthropol.* <https://doi.org/10.1146/annurev.anthro.34.081804.120613> (2007).
- Ross, A. Procrustes analysis. *Course report, Department of Computer Science and Engineering, University of South Carolina*, 26, 1–8 (2004).
- Adams, D. C., Rohlf, F. J. & Slice, D. E. Geometric morphometrics: Ten years of progress following the ‘revolution’. *Ital. J. Zool.* **71**(1), 1. <https://doi.org/10.1080/11250000409356545> (2004).
- Rohlf, F. J. Procrustes problems. *J. Am. Stat. Assoc.* **100**(471), 1097–1097. <https://doi.org/10.1198/jasa.2005.s40> (2005).
- Walker, J. A. Ability of geometric morphometric methods to estimate a known covariance matrix. *Syst. Biol.* **49**(4), 686–696. <https://doi.org/10.1080/106351500750049770> (2000).

26. Gkantidis, N. & Halazonetis, D. J. Morphological integration between the cranial base and the face in children and adults: Cranial base and face integration. *J. Anat.* **218**(4), 426–438. <https://doi.org/10.1111/j.1469-7580.2011.01346.x> (2011).
27. Freudenthaler, J., Čelar, A., Ritt, C. & Mitteröcker, P. Geometric morphometrics of different malocclusions in lateral skull radiographs. *J. Orofac. Orthop. Fortschr. Kieferorthopädie* **78**(1), 11–20. <https://doi.org/10.1007/s00056-016-0057-x> (2017).
28. Lieberman, D. E., Pearson, O. M. & Mowbray, K. M. Basicranial influence on overall cranial shape. *J. Hum. Evol.* **38**(2), 291–315. <https://doi.org/10.1006/jhev.1999.0335> (2000).
29. Thurzo, A., Strunga, M., Urban, R., Surovková, J. & Afrashtehfar, K. I. Impact of artificial intelligence on dental education: A review and guide for curriculum update. *Educ. Sci.* **13**(2), 2. <https://doi.org/10.3390/educsci13020150> (2023).
30. Wang, H., Wang, Z., Du, M., Yang, F., Zhang, Z., Ding, S., Mardziel, P., & Hu, X. Score-CAM : Score-Weighted Visual Explanations for Convolutional Neural Networks. [arXiv:1910.01279](https://arxiv.org/abs/1910.01279) [cs]. <http://arxiv.org/abs/1910.01279> (2020).
31. Kim, I., Rajaraman, S. & Antani, S. Visual interpretation of convolutional neural network predictions in classifying medical image modalities. *Diagnostics* **9**(2), 2. <https://doi.org/10.3390/diagnostics9020038> (2019).
32. Daanouni, O., Cherradi, B. & Tmiri, A. Automatic detection of diabetic retinopathy using custom CNN and Grad-CAM. In *Advances on Smart and Soft Computing* Vol. 1188 (eds Saeed, F. et al.) 15–26 (Springer, 2021). [https://doi.org/10.1007/978-981-15-6048-4\\_2](https://doi.org/10.1007/978-981-15-6048-4_2).
33. Masud, M., Eldin Rashed, A. E. & Hossain, M. S. Convolutional neural network-based models for diagnosis of breast cancer. *Neural Comput. Appl.* **34**(14), 11383–11394. <https://doi.org/10.1007/s00521-020-05394-5> (2022).
34. Irie, R. et al. A novel deep learning approach with a 3D convolutional ladder network for differential diagnosis of idiopathic normal pressure hydrocephalus and Alzheimer's disease. *Magn. Reson. Med. Sci.* **19**(4), 351–358. <https://doi.org/10.2463/mrms.mp.2019-0106> (2020).
35. Liu, X. et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *Lancet Digit. Health* **1**(6), e271–e297. [https://doi.org/10.1016/S2589-7500\(19\)30123-2](https://doi.org/10.1016/S2589-7500(19)30123-2) (2019).
36. Diao, J. A. et al. Human-interpretable image features derived from densely mapped cancer pathology slides predict diverse molecular phenotypes. *Nat. Commun.* **12**(1), 1613. <https://doi.org/10.1038/s41467-021-21896-9> (2021).
37. Lewis, J. E. & Kemp, M. L. Integration of machine learning and genome-scale metabolic modeling identifies multi-omics biomarkers for radiation resistance. *Nat. Commun.* **12**(1), 2700. <https://doi.org/10.1038/s41467-021-22989-1> (2021).
38. Finlayson, S. G., Chung, H. W., Kohane, I. S., & Beam, A. L. *Adversarial Attacks Against Medical Deep Learning Systems* (2019). [arXiv:1804.05296](https://arxiv.org/abs/1804.05296). <https://doi.org/10.48550/arXiv.1804.05296>.
39. Silva, C. & Ferreira, A. P. Frankfort plane vs. natural head posture in cephalometric diagnosis. *Dent. Med. Probl.* **40**(1), 129–134 (2003).
40. Jacobson, A., & Jacobson, R. L. *Radiographic cephalometry : From basics to 3-D imaging* (2nd ed) (Quintessence, 2006).
41. Makaremi, M., Lacaule, C. & Mohammad-Djafari, A. Deep learning and artificial intelligence for the determination of the cervical vertebra maturation degree from lateral radiography. *Entropy* **21**(12), 12. <https://doi.org/10.3390/e21121222> (2019).
42. Dawson, C. W. & Wilby, R. An artificial neural network approach to rainfall-runoff modelling. *Hydrol. Sci. J.* **43**(1), 1. <https://doi.org/10.1080/02626669809492102> (1998).
43. Gao, W., Zhang, X., Gilpin, L. H., & Liu, H. An improved Sobel edge detection. In *2010 3rd International Conference on Computer Science and Information Technology* 67–71 (2010). <https://doi.org/10.1109/ICCSIT.2010.5563693>
44. Bengio, Y. & Delalleau, O. On the expressive power of deep architectures. *Discov. Sci.* [https://doi.org/10.1007/978-3-642-24477-3\\_1](https://doi.org/10.1007/978-3-642-24477-3_1) (2011).
45. Eldan, R., & Shamir, O. The power of depth for feedforward neural networks. In *Conference on Learning Theory*, 907–940 (2016). <https://proceedings.mlr.press/v49/eldan16.html>
46. He, J., Li, L., Xu, J. & Zheng, C. ReLU deep neural networks and linear finite elements. *J. Comput. Math.* **38**(3), 502–527. <https://doi.org/10.4208/jcm.1901-m2018-0160> (2020).
47. Kingma, D. P., & Ba, J. *Adam: A Method for Stochastic Optimization* (2017). [arXiv:1412.6980](https://arxiv.org/abs/1412.6980). <https://doi.org/10.48550/arXiv.1412.6980>
48. Zhang, X., Zhou, F., Lin, Y., & Zhang, S. Embedding label structures for fine-grained feature representation. [arXiv:1512.02895](https://arxiv.org/abs/1512.02895) [cs]. <https://arxiv.org/abs/1512.02895> (2016)
49. Montavon, G., Samek, W. & Müller, K.-R. Methods for interpreting and understanding deep neural networks. *Digit. Signal Process.* **73**, 1–15. <https://doi.org/10.1016/j.dsp.2017.10.011> (2018).
50. Selvaraju, R. R. et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Int. J. Comput. Vis.* **128**(2), 336–359. <https://doi.org/10.1007/s11263-019-01228-7> (2020).
51. Bach, S. et al. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS ONE* **10**(7), e0130140. <https://doi.org/10.1371/journal.pone.0130140> (2015).
52. Selvaraju, R. R., Das, A., Vedantam, R., Cogswell, M., Parikh, D., & Batra, D. Grad-CAM : Why did you say that? [arXiv:1611.07450](https://arxiv.org/abs/1611.07450) [cs, stat]. <http://arxiv.org/abs/1611.07450> (2017).
53. Erten, O. & Yilmaz, B. N. Three-dimensional imaging in orthodontics. *Turk. J. Orthod.* **31**(3), 86–94. <https://doi.org/10.5152/TurkJOrthod.2018.17041> (2018).
54. Singh, S. P. et al. 3D Deep learning on medical images: A review. *Sensors* **20**(18), 18. <https://doi.org/10.3390/s20185097> (2020).
55. Kaźmierczak, S. et al. Prediction of the facial growth direction is challenging. In *Neural Information Processing* (eds Mantoro, T. et al.) 665–673 (Springer, 2021). [https://doi.org/10.1007/978-3-030-92310-5\\_77](https://doi.org/10.1007/978-3-030-92310-5_77).
56. Chauhan, D., Sachdev, V., Chauhan, T. & Gupta, K. A study of malocclusion and orthodontic treatment needs according to dental aesthetic index among school children of a hilly state of India. *J. Int. Soc. Prev. Community Dent.* **3**(1), 32–37. <https://doi.org/10.4103/2231-0762.115706> (2013).
57. Lieberman, D. E., Ross, C. F. & Ravosa, M. J. The primate cranial base: Ontogeny, function, and integration: Primate Cranial Base. *Am. J. Phys. Anthropol.* **113**(S31), 117–169. [https://doi.org/10.1002/1096-8644\(2000\)43:31+%3c117::AID-AJPA5%3e3.0.CO;2-1](https://doi.org/10.1002/1096-8644(2000)43:31+%3c117::AID-AJPA5%3e3.0.CO;2-1) (2000).
58. Lieberman, D. E. Sphenoid shortening and the evolution of modern human cranial shape. *Nature* **393**(6681), 6681. <https://doi.org/10.1038/30227> (1998).

## Author contributions

M.M.: came up with the idea of research, participated in the collection of data and their labeling, bibliographic research and article writing, team coordination A.V.S.: implemented the neural networks architecture, interpretation of results and article writing B.M.: participated in the labeling of data, bibliographic research, and article writing I.C.K.: participated in the bibliographic research and the writing of the article A.M.D.: article writing S.S.: collection of data, bibliographic research F.B.: Article writing B.N.: Article writing, team coordination.

## Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to M.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023